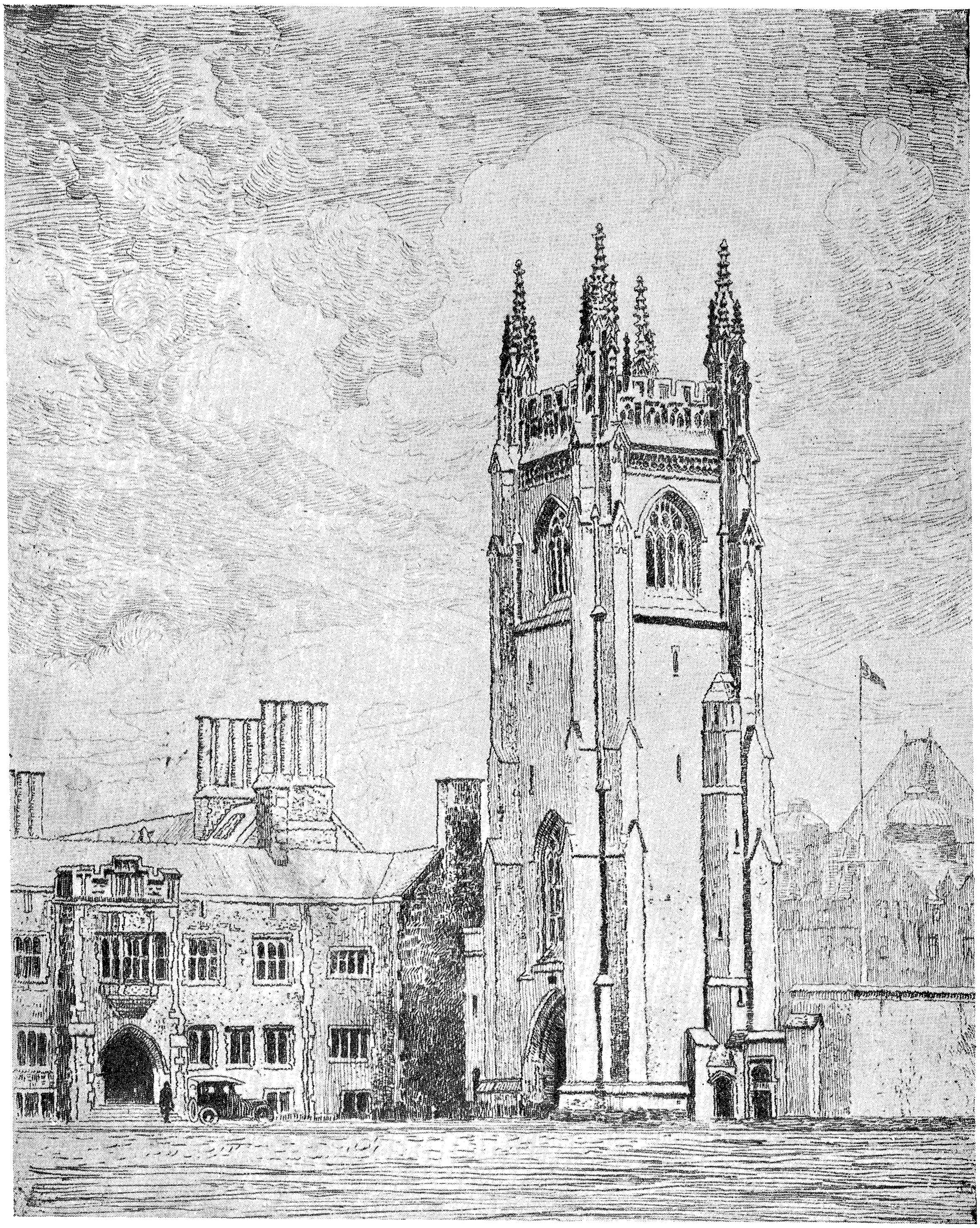


Memorial Tower, University of Toronto

Owen Staples

PROCEEDINGS

OF THE

INTERNATIONAL MATHEMATICAL CONGRESS

TORONTO, 1924

PHYSICS BUILDING—LECTURES. CONVOCATION HALL—OPENING AND CLOSING MEETINGS OF THE CONGRESS

PROCEEDINGS

OF THE

INTERNATIONAL MATHEMATICAL CONGRESS

HELD IN

TORONTO, AUGUST 11-16, 1924

EDITED BY

J. C. FIELDS

RESEARCH PROFESSOR OF MATHEMATICS IN THE UNIVERSITY OF TORONTO

WITH THE COLLABORATION OF AN EDITORIAL COMMITTEE

VOL. II

COMMUNICATIONS TO SECTIONS III, IV, V AND VI.

TORONTO:
THE UNIVERSITY OF TORONTO PRESS
1928

TORONTO:
THE UNIVERSITY OF TORONTO PRESS

PRINTED IN CANADA

The Congress met in Toronto by invitation of the University of Toronto and the Royal Canadian Institute, its sessions being held in the buildings of the University. In its organization and the conduct of its proceedings it conformed to the regulations of the International Research Council, and the International Mathematical Union.

COMMUNICATIONS

SECTION III

MECHANICS, PHYSICS, ASTRONOMY, GEOPHYSICS

SUR LA STABILITÉ ORDINAIRE DES ELLIPSOÏDES DE JACOBI

PAR M. ÉLIE CARTAN,
Professeur à la Sorbonne, Paris, France.

1. Comme on sait, c'est dans son mémoire célèbre: *Sur l'équilibre d'une masse fluide animée d'un mouvement de rotation** que H. Poincaré a abordé l'étude de la stabilité *ordinaire* des ellipsoïdes de Maclaurin et de Jacobi en la ramenant à l'étude de leurs petits mouvements fondamentaux. Il a montré que, alors que la stabilité *séculaire* des ellipsoïdes de Maclaurin cesse à l'ellipsoïde de bifurcation de Jacobi, la stabilité ordinaire s'étend encore au-delà, c'est-à-dire à des ellipsoïdes plus aplatis. Quant à la stabilité ordinaire des ellipsoïdes de Jacobi, H. Poincaré s'est contenté de déclarer comme probable une propriété analogue, à savoir que la stabilité ordinaire s'étend au-delà de l'ellipsoïde de bifurcation sur lequel s'embranchent les figures piriformes (ellipsoïde piriforme).

J'ai montré récemment† que la stabilité ordinaire s'étendait à tous les ellipsoïdes de Maclaurin dont l'aplatissement est inférieur à 0.697 (ou l'excentricité inférieure à 0.953): c'est la limite qui avait déjà été indiquée par B. Riemann‡, mais dans l'hypothèse restreinte où la masse fluide conservait une forme ellipsoïdale§.

La question de la stabilité ordinaire des ellipsoïdes de Jacobi n'a pas, à ma connaissance, été résolue jusqu'à présent. Or, il se trouve, chose curieuse, que sa solution est très simple. Contrairement à ce que pensait H. Poincaré, *la stabilité ordinaire cesse, pour les ellipsoïdes de Jacobi, en même temps que la stabilité séculaire*, c'est-à-dire pour l'ellipsoïde piriforme.

2. La propriété qui vient d'être indiquée est un cas particulier d'un théorème plus général. Imaginons la masse fluide soumise à des liaisons l'obligeant à ne subir que des déformations d'ordre donné n: cela signifie que la déformation de sa surface libre peut s'exprimer au moyen des seuls polynomes de Lamé d'ordre n. On sait qu'elle comporte alors $2n+1$ coefficients de stabilité, dont $2n$ sont toujours positifs, le dernier seul étant susceptible de s'annuler|| (nous l'appellerons

*Acta Mathematica, t. 7 (1885), p. 259-380.

†Bull. Sci. Math., t. 46 (1922), p. 332-342.

‡Göttingen, Abhandlungen, t. 9 (1860), p. 3.

§Dans un mémoire des Phil. Trans., t. 180 (1888), p. 187, M. Bryan, étudiant en détail les petits mouvements des ellipsoïdes de Maclaurin, regardait déjà comme très probable la validité générale de cette limite, mais sans pouvoir en donner une démonstration.

||Le cas $n=2$ fait exception, l'un des coefficients de stabilité étant toujours négatif. Ce qui suit ne s'applique pas à ce cas. La stabilité séculaire, et par suite ordinaire, d'ordre 2, est toujours assurée.

le coefficient de stabilité *caractéristique* d'ordre n); on sait aussi qu'il y a un seul ellipsoïde de bifurcation d'ordre n.

Cela posé, on sait que la stabilité séculaire d'ordre n est assurée quand, partant de l'ellipsoïde de Jacobi de révolution, on suit la série de Jacobi jusqu'à l'ellipsoïde de bifurcation d'ordre n; cette stabilité cesse aussitôt après.

Le théorème annoncé et qui concerne la stabilité *ordinaire* est le suivant:

Si n est impair, la stabilité ordinaire d'ordre n cesse en même temps que la stabilité séculaire;

Si n est pair, la stabilité ordinaire d'ordre n se maintient encore pendant un certain temps après que la stabilité séculaire a disparu.

Autrement dit, l'ellipsoïde de bifurcation d'ordre n est *critique* ou *non critique* pour la stabilité ordinaire suivant que n est impair ou pair.

3. Avant de démontrer ce théorème, il ne sera pas inutile de signaler ce qu'il a de paradoxal, quand on le compare aux résultats classiques relatifs à la stabilité ordinaire d'un système matériel n'ayant qu'un nombre fini r de degrés de liberté, au voisinage d'une position d'équilibre relatif. On sait que, si l'on ne tenait compte que des forces directement appliquées au système (forces supposées dérivées d'un potentiel) et des forces d'inertie d'entraînement (forces centrifuges), la condition nécessaire et suffisante de stabilité serait que les r coefficients de stabilité fussent positifs. Mais si l'on tient compte, comme c'est nécessaire, des forces centrifuges composées, la stabilité peut avoir lieu sans que cette condition soit remplie; *néanmoins, il est nécessaire que les coefficients de stabilité négatifs soient en nombre pair.*

Dans le cas d'une masse fluide au voisinage d'une position ellipsoïdale d'équilibre relatif, assujettie à n'avoir que des déformations d'ordre n, le nombre des degrés de liberté est infini, mais il n'y a qu'un nombre fini, à savoir $2n+1$, de coefficients de stabilité, et *l'un au plus* de ces coefficients de stabilité est susceptible de devenir négatif. Il semblerait donc, par extension naturelle du cas d'un nombre fini au cas d'un nombre infini de degrés de liberté, que la stabilité ordinaire d'ordre n des ellipsoïdes de Jacobi devrait toujours cesser à l'ellipsoïde de bifurcation*; or, le théorème énoncé plus haut montre que cette conclusion n'est exacte que pour n impair, elle est fausse pour n pair.

4. Avant d'entrer dans le détail de la démonstration du théorème annoncé, résumons-en les grandes lignes. L'équation qui donne les fréquences fondamentales λ d'ordre n est de degré n^2+4n+1; mais comme les racines de cette équation sont deux à deux égales et opposées, λ est en facteur dans le premier membre si n est pair, et la racine $\lambda=0$ est *simple*, même si le coefficient de stabilité caractéristique est nul; par suite, les racines restent réelles pour des valeurs négatives suffisamment petites de ce coefficient de stabilité; au contraire, pour n impair, l'équation en λ n'admet pas de racine nulle, et le produit des racines de cette équation est, à un facteur près qui ne s'annule jamais, égal au produit des $2n+1$

*On sait (v. par. ex. le mémoire cité plus haut de M. Bryan) qu'aucun ellipsoïde de bifurcation de Maclaurin n'est critique pour la stabilité ordinaire ,mais cela n'est pas en contradiction avec le théorème relatif aux systèmes à un nombre fini de degrés de liberté, parce que les coefficients de stabilité qui s'annulent sont *doubles*.

coefficients de stabilité; il en résulte que le coefficient de stabilité caractéristique devenant négatif, le produit des $\frac{n^2+4n+1}{2}$ valeurs de λ^2 devient négatif, et, par suite, les racines ne peuvent rester toutes réelles. On voit de plus que l'une des fréquences fondamentales s'annule pour l'ellipsoïde de bifurcation de l'ordre considéré.

5. Arrivons à la démonstration proprement dite du théorème annoncé. Les petits mouvements fondamentaux d'ordre inférieur ou égal à n s'obtiennent en intégrant les équations*:

$$(1)\quad \begin{cases} \dfrac{\partial\psi}{\partial x} = \lambda^2\xi + 2i\omega\lambda\eta, \\[2ex] \dfrac{\partial\psi}{\partial y} = \lambda^2\eta - 2i\omega\lambda\xi, \\[2ex] \dfrac{\partial\psi}{\partial z} = \lambda^2\zeta, \\[2ex] \dfrac{\partial\xi}{dx} + \dfrac{\partial\eta}{\partial y} + \dfrac{\partial\zeta}{\partial z} = 0, \end{cases}$$

dans lesquelles ψ, ξ, η, ζ sont des polynomes entiers en x, y, z, le premier de degré n, les trois derniers de degré $n-1$, avec les conditions aux limites

$$(2)\qquad \int P_k\psi l d\sigma = 2\pi f\rho(H_3 - H_k)\int P_k\left(\frac{x}{a^2}\xi + \frac{y}{b^2}\eta + \frac{z}{c^2}\zeta\right) l d\sigma.$$

Ces intégrales sont étendues à la surface de l'ellipsoïde; les P_k désignent les différents polynomes de Lamé d'ordre inférieur ou égal à n. En éliminant entre (1) et (2) les coefficients des polynomes ψ, ξ, η, ζ, on obtient l'équation algébrique entière en λ qui donne les fréquences fondamentales d'ordre inférieur ou égal à n. Les fréquences d'ordre n s'obtiendront, on le voit facilement, en éliminant les coefficients des termes *de plus haut degré* de ψ, ξ, η, ζ entre les équations (1) et les $2n+1$ équations (2) qui se rapportent aux polynomes de Lamé d'ordre n. Autrement dit, on obtient l'équation $D_n(\lambda)=0$ qui donne les fréquences fondamentales d'ordre n en exprimant la compatibilité des équations (1) et (2) où ψ, ξ, η, ζ sont des polynomes *homogènes* (le premier de degré n, les autres de degré $n-1$) et où on remplace P_k par les $2n+1$ polynomes de Lamé d'ordre n.

Pour faire l'élimination, voici comment on peut procéder. Posons $\psi = \lambda\phi$. Les équations deviennent:

*Je conserve les notations employées dans mon article du Bull. Sci. Math. cité plus haut. J'ai écrit H_k à la place de $\frac{R_k S_k}{2n+1}$ (notation de Poincaré); f est la constante de l'attraction universelle, ρ la densité de la masse fluide.

$$(1') \quad \begin{cases} \dfrac{\partial \phi}{\partial x} = \lambda \xi + 2i\omega\eta, \\ \dfrac{\partial \phi}{\partial y} = \lambda \eta - 2i\omega \xi, \\ \dfrac{\partial \phi}{\partial z} = \lambda \zeta, \\ \dfrac{\partial \xi}{\partial x} + \dfrac{\partial \eta}{\partial y} + \dfrac{\partial \zeta}{\partial z} = 0; \end{cases}$$

$$(2') \qquad \lambda \int P_k \phi l d\sigma = 2\pi f \rho (H_3 - H_k) \int P_k \left(\frac{x}{a^2}\xi + \frac{y}{b^2}\eta + \frac{z}{c^2}\zeta \right) l d\sigma.$$

Éliminons d'abord les coefficients du polynome ϕ. Les équations (1′) fourniront, en fonction linéaire des coefficients de ξ, η, ζ, et aussi en fonction linéaire de λ, les différents coefficients de ϕ. Mais, tandis que le coefficient de x^n (ou de y^n, ou de z^n) dans ϕ ne sera fourni qu'une fois, le coefficient d'un monôme tel que $x^\alpha y^\beta (\alpha, \beta \neq 0)$ sera fourni deux fois, et le coefficient d'un monôme tel que $x^\alpha y^\beta z^\gamma$ $(\alpha, \beta, \gamma \neq 0)$ sera fourni trois fois.

Én égalant les différentes valeurs obtenues pour un même coefficient de ϕ, on aura entre les coefficients de ξ, η, ζ, un nombre de relations égal à trois fois le nombre total des coefficients d'une forme binaire de degré $n-2$, augmenté du double du nombre des coefficients d'une forme ternaire de degré $n-3$. Le nombre de ces relations est

$$3(n-1) + 2\frac{(n-2)(n-1)}{2} = n^2 - 1,$$

et, dans chacune d'elles, le paramètre λ entre linéairement.

Les relations (2′), quand on y remplacera les coefficients de ϕ par leurs valeurs précédemment déterminées, contiendront évidemment λ au second degré.

En définitive, l'équation cherchée s'obtiendra en annulant le déterminant $D_n(\lambda)$ des relations obtenues entre les coefficients de ξ, η, ζ; or, ce déterminant possède n^2-1 lignes faisant intervenir λ au premier degré, $2n+1$ lignes le faisant intervenir au second degré, les autres lignes ne dépendant pas de λ. Par suite $D_n(\lambda)$ est au plus de degré

$$n^2 - 1 + 2(2n+1) = n^2 + 4n + 1.$$

6. Il est facile de montrer que ce degré est effectivement atteint. Dire en effet que le coefficient de λ^{n^2+4n+1} dans $D_n(\lambda)$ est nul, c'est dire que les relations précédemment formées entre les coefficients de ξ, η, ζ sont compatibles quand, dans chacune d'elles, on ne conserve que les termes de plus haut degré en λ. Cela revient à dire qu'il existe des polynomes χ, ξ, η, ζ satisfaisant aux relations:

$$(3)\qquad \begin{cases} \dfrac{\partial\chi}{\partial x}=\xi, \ \dfrac{\partial\chi}{\partial y}=\eta, \ \dfrac{\partial\chi}{\partial z}=\zeta, \\ \dfrac{\partial\xi}{\partial x}+\dfrac{\partial\eta}{\partial y}+\dfrac{\partial\zeta}{\partial z}=0, \\ \displaystyle\int P_k\chi l d\sigma=0; \end{cases}$$

on aurait donc un polynome harmonique homogène χ de degré n orthogonal à tous les polynomes de Lamé d'ordre n, ce qui est absurde. Le polynome $D_n(\lambda)$ est donc effectivement de degré n^2+4n+1.

Étudions maintenant le terme constant de $D_n(\lambda)$.

Chacune des $2n+1$ dernières lignes du déterminant $D_n(\lambda)$, quand on y fait $\lambda=0$, contient en facteur le coefficient de stabilité correspondant. On a donc

$$D_n(0)=M\prod_{k=1}^{2n+1}(H_3-H_k);$$

quant au coefficient M, voyons s'il peut être nul. Il faut et il suffit évidemment, pour cela, que les équations:

$$(4)\qquad \begin{cases} \dfrac{\partial\phi}{\partial x}=2i\omega\eta, \ \dfrac{\partial\phi}{\partial y}=-2i\omega\xi, \ \dfrac{\partial\phi}{\partial z}=0, \\ \dfrac{\partial\xi}{\partial x}+\dfrac{\partial\eta}{\partial y}+\dfrac{\partial\zeta}{\partial z}=0, \\ \displaystyle\int P_k\left(\frac{x}{a^2}\xi+\frac{y}{b^2}\eta+\frac{z}{c^2}\zeta\right)ld\sigma=0, \quad (k=1, 2, \ldots, 2n+1), \end{cases}$$

soient compatibles. On voit d'abord que les polynomes ϕ, ξ, η, ζ ne doivent pas dépendre de z. Les $2n+1$ dernières équations (4) expriment ensuite que le polynome du premier degré en z

$$-\frac{1}{2i\omega}\left(\frac{x}{a^2}\frac{\partial\phi}{\partial y}-\frac{y}{b^2}\frac{\partial\phi}{\partial x}\right)+\frac{z}{c^2}\zeta$$

doit se réduire, sur l'ellipsoïde, à une somme de polynomes de Lamé d'ordre inférieur à n; par suite, à un multiple près de $\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}+\dfrac{z^2}{c^2}-1$, il doit être identique à un polynome de degré inférieur à n; cela revient à dire qu'il doit être divisible par $\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}+\dfrac{z^2}{c^2}$, ce qui est impossible, à moins qu'il ne soit identiquement nul. On a donc $\zeta=0$ et ϕ ne doit dépendre que de $\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}$ et, par suite, être une puissance de cette expression. Cela exige que n soit pair, auquel cas on a, à un facteur constant près,

$$\phi=\left(\frac{x^2}{a^2}+\frac{y^2}{b^2}\right)^{\frac{n}{2}}.$$

Il résulte de cette analyse que si n est impair, le coefficient M est essentiellement différent de zéro, et, à ce facteur près, $D_n(0)$ est bien égal au produit des coefficients de stabilité d'ordre n, conformément à ce qui avait été annoncé au N° 4.

7. Si n est pair, le coefficient M est nul, puisque les équations (4) sont compatibles et l'équation $D_n(\lambda)=0$ admet la racine $\lambda=0$. Nous allons montrer que, pour l'ellipsoïde de bifurcation d'ordre n, cette racine est simple. En effet, pour cet ellipsoïde, le coefficient de stabilité caractéristique, que nous supposerons être H_3-H_{2n+1}, s'annule; il en résulte que, dans le déterminant $D_n(\lambda)$, la ligne qui se rapporte à la dernière des équations (2′) contient λ en facteur. Dire que $\frac{D_n(\lambda)}{\lambda}$ contient encore λ en facteur, c'est dire que les équations:

$$(5)\qquad \begin{cases} \dfrac{\partial\phi}{\partial x}=2i\omega\eta, \\ \dfrac{\partial\phi}{\partial y}=-2i\omega\xi, \\ \dfrac{\partial\phi}{\partial z}=0, \\ \dfrac{\partial\xi}{\partial x}+\dfrac{\partial\eta}{\partial y}+\dfrac{\partial\zeta}{\partial z}=0, \\ \displaystyle\int P_k\left(\frac{x}{a^2}\xi+\frac{y}{b^2}\eta+\frac{z}{c^2}\zeta\right)l\,d\sigma=0, \quad (k=1,\,2,\,\ldots,\,2n), \\ \displaystyle\int P_{2n+1}\phi l\,d\sigma=0 \end{cases}$$

sont compatibles. Par suite, on a, sur l'ellipsoïde, si l'on tient compte des équations (5) autres que la dernière,

$$(6)\qquad -\frac{1}{2i\omega}\left(\frac{x}{a^2}\frac{\partial\phi}{\partial y}-\frac{y}{b^2}\frac{\partial\phi}{\partial x}\right)+\frac{z}{c^2}\zeta=hP_{2n+1}+Q,$$

h étant une constante et Q un polynome de degré inférieur à n.

Considérons maintenant le polynome obtenu en remplaçant, dans le polynome de Legendre $L_n(t)$ d'ordre n, t^2 par $1-\frac{x^2}{a^2}-\frac{y^2}{b^2}$:

$$P'(x,\,y)=L_n\left(\sqrt{1-\frac{x^2}{a^2}-\frac{y^2}{b^2}}\right);$$

P' est, sur l'ellipsoïde, une combinaison linéaire de polynomes de Lamé d'ordre n, et il satisfait identiquement à l'équation

$$\frac{x}{a^2}\frac{\partial P'}{\partial y}-\frac{y}{b^2}\frac{\partial P'}{\partial x}=0.$$

Multiplions les deux membres de l'équation (6) par $P'ld\sigma$ et intégrons sur toute la surface de l'ellipsoïde; l'intégrale du premier membre sera nulle, en vertu de l'identité

$$\int\left[P'\left(\frac{x}{a^2}\frac{\partial\phi}{\partial y}-\frac{y}{b^2}\frac{\partial\phi}{\partial x}\right)+\phi\left(\frac{x}{a^2}\frac{\partial P'}{\partial y}-\frac{y}{b^2}\frac{\partial P'}{\partial x}\right)\right]ld\sigma=0;$$

on aura donc:

$$h\int P'P_{2n+1}ld\sigma=0.$$

Or, les deux polynomes P' et P_{2n+1} *ne sont pas* orthogonaux. En effet, considérons-les sur la sphère homographique de l'ellipsoïde. Ils se réduisent à des fonctions sphériques d'ordre n, P' devenant le polynome de Legendre de degré n. On sait que les fonctions sphériques orthogonales au polynome de Legendre s'annulent toutes aux pôles de la sphère; par suite le polynome P_{2n+1}, s'il était orthogonal à P', s'annulerait aux deux pôles de l'ellipsoïde ($x=y=0$, $z=\pm c$). Cela n'est pas puisque, comme on sait, on a

$$P_{2n+1}=\prod_{i=1}^{n/2}\left(\frac{x^2}{a^2+h_i}+\frac{y^2}{b^2+h_i}+\frac{z^2}{c^2+h_i}-1\right),\quad(-a^2<h_i<-b^2<-c^2),$$

$$P_{2n+1}(0,0,c)=\prod_{i=1}^{n/2}\frac{-h_i}{c^2+h_i}\neq 0.$$

Les polynomes P' et P_{2n+1} n'étant pas orthogonaux, la constante h est nulle; autrement dit, comme il a été montré au N° 6, le polynome $\frac{x}{a^2}\xi+\frac{y}{b^2}\eta+\frac{z}{c^2}\zeta$ est identiquement nul, et on a

$$\zeta=0,\quad \phi=k\left(\frac{x^2}{a^2}+\frac{y^2}{b^2}\right)^{\frac{n}{2}},$$

k étant un coefficient constant; on peut encore dire que sur l'ellipsoïde ϕ est de la forme kP', à un polynome près de degré inférieur à n. La dernière équation (5) donne alors

$$k\int P'P_{2n+1}ld\sigma=0,$$

et, par suite, $k=0$. Le système (5) est donc finalement incompatible, puisqu'il n'admet que la solution $\phi=\xi=\eta=\zeta=0$. C'est ce que nous voulions démontrer.

8. La racine $\lambda=0$ n'étant pas double lorsque le coefficient de stabilité caractéristique s'annule, est-il possible qu'une autre racine soit double, de manière à donner des racines imaginaires pour des valeurs négatives très petites de ce coefficient de stabilité? Nous allons montrer que cette dernière possibilité est exclue.

Considerons en effet une solution d'ordre n des équations (1) et (2), ψ étant un polynome de degré n (non nécessairement homogène), ξ, η, ζ des polynomes de degré $n-1$, les constantes de Fourier d'ordre inférieur à n de ψ et de $\frac{x}{a^2}\xi+\frac{y}{b^2}\eta+\frac{z}{c^2}\zeta$ étant toutes nulles.

Désignons par ξ_0, η_0, ζ_0 les polynomes imaginaires conjugués de ξ, η, ζ et supposons qu'on ait sur l'ellipsoïde

$$\frac{x}{a^2}\xi+\frac{y}{b^2}\eta+\frac{z}{c^2}\zeta=\sum_{k=1}^{2n+1}c_kP_k\ (x,\ y,\ z); \tag{6}$$

on aura également sur l'ellipsoïde

$$\psi=2\pi f\rho\sum c_k(H_3-H_k)P_k(x,\ y,\ z). \tag{7}$$

On a la formule

$$\int\psi\left(\frac{x}{a^2}\xi_0+\frac{y}{b^2}\eta_0+\frac{z}{c^2}\zeta_0\right)ld\sigma=\int\left(\xi_0\frac{\partial\psi}{\partial x}+\eta_0\frac{\partial\psi}{\partial y}+\zeta_0\frac{\partial\psi}{\partial z}\right)dxdydz, \tag{8}$$

l'intégrale du premier membre étant étendue à la surface de l'ellipsoïde, celle du second membre au volume de cet ellipsoïde. En tenant compte des formules (6) et (7) d'une part, (1) d'autre part, la relation (8) s'écrit

$$2\pi f\rho\sum(H_3-H_k)\ |c_k|^2\int P_k^2ld\sigma=\lambda^2\int(\xi\xi_0+\eta\eta_0+\zeta\zeta_0)dxdydz+\lambda\int 2i\omega(\xi_0\eta-\xi\eta_0)dxdydz,$$

ou encore

$$A\lambda^2+B\lambda-C=0, \tag{9}$$

les quantités A, B, C étant *réelles*, avec $A>0$*.

9. Cela posé, supposons que, pour des valeurs négatives suffisamment petites du coefficient de stabilité caractéristique, l'équation $D_n(\lambda)=0$ admette une racine imaginaire. Elle peut être suivie par continuité, en même temps que les polynomes ψ, ξ, η, ζ correspondants, jusqu'à l'ellipsoïde de bifurcation. A chaque instant, elle peut être considérée comme une des racines de l'équation du second degré à coefficients *réels* (9). Cela n'est possible que si l'on a

$$C<0,\ B^2+4AC<0.$$

Or, à la limite, la quantité C, d'après son expression même

$$C=2\pi f\rho\sum(H_3-H_k)\ |c_k|^2\int P_k^2ld\sigma$$

est *positive ou nulle*; il faut donc qu'à la limite, on ait $C=0$ et par suite $B=0$; mais alors la racine imaginaire considérée serait nulle à la limite, ce qui est contraire à l'hypothèse.

Il est donc démontré rigoureusement que *les ellipsoïdes de bifurcation d'ordre pair ne sont pas critiques pour la stabilité ordinaire du même ordre.*

10. Il reste à indiquer une interprétation physique de la racine nulle de l'équation qui donne les fréquences fondamentales d'ordre pair. Elle corres-

*La formule (9) démontre *a posteriori* que toutes les racines de l'équation $D_n(\lambda)=0$ sont réelles lorsque les coefficients de stabilité sont tous positifs.

pond, comme il est facile de le démontrer, à un petit mouvement *permanent* dans lequel chaque molécule est animée de la vitesse

$$u = k \frac{\partial P'}{\partial y}, \qquad v = -k \frac{\partial P'}{\partial x}, \qquad w = 0,$$

k étant une constante très petite et P' le polynome $L_n\left(\sqrt{1-\frac{x^2}{a^2}-\frac{y^2}{b^2}}\right)$. Chaque molécule décrit lentement une ellipse homothétique de l'ellipse équatoriale. Quant à la surface libre, elle conserve une forme invariable résultant d'une déformation d'ordre n de la surface de l'ellipsoïde. Ces petits mouvements, qui ne se ramènent pas aux mouvements fondamentaux de H. Poincaré, ont été depuis longtemps signalés pour $n=2$ par Dedekind*. Il n'existe pas, pour n impair, de petits mouvements permanents analogues.

*Journal für Mathematik, t. 58 (1861), p. 621.

SUR L'ARRIVÉE DANS LE SYSTÈME SOLAIRE D'UN ASTRE ÉTRANGER

PAR M. JEAN CHAZY,
Professeur à l'Université de Lille, Lille, France.

Dans un travail précédent*, j'ai étudié certaines trajectoires du problème des trois corps: je veux ici compléter les résultats obtenus dans ce travail.

J'ai démontré† que dans le problème des trois corps, quand le temps croît indéfiniment sur une trajectoire donnée dans un sens donné, chacune des trois distances mutuelles peut présenter seulement les quatre cas qui suivent: ou bien elle est infiniment grande d'ordre 1 par rapport au temps; ou bien elle est infiniment grande d'ordre $\frac{2}{3}$; ou bien elle est bornée supérieurement: tout comme le rayon vecteur du problème des deux corps; à l'exception peut-être de trajectoires, dont l'existence reste douteuse, et sur lesquelles indéfiniment, tantôt les trois distances mutuelles sont inférieures à une longueur fixe, tantôt l'une d'entre elles est inférieure à une longueur fixe, et les deux autres deviennent supérieures à toute longueur fixée à l'avance. La question de savoir si ces dernières trajectoires existent ou n'existent pas est la partie la plus célèbre de la question de la *stabilité dans le problème des trois corps*.

Ces résultats comportent le complément suivant. Représentons le mouvement des trois masses m_1, m_2, m_3 dans l'espace à douze dimensions, en prenant comme coordonnées d'un point de cet espace les trois coordonnées rectangulaires x, y, z, de la masse m_2 par rapport à la masse m_1, les trois coordonnées ξ, η, ζ de la masse m_3 par rapport au centre de gravité des masses m_1 et m_2, et les six dérivées par rapport au temps x', y', z', ξ', η', ζ'. Et posons:

*Annales de l'École Normale Supérieure, 3[e] série, t. 39, 1922, pp. 29-130: une proposition énoncée sans démonstration (p. 32, en note) est erronée, et rectifiée par les résultats énoncés ici.

†Nous écartons dans cette Communication les trajectoires aboutissant à un choc des trois corps, et situées dans l'espace à douze dimensions sur une multiplicité algébrique à neuf dimensions (puisque sur ces trajectoires les trois constantes des aires sont nulles). Les trajectoires conduisant à un choc de deux corps peuvent être continuées au-delà de ce choc par le prolongement analytique de M. Sundman: les résultats énoncés s'étendent aux trajectoires ainsi prolongées. Mais il est clair que les conséquences qu'on peut tirer de ces résultats ont d'autant moins de valeur pratique que les trois corps ou deux d'entre eux deviennent plus voisins, et que la loi de Newton et le prolongement analytique de M. Sundman représentent moins exactement les actions mutuelles et les mouvements de la réalité.

$$h=\frac{1}{2}\frac{m_1m_2}{m_1+m_2}Sx'^2+\frac{1}{2}\frac{(m_1+m_2)m_3}{m_1+m_2+m_3}S\xi'^2-\frac{m_2m_3}{r_{23}}-\frac{m_3m_1}{r_{31}}-\frac{m_1m_2}{r_{12}},$$

r_{23}, r_{31}, r_{12} désignant les trois distances mutuelles, et la constante de l'attraction universelle étant prise égale à l'unité.

Les trajectoires du problème des trois corps divisent l'espace à douze dimensions en cinq continua. Dans l'un, *le continuum extérieur,* situé dans la région $h>0$, chaque trajectoire suivie dans les deux sens est *hyperbolique:* les trois distances mutuelles y sont des infiniment grands d'ordre 1 par rapport au temps. Dans le *continuum intérieur,* situé dans la région $h<0$, dans chacun des deux sens où l'on peut suivre chaque trajectoire, ou bien les trois distances mutuelles sont bornées, ou bien cette trajectoire est de l'espèce exceptionnelle définie plus haut. Dans les trois autres continua, traversés chacun par la multiplicité algébrique $h=0$, chaque trajectoire suivie dans les deux sens est *hyperbolique-elliptique:* deux distances mutuelles sont des infiniment grands d'ordre 1, et la troisième, la même dans un même continuum, est bornée.

Les frontières de ces trois dernières continua, multiplicités analytiques à onze dimensions, sont engendrées, dans la région $h>0$ par les trajectoires hyperboliques-paraboliques, et dans la région $h<0$ par les trajectoires paraboliques-elliptiques: ces frontières se rejoignent deux à deux sur la multiplicité $h=0$ le long des trois multiplicités analytiques à dix dimensions engendrées par les trajectoires *paraboliques* (où les trois distances mutuelles sont des infiniment grands d'ordre $\frac{2}{3}$).

Les résultats précédents appellent différentes remarques. D'abord, au point de vue théorique, dans un problème de Dynamique, il est toujours important de relier les valeurs négatives très grandes et les valeurs positives très grandes du temps: qu'on se rappelle, par exemple, les efforts que Poincaré a consacrés à démontrer l'existence des solutions doublement asymptotiques dans le problème des trois corps.

En outre, par les propositions énoncées, nous restreignons la région de l'espace à douze dimensions où peuvent être situées les trajectoires dont l'existence est l'un des objets de la question de la stabilité dans le problème des trois corps: nous montrons que *cette région est une partie seulement de la région* $h<0$.

Enfin, au point de vue pratique, les propositions énoncées comportent les conséquences suivantes, dont l'application au système solaire est immédiate, avec une approximation supérieure aux approximations habituellement admises dans les hypothèses cosmogoniques.

Soit un système de deux corps, S et J, de masses quelconques, animés d'un mouvement elliptique d'excentricité quelconque et soit un troisième corps M, de masse quelconque, arrivant de l'infini dans une direction quelconque au voisinage des corps S et J.

Il est impossible que le corps M devienne satellite de l'un des corps S et J restant voisins, et plus généralement que le corps M reste indéfiniment au voisinage des corps S et J. C'est l'extension au cas général du problème des trois corps d'une proposition* démontrée par Schwarzschild dans le cas où la masse du corps M est nulle. Les conséquences qui suivent sont plus inattendues.

Il est impossible que le système des corps S et J soit disloqué par l'arrivée du corps M, et que les trois corps s'écartent indéfiniment dans des directions divergentes.

Il est impossible enfin que le corps M devienne satellite du corps S (ou J) et que le système de deux corps ainsi formé s'écarte indéfiniment du troisième corps J (ou S).

Après l'arrivée du corps M au voisinage des corps S et J, il ne peut se produire d'autre circonstance que la suivante: *le corps M reste au voisinage des corps S et J pendant un intervalle de temps plus ou moins long, mais fini, puis s'en éloigne indéfiniment:* les deux corps S et J restent voisins, les éléments osculateurs de leur mouvement relatif peuvent être plus ou moins profondément modifiés, mais restent ou du moins redeviennent elliptiques.

*Astronomische Nachrichten, Band 141, 1896, p. 7. La proposition de Schwarzschild comportait une exception possible pour des trajectoires dont les points forment un ensemble de mesure nulle dans tout volume fini de l'espace à six dimensions où est représenté le mouvement de la masse nulle; mais il y a continuité de l'allure finale des trajectoires en fonction des conditions initiales, et l'exception indiquée par Schwarzschild ne saurait se présenter, quelle que soit la valeur de la troisième masse.

LES RECHERCHES POSTHUMES DE LIAPOUNOFF SUR LES FIGURES D'ÉQUILIBRE D'UN LIQUIDE HÉTÉROGÈNE EN ROTATION

PAR M. WLADIMIR STEKLOFF,

Vice-Président de l'Académie des Sciences de Russie, Léningrad, Russie.

1. Dans la présente communication je me permets de soumettre à votre attention un bref résumé de l'œuvre posthume de notre illustre géomètre Liapounoff, portant le titre: *Sur certaines séries de figures d'équilibre d'un liquide hétérogène en rotation.*

Ces recherches ont été commencées en 1915 et achevées quelques jours avant la mort de l'illustre membre de notre Académie à Odessa, à la fin de 1919. Le manuscrit presque prêt pour l'impression n'a été remis entre mes mains qu'au commencement de 1922.

Jusqu'à ces derniers temps, l'Académie des Sciences de Russie n'a pas eu de moyens pour imprimer cette œuvre classique et ce n'est que dans un an que l'occasion se présentera de la faire paraître comme édition extraordinaire à l'occasion du 200^{e} anniversaire de notre Académie des Sciences qui sera célébré au commencement de l'été 1925.

Cette œuvre contiendra plus de 50 feuilles d'impression in 4°. Elle est rédigée d'une manière extrêmement concise, il n'y est donné que la marche générale des opérations analytiques qui amènent à la solution du problème; de nombreux détails, souvent intéressants, sont omis.

Les calculs détaillés, extrêmement compliqués, se rapportant à cette œuvre posthume, ainsi qu'aux quatre parties du travail *Sur les figures d'équilibre peu différentes des ellipsoïdes d'une masse liquide homogène douée d'un mouvement de rotation* (1906-1914), publié pendant la vie de l'auteur, soigneusement écrits et plusieurs fois vérifiés de différentes manières, remplissent près de 1000 feuilles de papier et sont gardés dans une armoire spéciale dans la Grande Salle de l'Académie. Dans les circonstances présentes il n'y a aucune possibilité de les publier.

2. L'œuvre posthume de Liapounoff, dont je parlerai, peut être divisée en trois parties.

Dans la première partie le problème est posé et sa possibilité est démontrée, sous certaines conditions. La première partie contient les §§ 1-23 (80 pages de manuscrit).

Le problème est posé comme il suit:

Soit E un ellipsoïde à demi-axes

$$\sqrt{\rho+1},\ \sqrt{\rho+q},\ \sqrt{\rho}$$

appartenant à la série, soit des ellipsoïdes de Jacobi (où $q<1$), soit des ellipsoïdes de Maclaurin (où $q=1$). Soit ω la vitesse qui lui correspond.

Trouver la figure F d'équilibre d'un liquide hétérogène, dont les particules s'attirent suivant la loi de Newton et qui tourne autour du même axe avec la même vitesse angulaire.

Supposons que les surfaces de niveau soient les ellipsoïdes homothétiques à l'ellipsoïde E et a un paramètre définissant les surfaces de niveau.

Supposons que la densité k du liquide s'exprime comme il suit:

$$k=1+\delta\phi(a), \tag{1}$$

δ étant un paramètre, supposé petit, et $\phi(a)$ une fonction donnée de a.

Liapounoff considère le problème dans la supposition que les surfaces de niveau de la figure cherchée F différent peu des ellipsoïdes homothétiques à l'ellipsoïde E et les définit par les équations

$$\begin{aligned} x&=a(1+\zeta)\sqrt{\rho+1}\ \sin\theta\cos\psi,\\ y&=a(1+\zeta)\sqrt{\rho+q}\ \sin\theta\sin\psi,\\ z&=a(1+\zeta)\sqrt{\rho}\cos\theta, \end{aligned} \tag{2}$$

ζ étant une fonction de a, θ et ψ, dont toutes les valeurs sont petites.

Le problème se ramène à la détermination de la fonction $\zeta(a,\theta,\psi)$. Faisant dans cette fonction $a=1$, nous obtiendrons

$$\bar{\zeta}=\zeta(1,\theta,\psi)$$

et les équations (2) donneront, pour $a=1$, l'équation de la surface de la figure F

Soit f la constante de Newton et

$$U=\int\frac{k'd\tau'}{r'} \tag{3}$$

le potentiel de la masse liquide.

L'équation fondamentale du problème sera

$$U+\Omega(x^2+y^2)=\text{fonc. de } a, \tag{A}$$

où

$$\Omega=\frac{\omega^2}{2f}.$$

Faisant $\delta=0$, on retrouve le cas du liquide homogène.

3. Pour résoudre le problème il faut non-seulement démontrer l'existence de la fonction ζ satisfaisant à l'équation (A), mais encore trouver son expression analytique valable pour $a=1$.

C'est un problème qui paraissait, il y a peu de temps, presque insoluble, même dans le cas du liquide homogène.

La solution approchée à l'aide de la méthode des approximations successives a été donnée pour la première fois par Poincaré; G. Darwin a poussé les calculs plus loin, jusqu'à l'approximation du second ordre. Liapounoff a résolu le problème dans toute sa généralité, dans son œuvre colossale citée plus haut.

Je dois faire remarquer qu'il a donné non-seulement le moyen de calculer les termes successifs du développement de la fonction cherchée, mais encore demontré la convergence des approximations successives.

Je saisis l'occasion pour remarquer encore que Liapounoff a étudié le problème de la stabilité des figures d'équilibre et a démontré rigoureusement que la figure piriforme est instable.

Il a établi que la stabilité dépend d'une quantité A, pour laquelle il a trouvé non-seulement une expression approchée, mais une équation algébrique exacte, à laquelle doit satisfaire cette quantité.

Il a trouvé les limites supérieure et inférieure de la racine A de cette équation et de cette manière a démontré en toute rigueur l'instabilité des figures piriformes (pear-shaped).

Neuf ans après (en 1917), M. Jeans, en appliquant la formule correspondante de Poincaré jusqu'à la troisième approximation, a retrouvé avec cette dernière approximation le résultat général de Liapounoff.

4. Il est vrai, cependant, comme le dit P. Appell, que l'appareil analytique nécessaire à la démonstration rigoureuse des résultats obtenus par Liapounoff est très compliqué, ce qui rend la lecture de ses ouvrages très difficile, mais cela dépend de la nature même du problème.

Il est d'autant plus difficile d'exposer dans une courte communication l'analyse encore beaucoup plus compliquée de l'œuvre posthume de Liapounoff concernant la théorie des figures de l'équilibre d'un liquide hétérogène.

Je regrette de ne pouvoir soumettre à votre attention les détails de cette analyse, aussi importante qu'ingénieuse, et je me bornerai à l'indication des principaux résultats.

5. Pour exclure le cas d'un liquide homogène, Liapounoff cherche la fonction inconnue ζ sous la forme de la série

$$(a) \qquad \zeta=\zeta_1\delta+\zeta_2\delta^2+\zeta_3\delta^3+\dots$$

Faisant

$$\Delta(\rho) = \Delta=\sqrt{\rho(\rho+1)(\rho+q)},\ d\sigma'=\sin\theta d\theta d\psi,\ z=D(a+a\zeta,\ a'+a'\zeta') = D(\zeta,\ \zeta'),$$

il trouve

$$U=\Delta\int_0^1 k'a'^2da'\int\frac{(1+\zeta')^2\left(1+\dfrac{\partial a'\zeta'}{\partial a'}\right)}{D(\zeta,\zeta')}d\sigma'.$$

Pour trouver la solution ζ de l'équation fondamentale (A) il faut tout d'abord représenter la fonction U par une série procédant suivant les termes d'ordres différents relatifs à ζ.

Pour que pareille représentation soit possible, il faudra faire à l'égard de ζ certaines hypothèses que Liapounoff exprime par les inégalités

$$|\zeta|<l, \quad \frac{\sqrt{\rho+1}(a+a')}{2}\frac{|\zeta'-\zeta|}{D(0,0)}<g, \tag{4}$$

l et g étant des nombres indépendants de a, a', θ, θ', ψ et ψ'.

Il obtient le développement cherché en remplaçant dans U, ζ et ζ' par $\epsilon\zeta$ et $\epsilon\zeta'$, ϵ étant un paramètre, et démontre ensuite, moyennant les fonctions majorantes construites d'une manière convenable, que, pour ϵ assez petit, $U(\epsilon)$ se développe en série uniformément et absolument convergente de la forme

$$U(\epsilon)=U_0+U_1\,\epsilon+U_2\epsilon^2+\ldots \tag{5}$$

et que ce développement subsiste pour $\epsilon=1$ toutes les fois que

$$g+l<1. \tag{4_1}$$

Dans cette supposition au sujet des nombres g et l, il obtient le développement cherché:

$$U=U_0+U_1+U_2+\ldots, \tag{5_1}$$

où U_0 est le potentiel de l'ellipsoïde E supposé hétérogène, et U_i la fonction homogène de degré i par rapport à ζ et à sa dérivée $\dfrac{\partial\zeta}{\partial a}$.

6. Dans ses recherches ultérieures il se borne aux hypothèses suivantes:

(1) k est une fonction décroissante de a;

(2) ζ et, par suite, les ζ_i sont des fonctions paires de

$$a\sin\theta\cos\psi,\ a\sin\theta\sin\psi,\ a\cos\theta.$$

Remplaçant dans U la fonction k' par son expression (1), il écrit

$$U=V+\Phi\delta.$$

Il vient alors

$$V=V_0+V_1+V_2+\ldots+V_n+\ldots,$$

et

$$U_n=V_n+\Phi_n\delta.$$

L'équation fondamentale (A) se réduit à la suivante:

$$R\zeta-\frac{1}{4\pi a^2}\int\frac{\bar{\zeta}'d\sigma'}{D(a,1)}=W+\text{fonc. de } a,\quad R=\frac{\rho}{2}\int_\rho^\infty\frac{dt}{t\Delta(t)}, \tag{6}$$

où

$$W=\frac{\Omega}{4\pi\Delta}(\rho+\cos^2\psi+q\sin^2\psi)\zeta^2\sin^2\theta+\frac{1}{4\pi\Delta a^2}[V_1+V_2+\ldots-V_2(0)+(\Phi-\Phi(0))\delta]. \tag{7}$$

En portant l'expression (a) de ζ dans (6) et (7) on obtient

$$W = W_1\delta + W_2\delta^2 + W_3\delta^3 + \dots$$

et ces équations pour la définition successive de ζ_i:

(8) $$R\zeta_i - \frac{1}{4\pi a^2}\int \frac{\bar{\zeta}_i' d\sigma'}{D(a,1)} = W_i + \text{fonc. de } a.$$

Supposant que le volume limité par la surface de niveau de la figure cherchée est égal à celui de l'ellipsoïde à demi-axes

$$a\sqrt{\rho+1},\ a\sqrt{\rho+q},\ a\sqrt{\rho},$$

Liapounoff arrive à cette condition pour ζ:

(9) $$\int \zeta d\sigma = -\int \zeta^2 d\sigma - \frac{1}{3}\int \zeta^3 d\sigma,$$

qui conduit à la conclusion que

(10) $$\int \zeta_i d\sigma = N$$

ne dépend que des fonctions

(11) $$\zeta_1,\ \zeta_2,\ \dots,\ \zeta_{i-1}.$$

Cela permet de transformer l'équation (8) en la suivante:

(12) $$R\zeta_i - \frac{1}{4\pi a^2}\int \bar{\zeta}_i'\left(\frac{1}{D(a,1)} - \frac{1}{2a}\int_{\bar{\tau}'}^{\infty}\frac{dt}{\Delta(t)}\right)d\sigma = W_i - \frac{1}{4\pi}\int W_i d\sigma + \frac{1}{4\pi}RN_i,$$

où $\bar{\tau}'$ est une racine positive de l'équation

$$\frac{\rho+1}{\bar{\tau}'+1}\sin^2\theta'\cos^2\psi' + \frac{\rho+q}{\bar{\tau}'+q}\sin^2\theta'\sin^2\psi' + \frac{\rho}{\bar{\tau}'}\cos^2\theta' = a^2.$$

Le second membre de l'équation (12) ne dépend que des fonctions (11) et devient connu quand ces fonctions sont connues.

7. Pour déterminer ζ_i il ne reste qu'à trouver $\bar{\zeta}_i$.

Faisant dans (8) $a=1$ Liapounoff arrive à l'équation intégrale

(13) $$R\bar{\zeta}_i - \frac{1}{4\pi}\int \frac{\bar{\zeta}_i' d\sigma'}{D(1,1)} = \overline{W}_i + \text{const.}$$

$\overline{W}_i$ désignant la valeur de W_i pour $a=1$.

C'est précisément l'équation qu'il a étudiée dans son ouvrage: *Sur les figures d'équilibre*, etc., bien connu, où il a montré, moyennant la théorie des fonctions sphériques et celles de Lamé, que ζ_i' se présente sous la forme

$$\bar{\zeta}_i = c_0\cos\theta + u_i, \qquad \text{si } q=1,$$

$$\bar{\zeta}_i = c_0\cos\theta + c_1\sin^2\theta\sin^2\psi + u_i, \text{ si } q<1,$$

où les u_i sont les fonctions bien déterminées vérifiant respectivement les conditions :

$$\int u_i Y_{10} d\sigma = 0,$$

$$\int u_i Y_{10} d\sigma = 0, \quad \int u_i Y_{23} d\sigma = 0,$$

Y_{ns} désignant les fonctions sphériques élémentaires d'ordre n.

Il démontre ensuite que dans le cas considéré on peut poser $c_0 = c_1 = 0$.

Les fonctions $\bar{\zeta}_i$ et (11) étant connues, on obtient ζ_i à l'aide de l'équation (13).

8. Pour que ζ, défini par la série (a), représente en effet la solution du problème, il faut démontrer non-seulement la convergence de la série (a) pour toutes les valeurs de a satisfaisant aux conditions

$$0 < a \leqq 1,$$

mais encore que ζ reste continue même pour $a = 0$.

La fin de la première partie du manuscrit de Liapounoff (jusqu'à la page 79) est consacrée à la démonstration rigoureuse des propositions signalées.

L'analyse qui lui a permis d'achever le problème est très délicate et très ingénieuse, mais en même temps si compliquée que je ne peux pas la reproduire dans ma communication.

Le monde scientifique aura le plaisir de connaître celle-ci dès que l'ouvrage de notre illustre géomètre aura paru dans les éditions de l'Académie des Sciences de Russie.

De cette manière Liapounoff a réussi à établir l'existence d'une série de figures d'équilibre de la masse fluide hétérogène, peu différentes des ellipsoïdes (non singuliers) de Jacobi et de Maclaurin, dans l'hypothèse que la densité du liquide est de la forme

$$k = 1 + \delta\phi(a)$$

et dans certaines hypothèses générales au sujet de la fonction $\phi(a)$.

9. Cependant, Liapounoff n'a pas pu se borner à la seule démonstration de la possibilité du problème, d'un théorème sur l'existence des figures de l'équilibre. Il s'est proposé d'aller jusqu'à la fin sans tenir compte des difficultés énormes que présente la solution effective du problème, c'est-à-dire le calcul effectif des fonctions, et dans la seconde partie de son ouvrage, il trouve les expressions analytiques de ζ_1, ζ_2, ζ_3, la méthode générale du calcul successif de toutes les fonctions ζ_i et en déduit leur propriétés les plus importantes.

Cette partie de son ouvrage embrasse les pp. 80-320.

Dans l'hypothèse la plus intéressante où $\phi(a)$ est un polynome de a de degré k, il arrive, en appliquant les méthodes les plus variées de l'Analyse, à la conclusion suivante:

La fonction ζ_i, quel que soit l'indice i, peut se représenter sous la forme

$$\zeta_i = Z_0 + Z_2 + \ldots + Z_{2i} + Z_{2i+2} + \ldots + Z_{2ik+2i},$$

où tous les Z_{2m} sont des fonctions sphériques, dont l'ordre est indiqué par l'indice, et des polynomes de a^2 de degré ik; les fonctions

$$Z_{2i+2},\ Z_{2i+4},\ \ldots,\ z_{2i+2ik},$$

sont divisibles respectivement par a^2, a^4, ..., a^{2ik}.

Liapounoff donne les expressions des fonctions Z_{2m} et pour arriver au résultat énoncé fait usage de la théorie des fonctions sphériques, celles de Lamé, des polynomes de Jacobi, des séries hypergéométriques, des polynomes analogues à ceux de Legendre en ajoutant à ces théories des contributions nouvelles; pour déduire les développements en séries de certaines quantités auxiliaires, nécessaires pour le calcul des ζ_i, il emploie souvent une méthode particulière en démontrant que ces quantités cherchées peuvent être considérées comme les racines de certaines équations algébriques dont la solution peut être obtenue à l'aide de la série de Lagrange.

L'une des difficultés principales dans la solution du problème consistait à trouver un développement convenable du potentiel d'une simple couche étalée sur un ellipsoïde homothétique à l'ellipsoïde donné, en série de fonctions sphériques. Cette difficulté paraissait insurmontable si nous allions comparer selon l'habitude, la surface d'équilibre à l'ellipsoïde donné.

Liapounoff a réussi à écarter cette difficulté en introduisant au lieu de celui-ci un ellipsoïde variable, passant par le point de la surface d'équilibre cherchée, pour lequel on trouve la valeur de la fonction potentielle. Cette idée a conduit au développement cherché et puis à la démonstration de la convergence des approximations successives conduisant à la solution du problème.

La démonstration de la convergence de plusieurs séries présentait de même de très grandes difficultés. Liapounoff a surmonté toutes ces difficultés à l'aide de ce théorème remarquable dont il a donné une démonstration dans l'article: *Sur les séries de polynomes*, publié dans le Bulletin de l'Académie des Sciences en 1915.

P_n étant un polynome entier des variables $x_1, x_2, \ldots, x_k$, de degré ne surpassant pas n, si l'on peut trouver un nombre fixe L assez grand pour qu'on ait

$$|P_n| < L,$$

quel que soit n et quelles que soient les valeurs réelles des variables satisfaisant à l'inégalité

$$x_1^2 + x_2^2 + \ldots\ldots + x_k^2 \leqq 1,$$

la série

$$P_0 + P_1\,\alpha + P_2\,\alpha^2 + \ldots\ldots,$$

où

$$|\alpha| < 1 + p - \sqrt{2p + p^2},$$

p étant un nombre positif donné, sera absolument et uniformément convergente pour toutes les valeurs complexes des variables que l'on obtient en considérant l'inégalité

$$|x_1-\xi_1|^2+|x_2-\xi_2|^2+\ldots+|x_k-\xi_k|^2\leqq\frac{p^2}{k},$$

et en faisant varier les nombres réels $\xi_1, \xi_2, \ldots, \xi_k$ sous la condition

$$\xi_1^2+\xi_2^2+\ldots+\xi_k^2\leqq 1,$$

de toutes les manières possibles.

Ces brèves remarques suffisent pour faire comprendre l'intérêt et l'importance des recherches posthumes de notre savant russe.

10. Dans la dernière partie de son ouvrage Liapounoff applique les méthodes développées dans les deux premières au cas particulier du liquide homogène, ce qui le conduit à des résultats nouveaux dans ce problème résolu par lui plus de dix ans auparavant.

Quelques-uns de ces résultats ont été publiés dans son Mémoire: *Nouvelles considérations relatives à la théorie des figures d'équilibre dérivées des ellipsoïdes dans le cas d'un liquide homogène*, paru en 1916 dans le Bulletin de l'Académie des Sciences de Saint-Pétersbourg.

11. Les considérations de Liapounoff ont un caractère purement analytique; plusieurs pages de son manuscrit ne contiennent que des formules mathématiques très compliquées et son œuvre rappelle souvent un chant sans paroles.

Il paraît que pour cet éminent analyste la langue mathématique était claire par elle-même; elle lui suffisait pour pénétrer *in naturam rerum*, mais il faut avouer que plusieurs illustrations et interprétations géométriques et physiques sont encore nécessaires pour donner une idée plus ou moins nette des faits mécaniques et physiques qui se cachent dans la foule de ces formules analytiques; d'un autre côté, il n'y a pas raison de douter que son analyse puisse être simplifiée sous certains rapports; il est hors de doute, enfin, que les recherches de Liapounoff ne sont que le premier pas, quoique très grand, vers la solution de cette importante question et que ses idées sont susceptibles de développement ultérieur et apporteront avec le temps de nouveaux fruits précieux dans le trésor des Sciences.

ON THE FUNCTIONAL DEPENDENCE OF PHYSICAL VARIABLES

BY PROFESSOR GIOVANNI GIORGI,
University of Rome (R. Scuola d'Ingegneria), Rome, Italy

Part 1. Preliminary remarks—Problems to be solved.

1. Consider an iron bridge, at any point of which a variable force $V(t)$, which is an arbitrary function of the time, is applied; we observe the displacement of another point of the bridge from its position of rest, and let it be $W(t)$. This function $W(t)$ is dependent on $V(t)$, but on account of the inertia of the framework, the dependence is not a point-to-point one; that is, at any time t, the value of $W(t)$ is not a function of the simultaneous value of $V(t)$ only, but it depends generally on all values which $V(t)$ has reached at preceding times. We call it a *functional dependence*.

In order to deal with the theory of these dependences, the following symbols will be employed: operators of a functional dependence will be represented by Greek capital letters, so that they will be clearly distinguished from Latin letters and small Greek letters, to be used in their ordinary mathematical signification. Thus we write

(1) $$W(t)=\Theta V(t),$$

where Θ is the operator which changes the curve $V(t)$ into the curve $W(t)$ in such a way that each point of $W(t)$ is dependent on the whole behaviour of $V(t)$. As a particular case, we shall write Δ for the differential operator $\frac{d}{dt}$ (the symbol D employed by Cauchy and many analysts is not suitable in the formulae of mathematical physics, as it may give rise to confusion): Accordingly $\Delta^n=\frac{d^n}{dt^n}$ will represent the differentiator of *nth* order; Δ^{-1} will represent an integration; for instance

$$\Delta^{-1}V(t)=\int_{-\infty}^{t}V(t)dt.$$

Other examples of Θ operators are: Casorati's operator, defined by

$$\theta V(t)=V(t+1)$$

and similarly the transformations of hereditary physics (Volterra's transformations):

$$\Psi V(t)=\int_{-\infty}^{t}G(t,\ \tau)V(\tau)d\tau.$$

It is to be noted, on the contrary, that hysteresis phenomena in magnetic substances do not belong to this class; they show a simplified form of dependen e.

2. We ask whether it is possible with such Θ operators to build up an algebra enabling us to operate with them, and employ them for the solution of physical problems, in the same way as it is possible, for instance, to write $\Delta^m\Delta^n=\Delta^{m+n}$, etc. This would be a generalization of the infinitesimal cacullus, which is particularly concerned with the operators Δ, Δ^m, Δ^{-1}, Δ^{-n}, that is with particular cases of the Θ class. Very interesting instances are known of problems solved in this way: *e.g.*, Boole* has shown how to deal with certain expressions containing Δ, for solving differential equations; other similar solutions have been employed by Forsyth†; Oltramare‡ has dealt with problems which are closely related to our present ones; and many others may be quoted.

Electrodynamics is an important source of problems of this class. Suppose we have a variable current I through a circuit containing a resistance R and an inductance L, the impressed voltage being $V=V(t)$, an arbitrary function of the time. Obviously

$$V=RI+L\,\frac{dI}{dt}\,.$$

When this formula is written

$$V=(RL+\Delta)I,$$

we see that for variable currents the operator $(R+L\Delta)$ plays the same part as the simple ohmic resistance R would play if the current and voltage were steady; so that we are induced to call it the *functional resistance* of the circuit. Reciprocally, we may write:

$$I=(R+L\Delta)^{-1}V,$$

the symbol $(R+L\Delta)^{-1}$ meaning the inverse operation; this operator may be called the *functional conductance*, the formula being "evaluated" by

$$I=(R+L\Delta)^{-1}V=\frac{1}{L}\,e^{\rho t}\int_{-\infty}^{t}Ve^{-\rho t}dt$$

where $\rho=-\dfrac{R}{L}$, as it is obtained by solving the differential equation.

It will be remembered that Heaviside§ has shown how a great many very remarkable results may be obtained in problems of electrodynamic propagation by employing the operator $\Delta=\dfrac{d}{dt}$ as if it were the symbol of a numerical quantity: he has employed it in infinite series and other transcendental expressions in the same way as Boole did in rational fractions. The proper rules however for

**A treatise on differential Equations*, London 1877 and Supplementary Volume.
†*A treatise on differential Equations*, London, 1903; Ch. III.
‡*Leçons sur le calcul de généralisation*, Paris, 1899.
§Especially in *Electrical Papers* and in *Electromagnetic Theory*, vol. II.

obtaining correct results, and the theory of these operational methods have never been given.

Again, the attention of scientists has been awakened by the striking application which Silberstein* has made of the operators cos(*vtcurl*) and sin(*vtcurl*) for solving the problem of electromagnetic waves in space. These operators are remarkably akin to the $f(\Delta)$ ones, as they contain the *curl* which is a particular space-differentiator, instead of Δ which is the differentiator with respect to time.

3. On examining these results and on thinking about those symbols which I have called the functional resistance and conductance I was led some time ago to consider the following problem, which seems interesting: what class of Θ operators may be suitably represented in the form of a finite or infinite expansion containing Δ and its powers, or more generally in the form $f(\Delta)$ where $f(\ )$ stands for the symbol of an analytical expression containing Δ along with constant or variable quantities, subject to the condition that it may be possible to handle such expressions by rules similar to those of ordinary algebra? In my memoir: *Il calcolo simbolico nello studio delle correnti variabili*† I investigated this problem and the behaviour of functional operators when dealing with systems of circuits subject to variable voltages and currents of any kind (of which of course the systems subject to alternating or to periodic currents are quite particular cases, reducible to more elementary methods). There I have shown that, provided the physical properties of the circuits are not variable with the time, all variable-current problems may be simply reduced to steady current problems, by the use of operators of the $f(\Delta)$ type, which behave according to the rules of ordinary algebra: all Steinmetz and other symbols of the form $R+Lj$ which are those used when dealing with simple harmonic alternating currents are elementary forms of those $f(\Delta)$ operators, which in their general forms are valid for currents of any kind whatever. I have further developed the theory of the functional operators in the memoir *Sul calcolo delle soluzioni funzionali*‡, and there I gave, among other things, the reason for Heaviside's results, and explained why the series arising from the infinite expansions of $f(\Delta)$ are sometimes valid and sometimes not. I wish now to deal with the matter from a wider point of view, and at the same time as I summarize the results for the use of English-speaking readers, to show also the path which has led me to build up the present method of operational calculus.

Part 2. Linear operators: their generating functions.

4. When dealing with functional operators of the Θ type, it is necessary first to distinguish between linear (I mean linear homogeneous, that is distributive)

* *Versuch einer Theorie der physikalischen Operatoren* in Annalen der Physik, Oct. 1901, p. 273 and in Annalen der Naturphilosophie 2 (1903), pp. 201-274.

†Read before the Electrotechnical Congress in Naples, October 11th, of 1903; printed in Atti dell' Associazione Elettrotecnica Italiana, vol. VIII (1904), pp. 65-141.

‡Read June 28th of 1905 before the Rome section of A.E.I.; printed in Atti dell' Associazione Elettrotecnica Italiana, vol. IX, fasc. 6° (dic. 1905), pp. 651-699.

operators, and non linear ones. We call Θ linear when it satisfies the condition of addition:

$$\Theta(V_1 + V_2 + \ldots + V_n) = \Theta V_1 + \Theta V_2 + \ldots + \Theta V_n$$

whatever the V functions are. In the present paper I shall consider linear operators only.

This being the case, suppose Θ is operating on a physical variable $V(t)$. The representative diagram* may be considered as the limiting form of an infinity of successive infinitesimal rectangles, having dt as a base, and $V(t)$ as height. Consider each individual rectangle as a single function, so that $V(t)$ is the limit of the sum of all these elementary functions. Then, by the assumption of linearity, the effect of Θ operating on $V(t)$ is the sum of the effects obtained when operating on these elementary functions separately.† So that we obtain an expression of the following kind:

$$(2)\qquad \begin{cases} \Theta V(t) = \displaystyle\int_{-\infty}^{+\infty} G(t,\tau)\, V(\tau) d\tau, \\ \text{or rather} \\ \Theta V(t) = \displaystyle\lim_{n=\infty} \int_{-\infty}^{+\infty} G_n(t,\tau)\, V(\tau) d\tau, \end{cases}$$

which may be written in the condensed form

$$(2^{\text{bis}})\qquad \Theta V(t) = \mathop{\mathbf{S}}_{\tau=-\infty}^{\tau=+\infty} G(t,\tau)\, V(\tau)$$

where $\mathbf{S}$ may be a sum, or an integral or a limit of an integral. The function G, which represents the effect of Θ on an elementary function of unit area located at the time $t=\tau$, is called the *generating* function (compare with the *kernel* of the integral equations).

Part 3. The use of impulsive functions.

5. To avoid unnecessary repetitions, we are led to introduce the notion of *improper functions*. The most important of them are the *impulsive functions* which were several times considered by Heaviside and are familiar to all writers on dynamics dealing with the theory of impulsions. By the symbol

$$Fu(t)$$

and by the name of *impulsive unitary function* I mean a function of t which is everywhere $=0$ for all values of t, except in an infinitesimal interval containing

*By the diagram of a function $V(t)$ will be understood the space included between the curve $V(t)$ and the line $V=0$, and limited, if necessary, by two ordinates.

†It may appear that by doing so we commutate Θ not only with a finite sum, but with an infinite sum, and therefore with the operation "lim", which may be not always legitimate. The intricate theory of the "vicinity" of functions is herein involved. To avoid the difficulty, the elementary functions may be supposed to be chosen, not precisely of the form of rectangles, but of the form $y = \dfrac{h}{\sqrt{\pi}} e^{-h^2t^2}$, or as functions of other kinds satisfying the conditions for obtaining the necessary degree of vicinity.

the point $t=0$, in which interval the function becomes infinitely great with such values that

$$\int_{-\infty}^{+\infty} Fu(t)dt=1. \tag{3}$$

This $Fu(t)$ may be considered as a limiting form of a rectangle or of a function of a kind $Ae^{-h^2t^2}$, or of others, according to the degree of continuity or of vicinity required. Following these ideas, we may, by differentiating $Fu(t)$, define impulsive functions of the second, third order, and so on.*

6. At first sight it may appear that the use of these impulsive functions is strange and illegitimate, as it involves the consideration of actual infinitesimals and infinities. But in fact it is very useful because it simplifies the formulae very greatly and removes the exceptions; in fact not one of the most rigorous writers on dynamics has refrained from introducing the impulsions. As regards the theoretical standpoint, it is to be remarked that actual infinitesimals and infinities may be introduced with perfect rigour as a class of non-archimedean numbers, involving special postulates; or, what amounts to the same thing, we may say that all formulae containing improper functions are formulae wherein a sign of *lim* is understood, so that $Fu(t)$ and similar symbols may be regarded as a kind of short-hand notation. For instance, we write as a fundamental formula, or we should rather say as a definition, the following:

$$V(t)=\int_{-\infty}^{+\infty} V(\tau)Fu(t-\tau)d\tau. \tag{4}$$

This is neither more nor less than an abridged form of writing

$$V(t)=\lim_{n=\infty}\int_{-\infty}^{+\infty} V(\tau)\phi_n(t-\tau)d\tau,$$

where $\phi_n(t-\tau)$ is a properly chosen function, for instance of the form

$$\phi_n(t-\tau)=\frac{n}{\sqrt{\pi}}e^{-n^2(t-\tau)^2},$$

it being always theoretically implied that ϕ_n be such that it is allowed to operate on the above representation of $V(t)$ by the Θ's which we have to deal with, under the sign of *lim* and under the sign of integration.

I cannot give here a more detailed theory about this. In my memoirs above referred to I have shown how to employ the $Fu(t)$ and the $j(n,t)$ properly; and that the definition and introduction of these symbols in that form is legitimate, has been shown elsewhere by Peano, whom I have to thank for the attention given to my papers.

*See my memoir above referred to, *Sul calcolo delle soluzioni funzionali*, etc., III, §9, where I defined and represented by the symbol $j(n,t)$ the impulsive unitary function of order $n+1$, the n being any eal number whatever. This is the same thing as the generating function of derivation with general index, Δ^n.

7. When we subject impulsive functions to an integration of a sufficient order, the result is a proper function.* So, by integrating $Fu(t)$ from $-\infty$ to t, we get a function which I have called $1(t)$ and which is thus defined:

$$1(t) = \begin{cases} 0 \text{ for all negative values of } t, \\ 1 \text{ for all positive values of } t, \end{cases} \tag{5}$$

the value given to it for $t=0$ being immaterial. We may write $1(t)=\Delta^{-1}Fu(t)$; this function $1(t)$ is the typical one for the study of all telegraphic functions, that is, functions having a zero value in certain intervals and a constant value in other intervals. Functions of this kind or ones very closely allied to them have been occasionally considered by Cauchy, Fourier, Dirichlet, Lebesgue and represented by various kinds of analytical expansions.

8. It is to be noted that when improper functions are included in the definition of functions, all functions get a differential coefficient, at least from a certain point of view. For instance, the derivative of the discontinuous function $1(t)$ is the impulsive function $Fu(t)$.

It is also to be noted that when the above implications are understood, we may operate by any Θ on both members of (4) as follows:

$$\Theta V(t) = \int_{-\infty}^{+\infty} V(t) \Theta Fu(t-\tau) d\tau,$$

this also being rather a definition or a short-hand writing than a new equation: so that we see that (2) may always be written in its simplest form

$$\Theta V(t) = \int_{-\infty}^{+\infty} G(t,\tau) V(\tau) d\tau, \tag{6}$$

provided we remember that here $G(t, \tau)$ may be a generalized or improper function, that is, a function containing improper elements (this being equivalent to saying that a sign *lim* is implied), this function being given by

$$G(t, \tau) = \Theta Fu(t-\tau). \tag{7}$$

We shall often use this formula, which is equivalent to saying that $G(t, \tau)$ is the result of the operation of Θ on the impulsive unitary function $Fu(t-\tau)$, or in other words that $G(t, \tau)$ is the generating function of Θ. This is a very abridged form of (2).

Part 4.—The restricted functional space and its transformations.

9. Consider a domain of functions, defined as follows: let $y_1(t), y_2(t), \ldots, y_n(t)$ be n linearly independent functions of t; we do not suppose them to be analytical; they may be any arbitrary functions of the real variable t; let the functional domain be composed of all functions of the form

$$y(t) = \sum_1^n p_n y_n(t), \tag{8}$$

*It is to be understood that by *proper* function I do not understand a finite function, but a function which has no impulsive elements.

where the p's are quantities not dependent on t. In a space of n dimensions, the $y_1(t)$, $y_2(t)$, $\ldots$, $y_n(t)$ may be represented by unit vectors drawn from the origin along the n Cartesian axes. Then $y(t)$ is represented by a general vector having p_1, p_2, $\ldots$, p_n as components along the axes. In this way, each point of the n-dimensional space is associated with a particular function $y(t)$, that is, not with a particular value of the function, but with the shape of the function: that is, each point corresponds to a particular form of a curve. When the domain is so represented, it is called *a functional space.*

For instance,

$$y(t) = p_1 \cos \omega t + p_2 \sin \omega t$$

is a general function in a two-dimensional space having $\cos \omega t$ and $\sin \omega t$ as fundamental functions: this is Steinmetz space of all simple-harmonic alternating currents of angular frequency ω, as usually considered in electrical engineering.

10. Let Θ be a linear operator, having the above functional space as an invariant, that is, such that it transforms all functions $y(t)$ of this space into other functions $Y(t)$ belonging to the same space. For instance, this applies to $\Delta = \dfrac{d}{dt}$ in Steinmetz space. Then Θ may be assimilated to a homogeneous linear transformation of that space, that is, to what has been also called a vectorial homographic correspondence, or a dyadic, or a matrix. The theory of these transformations has been dealt with under various forms by several analysts and vectorialists*; and the application to a linear space of *analytical* functions has been classically built up by Pincherle and Amaldi†.

As long as n is finite, the theory is not difficult to construct. All Θ's operating in the same functional domain, that is, having the same invariant, build up a group of transformations, wherein it is easy to define $\Theta_1 + \Theta_2$, $\Theta_1\Theta_2$, etc., according to the usual assumptions. These symbols are subject to the rules of ordinary algebra but for two peculiarities which have to be noted. First, the exception that products are generally not commutative, that is $\Theta_1\Theta_2$ may be different from $\Theta_2\Theta_1$. Second, the rule for the vanishing of products requires to be expressed in a more general form, that is, from a product being zero we cannot infer that some factor is zero, but only that some factor is *degenerate*‡. The latter is a well-known fact of all multiple algebras§.

*See CAYLEY, *A Memoir on the theory of matrices*, Phil. Trans., vol. CXLVIII, 1858, pp. 17-37, reprinted in his Coll. Math. Papers, vol. II, No. 152; GRASSMANN, *Ausdehnungslehre*, 1861; WHITEHEAD, *A treatise on Universal Algebra*, Cambridge, 1898, Ch. VI; GIBBS and WILSON, *Vector Analysis*, New Haven, 1913 (about dyadics); MARCOLONGO e BURALI-FORTI in all their books and papers on Vector Algebra, about the homographic correspondences between vectors; H. LAURENT, *Exposé d'une théorie nouvelle des substitutions linéaires.* Nouv. Ann. Math. III vol. 15 (1906) pp. 345-365.

†PINCHERLE E AMALDI, *Le operazioni distributive*, Bologna, 1901, and the papers on the same subject by PINCHERLE in various periodicals. During this Congress we have enjoyed very much hearing Professor PINCHERLE's lecture on *Sulle operazioni funzionali lineari.*

‡A degenerate matrix is a matrix whose determinant is zero; it was called *indeterminate* by CAYLEY, and the degree of degenerescence is called *vacuity* by WHITEHEAD.

§See SCORZA, *Corpi numerici e Algebre*, Messina, 1921.

Part 5. The definition of $f(\Theta)$ in the restricted functional space.

11. There is no need here to go over these theories, which may be supposed to be sufficiently known; but I want to extend them by framing an appropriate definition of

$$f(\Theta)$$

where f is a general analytical expression, such as $\log \Theta$, $e^{h\Theta}$, $\sin(c+\Theta)$, etc., containing Θ under any form along with any constant quantities, but not with functions of t. By an appropriate definition I mean one which is consistent with the laws of permanency of forms, and which enables us in practice to employ such expressions as if Θ were simply the symbol of an ordinary quantity.

To obtain this result, it is to be noticed that, as a consequence of the general theory of matrices, each Θ has in the functional n-dimensional space in which it operates, generally n independent invariant axes, that is, n independent vectors, the orientation of which is not changed by the transformation Θ; that is to say, there are generally n independent functions

$$\psi_1(t),\ \psi_2(t),\ \ldots,\ \psi_n(t),$$

such that when Θ operates on any one of them it behaves like a numerical multiplier ρ, without changing the form of the function. Thus we have the following equations:

$$(9)\qquad \begin{cases} \Theta\psi_1(t) = \rho_1\psi_1(t), \\ \Theta\psi_2(t) = \rho_2\psi_2(t), \\ \ldots\ldots\ldots\ldots\ldots \\ \Theta\psi_n(t) = \rho_n\psi_n(t). \end{cases}$$

The systems of the ψ's and of the ρ's are peculiar to the particular operator Θ and connected with it. We may call the ψ's the covariants, that is, the geometrical invariants, or the invariant functions of Θ (they have been called *latent points* by Whitehead), while the ρ's are the invariants, that is, the scalar invariants (*latent roots* of Whitehead) of Θ. When we say that the ρ's are scalar, we mean that they are not functions of t.

For instance, in a general functional space, the exponential functions $e^{\rho t}$ are the invariant functions with respect to Δ; indeed

$$(10)\qquad \Delta e^{\rho t} = \rho e^{\rho t},$$

and this fact explains why exponential functions are so important in analysis. Further, in the particular Steinmetz space as above referred to, the invariant functions are $e^{i\omega t}$, $e^{-i\omega t}$.

The regular case is when the ρ's are all distinct and non vanishing; in this case the ψ's also are all distinct. Attention is to be paid to the particular cases arising when some ρ's are equal with non-coalescent ψ's or with coalescent ψ's, and when some ρ's are vanishing, the latter case taking place when Θ is degenerate; even more complicated combinations may take place, see Pincherle's theories about normal roots and about the proper roots of the powers of a functional transformation. In the case of a degenerate Θ there are functions which reduce identically to zero when Θ operates on them. Thus $\Delta y=0$, when y is a constant, which fact exhibits the degeneration of the differential operation; and $\Delta^m y=0$ when $y=ct^n$, and n is not greater than $m-1$; so, according to Pincherle's definition, we say that t^{m-1} is a proper root for Δ^m.

12. Now, in order to define $f(\Theta)$, first let $\psi(t)$ be an invariant function with respect to Θ, and let ρ be its coefficient. So that $\Theta\psi(t)=\rho\psi(t)$. Then I define

$$f(\Theta)\psi(t)=f(\rho)\psi(t). \tag{11}$$

This definition is consistent with all requirements; obviously it satisfies all rules of ordinary algebra. For instance, let $\Theta=\Delta$, and $\psi(t)=e^{\rho t}$; then $f(\Delta)e^{\rho t}=f(\rho)e^{\rho t}$, including the elementary cases $\Delta^n e^{\rho t}=\rho^n e^{\rho t}$, and similar ones.

Next let $y(t)$ be any function whatever, belonging to our domain of functions, but generally not invariant with respect to Θ; and suppose first that all the ρ's of our Θ are distinct: then also the corresponding invariant functions $\psi_s(t)$ are distinct and linearly independent (a known theorem), that is they constitute a set of fundamental vectors, in terms of which any $y(t)$ belonging to our domain of functions may be represented as follows:

$$y(t)=A_1\psi_1(t)+A_2\psi_2(t)+\ldots+A_n\psi_n(t), \tag{12}$$

where the A's are constant quantities, that is, they do not contain t. In this case, as a natural consequence of the additive law, I am led to define

$$f(\Theta)y(t)=A_1f(\rho_1)\psi_1(t)+A_2f(\rho_2)\psi_2(t)+\ldots+A_nf(\rho_n)\psi_n(t), \tag{13}$$

and the first member will be considered as finite or infinite, single-valued or multi-valued, determinate or indeterminate, according to whether the second member is of the same character.

And now I recall that: the general condition for two operators enjoying the commutative property of multiplication is that they have the same set of invariant ψ's. All $f(\Theta)$'s defined by formula (13), whatever be the $f(\ )$, provided the Θ is always the same, have $\psi_1, \psi_2, \ldots, \psi_n$ as invariant functions, because each of these functions is unchanged by the $f(\Theta)$, it is simply multiplied by a numerical factor. It follows that all these $f(\Theta)$'s, derived from the same operation Θ, are commutative with each other and with Θ; therefore they constitute a sub-group of transformations, wherein the laws of ordinary algebra are obeyed, save for the necessity of taking proper account of the law for the vanishing of a product.

Reciprocally it may easily be shown that if Ω is any given linear transformation of our domain into itself, having as invariant functions the same ψ's which are invariant functions for Θ, it is always possible to express Ω in the form $f(\Theta)$, the number n of dimensions of the functional space being supposed to be finite (compare Laurent's theorem for matrices).

13. The next case is to suppose that some of the ρ's are equal, but that they do not give rise to so-called parabolic properties of the transformation: in this case it is still possible to find a fundamental system of invariant ψ's, allowing us to write formula (13) for any $y(t)$, and the remainder follows unchanged; only the recip-

rocal theorem requires to be so formulated, that in order that Ω may be reducible to the form $f(\Theta)$, it is necessary that *any* function invariant for Θ be also invariant for Ω.

As regards the parabolic case, that is, the case of several ψ's having coalesced into one, I do not need to consider it here, as it is not of special importance to our subject.

14. Something however must be said about the complications which arise when the function $f(\ )$ has a singular point corresponding to some of the values of the ρ's, so that in formula (13) some one of the factors $f(\rho)$ becomes infinite or indeterminate. This case includes the very interesting one wherein some ρ is zero, that is, Θ is degenerate, and $f(\ )$ has a pole in the zero-point of the complex plane. Then the result of $f(\Theta)$ acting on $y(t)$ will contain an arbitrary constant or a set of arbitrary constants. When, by a further particularization, we suppose $f(\Theta)=\Theta^{-1}$ or Θ^{-n}, we fall into the case of inverse operations, which has been fully treated by Pincherle, and which may be a good example for understanding what arises under more general assumptions.

The case of Δ operating on any functional domain containing either the function $t=$constant, or any function ct^m, belongs to this class. Accordingly, Δ^{-1}, and Δ^{-n} give rise to the arbitrary constants which are considered in the integral calculus, whenever it operates on any such space, but not so when its operation is restricted to a domain like $A_1e^{\rho_1 t}+\ldots+A_ne^{\rho_n t}$, where n is finite. The limits of the present paper do not allow us to enter into more details.

Part 6. Pincherle's functional field generalised.

15. The final extension is to consider a functional space with an infinite number of dimensions The theory of the Θ's operating in an enumerable infinity of dimensions is contained in Pincherle and Amaldi's and in Pincherle's books and papers above referred to; these authors have fully developed the application to the domain of Taylor's series, that is, to the functions

$$A_0+A_1t+A_2t^2+\ldots;$$

this is a functional space having 1, t, t^2, ... as fundamental components. We are not concerned with this space because our operands are non-analytical functions; but from general theorems given by the said authors we can easily deduce that for most cases our theory and definition for the $f(\Theta)$ operators remain valid without any change or with obvious changes only.

A functional domain with an enumerable infinity of dimensions is that of all alternating non-simple-harmonic currents of a given frequency. I have sketched out the theory of that domain elsewhere.* Therein are contained all alternating currents used in electrical engineering and most telephonic currents also. But transient currents, telegraphic currents along cables, and others which are not periodic, constitute a more complex domain. Therefore, before going further, it is necessary to precisely define what is the whole field of physical functions which we want to consider.

**Le correnti non sinusoidali*, read Dec. 15th, 1902, before the Rome section of A.E.I.; printed in Atti dell' Associazione Elettrotecnica Italiana, Vol. VII (1903) pp. 34-37.

16. What is a physical function? As a typical instance I take a mechanical force which is arbitrarily applied by the hand to some material body, or a voltage which is impressed on a circuit and is arbitrarily variable with time, such as the voltage at the origin of a telegraphic cable, or the displacements and currents arising therefrom; and so on. Mathematicians have not defined this class of functions. Dirichlet's definition of an arbitrary function of a real variable is far too wide, and the class of continuous functions or even of analytical functions is too narrow for us, and if it were accepted as an approximation, it would not be suitable for the study of the very interesting effects caused by discontinuities.

For general purposes I shall define a physical function of t as one which is arbitrarily given in the interval $-\infty < t < +\infty$, or in a part of it, subject to the conditions that the set of its points of discontinuity and of infinity is *reducible,** that $V(t)$ is of limited total fluctuation in any interval of continuity, and that the integral of the absolute value of the function in any finite interval is convergent; further I require that the function be identical with the derivative of its integral; and I assume that in all points of discontinuity the value attributed to the function can be disregarded, so that two different functions will be regarded a equal if they are equal in all points of continuity.

Part 7. The general definition for $f(\Delta)$.

17. Now, we wish to evaluate $f(\Delta)$, so as to obtain an appropriate definition for $f(\Delta)\,V(t)$, where $f(\Delta)$ is an analytical expression, and $V(t)$ is a physical function. The theory given above for a general $f(\Theta)$ will apply, provided $f(\quad)$ is only built up with Δ and quantities which are constants with respect to t. We shall suppose that the latter conditions are always fulfilled.

We want, accordingly, to get $V(t)$ developed linearly in terms of functions which are invariant with respect to Δ; these are the exponential functions. Whenever $V(t)$ is periodic, this is accomplished by a *Fourier's* series, written in the following form:

$$V(t) = \sum_{-\infty}^{+\infty} A_n e^{int}, \tag{14}$$

where for the sake of simplicity, we have supposed the period reduced to 2π; and then $f(\Delta)$ is evaluated by

$$f(\Delta)\,V(t) = \sum_{-\infty}^{+\infty} A_n f(in) e^{int}, \tag{15}$$

this formula being sufficient for the theory of all cases where alternating currents of any kind are involved.

18. In order to deal with the general case, Fourier's integral is necessary. Now, the physical function as we have defined it, satisfies in any finite interval the sufficient conditions for the expansion as a Fourier's integral and for its

*That is, that some derivative, or finite or transfinite order, of this set is zero; for physical purposes, a more restricted condition would be sufficient, but the theory would not be simplified.

validity at all points of continuity; but it does not generally satisfy the so-called "conditions at infinity" because we cannot put the restriction that $V(t)$ vanishes at infinity. In order to do away with them, we substitute $V(t)$ with a function $\overline{V}(t)$ which has the same value as $V(t)$ in an interval $-T<t<+T$, but has the value zero everywhere outside the interval. This $\overline{V}(t)$ obviously satisfies all required conditions at infinity: therefore by the theorem of Fourier's integral expansion, we may write:

$$\overline{V}(t) = \frac{1}{2\pi i}\int_{-i\infty}^{+i\infty} d\omega \int_{-\infty}^{+\infty} \frac{e^{\omega t}}{e^{\omega\tau}} \overline{V}(\tau)d\tau, \tag{16}$$

where the integration from $-i\infty$ to $+i\infty$ goes along the axis of the imaginary quantities in the plane of the complex variable ω. In my memoir above referred to, I have shown that the function which is subject to that integration, viz.,

$$A(\omega) = \int_{-\infty}^{+\infty} \frac{e^{\omega t}}{e^{\omega\tau}} \overline{V}(\tau)d\tau \tag{17}$$

is always a transcendental integral function of ω: therefore the path of integration with respect to ω may be changed, and it may be any line in the complex plane going from $\omega=-i\infty$ up to $\omega=+i\infty$; we must carefully remember this point.

Moreover, to integrate $\overline{V}(\tau)$ from $\tau=-\infty$ to $\tau=+\infty$ is the same thing as to integrate $V(\tau)$ from $-T$ to $+T$. Therefore the above formula may be written

$$\overline{V}(t) = \frac{1}{2\pi i}\int_{-i\infty}^{+i\infty} d\omega \int_{-T}^{+T} \frac{e^{\omega t}}{e^{\omega\tau}} V(\tau)d\tau.$$

Now, of course, we have identically, for any value of t

$$V(t) = \lim_{T=\infty} \overline{V}(t)$$

and this convergency of $\overline{V}(t)$ towards $V(t)$ is of an infinite (we should rather say of any transfinite) order for any fixed value of t, because from $T>t$ henceforward, $\overline{V}(t)$ is constantly equal to $V(t)$, and in any finite given interval this convergence satisfies also any required condition of uniformity.

Therefore we obtain our expansion:

$$V(t) = \frac{1}{2\pi i} \lim_{T=\infty} \int_{-i\infty}^{+i\infty} d\omega \int_{-T}^{+T} \frac{e^{\omega t}}{e^{\omega\tau}} V(\tau)d\tau \tag{18}$$

where the path of the integration with respect to ω is any line as above explained. This formula is valid for any physical function as defined, and satisfies any required conditions of convergency, which may be necessary to allow it to operate with any functional operation of the class with which we are concerned, under the sign of *lim* and under the signs of integration. The field of the physical functions $V(t)$ is thus represented as a space having a continuous infinity of dimensions.

19. Until now it has been supposed that $V(t)$, although it may have an infinite number of points at infinity, is a proper function, that is, it does not contain any impulsive element. If we wish to include the class of improper functions, it is necessary to add in (18) a convergence factor. Let $Q(\alpha,\omega)$ be a transcendental integral function of α and of ω, which for $\alpha=0$ and for all finite values of ω takes the value unity, and for $\alpha>0$, $\omega=\pm i\infty$ becomes infinitesimal of sufficiently high order; for instance $Q=e^{\alpha\omega^2}$, or $Q=e^{\alpha\omega^4}$, etc., as may be required to remove any case of indetermination caused by the impulsive elements of $V(t)$. Then the expansion may be written in the form:

$$V(t)=\frac{1}{2\pi i}\lim_{\substack{\alpha=0\\ T=\infty}}\int_{-i\infty}^{+i\infty}d\omega\int_{-\infty}^{+\infty}Q(\alpha,\omega)\frac{e^{\omega t}}{e^{\omega\tau}}V(\tau)d\tau. \tag{19}$$

20. On this expansion we operate with the operator $f(\Delta)$ according to our formula (13), and we are thus led to define

$$f(\Delta)V(t)=\frac{1}{2\pi i}\lim_{\substack{\alpha=0\\ T=\infty}}\int_{-i\infty}^{+i\infty}d\omega\int_{-\infty}^{+\infty}Q(\alpha,\omega)f(\omega)\frac{e^{\omega t}}{e^{\omega\tau}}V(\tau)d\tau, \tag{20}$$

provided $f(\Delta)$ has a field of existence extending from $-i\infty$ to $+i\infty$, and the factor $Q(\alpha,\omega)$ be suitably chosen in order to get the necessary convergence of the result; it being understood that the line of integration from $-i\infty$ to $+i\infty$ is arbitrarily chosen, and that $Q(\alpha,\omega)$ and both operations of *lim* may be dropped out when not necessary.

This definition is a particular case of (13), and it may be shown that notwithstanding the infinite number of dimensions it satisfies the same requirements; and all operators $f(\Delta)$ so defined are commutative with each other, and obey the laws of ordinary algebra*.

The formula may be written in the abridged notation of the improper functions

$$f(\Delta)V(t)=\int_{-\infty}^{+\infty}G(t-\tau)V(\tau)d\tau=\int_{-\infty}^{+\infty}G(\theta)V(t-\theta)d\theta, \tag{21}$$

where the generating function G( ), that is

$$G(t)=f(\Delta)Fu(t), \tag{22}$$

is given by

$$G(t)=\frac{1}{2\pi i}\int_{-i\infty}^{+i\infty}f(\omega)e^{\omega t}d\omega. \tag{23}$$

Here, it is understood that $G(t)$ may be an improper function, and that the "limits" and convergence factor are to be introduced when necessary. To avoid difficulties, the latter may for instance be written

$$G(t)=\frac{1}{2\pi i}\int_{-i\infty}^{+i\infty}Q(\alpha,\omega)e^{\omega t}d\omega \tag{24}$$

where α is a quantity which approaches zero.

*When speaking about the rules of ordinary algebra, it will always be understood that the generalized form of the rule for the vanishing of a product is always duly taken into account.

More complete details about the demonstrations and the results are given in my memoir: *Sul calcolo delle soluzioni funzionali* above referred to.

21. Now, compare (21) with (2). The difference is that the generating function instead of being a general function $G(t,\tau)$ of two variables, is a function of their difference $t-\tau$ only. What does that mean physically? Obviously it means that the physical system which has given rise to the functional dependence is invariable with respect to the time: its properties are not changing during the phenomena which are being considered. To affirm this, or to affirm that the functional dependence may be represented by an analytical expression $f(\Delta)$ which contains only constant quantities, or to say that the generating function $G(t,\tau)$ is a function of the single variable $t-\tau$, provided of course the dependence is linear, are equivalent statements. I have called *normal systems* and *normal operators* those satisfying this condition. In these systems, the effect and the after-effect of any given impulsion does not undergo a change of form by shifting the instant at which the impulsion acts, but it remains simply shifted with respect to time. That almost all physically interesting cases enter into this class, I take as self-evident. It may be remarked that transformations of the form (2) are those which occur in *hereditary physics*, a wide branch of science created by Professor Volterra and now extensively studied. Therein, my condition of normality corresponds almost entirely with Volterra's "condition of the closed cycle"; but I think that the name given by me is prior in time.

22. In order to show that any normal functional dependence is really reducible to the form $f(\Delta)$, we have to first write the remarkable formula which gives the evaluation of $e^{h\Delta}$. This formula:

$$e^{h\Delta}V(t)=V(t+h) \tag{25}$$

has sometimes been called Lagrange's form of symbolic Taylor's expansion, and is generally justified as an abridged form of Taylor's series; but may be deduced independently from any assumption about the analyticity of $V(t)$, by the use of our general formula (20), or, more shortly, it may be justified as I have shown in my memoir *Sul calcolo*, etc., 1st part, §1, No. 6.

Now, when any functional dependence of the form (21), that is

$$\Theta V(t)=\int_{-\infty}^{+\infty}G(\theta)\,V(t-\theta)d\theta, \tag{26}$$

is given, wherein $G(\theta)$ may be an improper function (that is, the integral may give rise to finite terms, etc.), we may write $e^{-\theta\Delta}\,V(t)$instead of $V(t)$; therefore

$$\Theta V(t)=\int_{-\infty}^{+\infty}G(\theta)e^{-\theta\Delta}\,V(t)d\theta$$

whatever $V(t)$ may be; and this is the same as writing

$$\Theta=\int_{-\infty}^{+\infty}G(\theta)e^{-\theta\Delta}d\theta, \tag{27}$$

and thus Θ is obtained in the form of a function $f(\Delta)$.

Part 8. Theoretical results.

23. Now, summarizing the results obtained, we have three theorems, the detailed proof of which was given by me in my memoir of 1905, while the theoretical grounds which lead to them, from the point of view of the theory of functional transformations, have been briefly shown here.

$V(t)$ being a proper or improper physical function of the real variable t, and $f(\Delta)$ being an analytical expression containing Δ along with constant quantities only, and having in the complex plane an uninterrupted field of existence from $-i\infty$ to $+i\infty$, the three theorems are expressed as follows:

1st. To any $f(\Delta)$ it is possible to associate a linear functional normal transformation, defined by formula (20), or by the equivalent formula (21), and whose generating function is given by (23).

2nd. All operators so defined obey the commutative laws of multiplication with each other, and the other laws of ordinary algebra, and therefore the expressions $f(\Delta)$ may be dealt with as if Δ were the symbol of a numerical quantity.

3rd. Any normal functional operation, that is, any linear functional dependence arising from physical systems whose properties are not changing during the phenomena to be studied may be expressed by a form $f(\Delta)$ which obeys the first and the second theorem, and which may be expressed by (27) when the (proper or improper) generating function G is known.

In this way a general theory of the symbols $f(\Delta)$ is originated; and I believe that this theory offers to the mathematical physicists a very powerful weapon.

Part 9. How to perform the evaluation practically.

24. It may at first sight appear that my formulae are rather complicated, especially when the convergence factor is needed. But these formulae are only necessary for establishing the theory. In order to secure a knowledge of the result of a given $f(\Delta)$, the only thing to be done is to discuss its generating function $G(t)$, which represents the effect of $f(\Delta)$ acting on an impulsive unitary function $Fu(t)$. Now, $f(\Delta)$ and $G(t)$ are related by the reciprocal formulae (23) and (27), which are neither more nor less than an extended form of the so-called Laplace's transformations. We have many theories of these transformations,* and the chief point which may be deduced from them is that: the two functions $f(\Delta)$ and $G(t)$ are so related, that any singularity cf either of them at a finite distance corresponds to a peculiar singularity of the other function at infinity. So, when $f(\Delta)$ vanishes asymptotically in the directions $+i\infty$ and $-i\infty$, $G(t)$ has no impulsive elements; when $f(\Delta)$ vanishes asymptotically for a given order, $G(t)$ is continuous and has continuous differential coefficients up to a given order; when $f(\Delta)$ has no singular points in the finite part of the complex plane, that is, when $f(\Delta)$ is an integral function, $G(t)$ does not contain any arbitrary constant; and so on. See for further further details, my memoir of 1905, part II, §5 and §6.

*A far reaching theory is that given by PINCHERLE, *Sur les fonctions déterminantes*, Annales de l'École Normale Supérieure, t. XXI. See also some very striking results given by LAURA PISATI in her paper *Sulle operazioni funzionali non analitiche originate da integrali definiti*. Rend. Circ. Mat. Palermo, Tomo XXV (1908) pp. 272-282.

But before explaining something more about the practical use of the $f(\Delta)$, a further point has to be noticed.

25. Whenever $f(\Delta)$ is an integral (polynomial or transcendental) function*, the result of the integration from $-i\infty$ to $+i\infty$ which occurs in the formulae is independent of the path of integration and therefore is single-valued. In the opposite case, several results are obtained by changing the path and by turning over the singular points of $f(\Delta)$; thereby arbitrary constants and even arbitrary functions appear in the result; this fact is an extension of the simpler facts which occur in the case of a finite or enumerable number of dimensions of the functional space, referred to above. So, differentiation Δ or Δ^n is single-valued, while integration Δ^{-1} or Δ^{-n} gives rise to arbitrary constants: when $f(\Delta)$ instead of poles, has critical points or essential singularities at a finite distance, more intricate results arise.

I have not yet worked out the complete theory of these results, but I have noticed that it is always possible to do something similar to what is usually done in the integration of linear differential equations; that is, the general result of $f(\Delta)$ operating on $V(t)$ or on $Fu(t)$ may be split into two parts, one of which is a specially chosen evaluation among the set of functions obtainable; and the other is a complementary term containing all the arbitrary elements. The latter may be represented by $f(\Delta)(0)$, because it is neither more nor less than the general result of the application of $f(\Delta)$ to a zero function.

In the theory of sound, or of vibrating systems, we speak of general vibration, forced vibration and free vibration, which are particular cases. Here, we cannot employ the word "vibration", because the functions may be non-oscillatory. I have introduced the word "variation" to be employed in the general case. So that it will be written

General variation = fundamental variation + free variation

and this equation holds good either for $f(\Delta)\,V(t)$ or for the generating functions, the term of the free variation being in both cases the same $f(\Delta)(0)$.

26. In order to define the fundamental variation, I remark that whenever

$$W(t) = f(\Delta)\,V(t)$$

and when $V(t)$ is really the independent or arbitrary function (for instance, an impressed force) and $W(t)$ is the dependent function (the displacement due to the force), in a physical system, any value of $W(t)$ depends on the values which $V(t)$ has received at the same times and at preceding times, and not on the values which $V(t)$ receives at subsequent times, because the effect follows the cause and cannot precede it. This condition I have called *the condition of succession*, and I have shown that in order to comply with it, that is, in order to get $G(t) = 0$ for all negative values of t, it is necessary to choose the path of integration from $-i\infty$ to $+i\infty$ in such a way that it leaves at the left side (that is, the side towards $-\infty$) all singular points of $f(\Delta)$. Not all $f(\Delta)$ operators admit of an evaluation

*Of course, I speak of $f(\Delta)$, as if in place of Δ a complex variable ω were substituted.

which is consistent with the condition of succession; but, whatever may be the case, provided a path of this kind exists, it gives rise to a function $f(\Delta)V(t)$ which is uniquely definite, and I select it as the *fundamental variation.* In the special cases where this particular path does not exist, for instance for $f(\Delta)=\dfrac{1}{\sin\Delta}$, no particular variation will be chosen as fundamental.

If a path satisfying the above condition is denoted by C, the fundamental generating function will be given by

$$G(t)=\frac{1}{2\pi i}\int_C f(\omega)e^{\omega t}d\omega, \tag{28}$$

with the convergence factor implied when necessary.

27. The general problem of evaluating a given $f(\Delta)$ is thus reduced to: 1st, describe the fundamental generating function which corresponds to this formula; 2nd, describe the free variation $f(\Delta)(0)$.

To give an example, let $f(\Delta)$ be Δ^{-1}. Then the fundamental generating function becomes $1(t)$, and the fundamental evaluation of $\Delta^{-1}V(t)$ is

$$\int_{-\infty}^{t} V(t)dt,$$

while the free variation $\Delta^{-1}(0)$ is an arbitrary constant c. The sum of the two is the general or complete variation

$$\Delta^{-1}V(t)=\int_{-\infty}^{t} V(t)dt+c.$$

28. I am now going to show that in the most important problems of practical interest it is sufficient to find out the fundamental variation.

Whenever $V(t)$ is really the independent function, and whenever the physical system satisfies the second principle of thermodynamics so that a dissipation of energy always takes place, a free vibration or a free variation must ultimately vanish when t increases to $+\infty$; otherwise the so-called "perpetual motion" would result. I call this the *condition of dissipation.* This condition is remarkably related to the behaviour of the function $f(\Delta)$ in the complex plane. I have shown in my former papers that all singular points of $f(\Delta)$ lying on the axis of imaginary quantities give rise to permanent, nonvanishing, variations, and that all singular points having a positive abscissa give rise to variations or vibrations which increase indefinitely with the time. Therefore, the second principle of thermodynamics requires that all singular points of $f(\Delta)$ lie in the negative part of the plane, that is, that they have a negative abscissa.

In this case, when the values of the independent variable $V(t)$ are known from $t=-\infty$ (that is, practically from a remote time) up to a certain instant $t=t_0$, they are sufficient to determine the whole result throughout the same interval, and the result consists of a fundamental variation alone. By a remote time we understand a time sufficient for the practical extinction of all free variations of ordinary amplitude; in most electric circuits this interval is not longer than a fraction of a second.

29. Thus, the conditions of linearity, of normality, of succession and of extinction being satisfied, the evaluation of any $f(\Delta)$ may be generally reduced to the single problem of describing its fundamental generating function, as given by (23). And then, owing to the condition of succession, the result of $f(\Delta)$ operating on $V(t)$ is simply given by

$$f(\Delta)\,V(t)=\int_0^\infty G(\theta)\,V(t-\theta)d\theta, \tag{29}$$

because $G(\theta)$ vanishes for $\theta<0$.

In order to discuss this $G(\theta)$ or $G(t)$ without having to calculate it from (23), the best way is to operate on $f(\Delta)$ by the rules of ordinary algebra and to split it into simple expressions which may be easy to evaluate. In my former papers I gave a set of rules for doing so, and for foreseeing the properties of the generating function from the analytical form of $f(\Delta)$. I quote here only a few of them, in addition to those theorems of reciprocal correspondence referred to above:

1st: the transposition theorem, applied to impulsive functions:

$$f(\Delta+h)\,Fu(t)=e^{-ht}f(\Delta)\,Fu(t). \tag{30}$$

2nd: the general transposition theorem:

$$f(\Delta+h)=e^{-ht}f(\Delta)\,.\,e^{ht}. \tag{31}$$

3rd: the exponential formula, which was our starting point:

$$f(\Delta)e^{ht}=f(h)e^{ht}. \tag{32}$$

4th: the application to a power of t:

$$f(\Delta)t^n=\frac{d^n}{dx^n}\,[f(x)e^{xt}]_{x=0}. \tag{33}$$

5th: the theorem of reciprocity:

$$f(\Delta)\,F(t)_{t=0}=F\left(\frac{d}{dx}\right)f(x)_{x=0}. \tag{34}$$

7th: the theorem for infinite expansions:

Any expansion of $f(\Delta)$ by a series or by an integral or by another infinite process, executed according to the rules of ordinary analysis, is allowable, provided it is valid, and satisfies the usual convergency tests (the sub-uniform convergence of Arzelà and Orlando being sufficient) along a line C as defined in No. 26.

8th: the second theorem for infinite expansions:

Any infinite expansion of $f(\Delta)$ which does not satisfy the above conditions, but which is able to represent $f(\Delta)$ asymptotically along the above line, gives rise to divergent series which when used for arithmetical evaluation may be able to represent the result asymptotically; the initial convergence of these series makes them sometimes more useful than the really convergent series of the 7th theorem.

9th: the test for diffusive operators:

Whenever $f(\Delta)$ vanishes regularly at the infinity point of the complex plane, or at least vanishes when it approaches to it in the directions $+i\infty$, $-i\infty$, its generating function $G(t)$ is a proper function, without impulsive elements: in this case the $f(\Delta)$ operator applied to a discontinuous $V(t)$ changes it into a continuous function.

10th: the theorem of the residues:

Whenever $f(\Delta)$ either is regular or has a pole of any order at the infinity point of the complex plane, or when at least it does not become exponentially infinite when approaching this point along any direction contained in the negative part of the complex plane (I mean, the half plane to the left of the axis of imaginary quantities); its fundamental generating function may be obtained by performing the integration along a circle of infinite radius; that is $G(t)$ is equal to Cauchy's residue of $f(\omega)e^{\omega t}$ around the point at infinity.

11th: if n is a negative real number $=-m$, the fundamental generating function of Δ^n is:

$$\Delta^n Fu(t) = \begin{cases} 0 \text{ for } t<0, \\ \text{is not impulsive for } t=0, \\ \dfrac{t^{m-1}}{\Gamma(m)} \text{ for } t>0, \end{cases}$$

where Γ is Gauss' gamma-function.

12th: if n is a positive real number $n=N-m$, where N is integral, and m is less than unity, the fundamental generating function of Δ^n is:

$$\Delta^n Fu(t) = \begin{cases} 0 \text{ for } t<0, \\ \text{impulsive function of } (N-1)th \text{ order, for } t=0, \\ \dfrac{t^{-n-1}}{\Gamma(-n)} \text{ for } t>0. \end{cases}$$

Part 10. Heaviside's and other authors' results—Their explanation.

30. By reason of the above theory, the very remarkable results obtained by Heaviside may be explained and accounted for. His main problem, in our language, consists of applying an operator $f(\Delta)$, arising from propagation problems, to the function which we have called $1(t)$, and describing or developing the function $H(t)$ obtained. As $1(t)=\Delta^{-1}Fu(t)$, this $H(t)$ is the same thing as the generating function of $f(\Delta)\Delta^{-1}$. So, when $f(\Delta)=\Delta^{\frac{1}{2}}$, the corresponding $H(t)$, as a consequence of the 11th theorem above, becomes:

$$H(t) = \begin{cases} 0 \text{ for } t\leq 0, \\ \dfrac{1}{\sqrt{\pi t}} \text{ for } t>0, \end{cases}$$

a formula which is of paramount importance for all Heaviside's expansions, and which he said he had obtained "by experimental trials"; we should rather say he obtained it by a genial divination.

In most cases, for obtaining the $H(t)$ corresponding to a given $f(\Delta)$, Heaviside expands this operator into a series of terms $a_n\Delta^{\frac{n}{2}}$, where n is a positive or negative integral number. These expansions may now be viewed as particular cases among the general expansions referred to above. Our theory then explains why the series arising therefrom are sometimes convergent, with a rapid or slow convergence, and sometimes divergent, with a rapid initial convergence (asymptotic series) and sometimes wholly divergent or even convergent toward wrong results, as Heaviside has shown.

31. The formulae for fractional differentiation, which have been given by Riemann and many other analysts* and which are different from each other, have likewise a theoretical explanation. Whenever n is fractional, the operation Δ^n is many-valued, because the function Δ^n then has a critical point at the origin. What the various authors have given are particular evaluations, which are all contained in our general formula (20). The fundamental evaluation is the one resulting from our 11th and 12th theorems above. The finding of the complete expression for the general evaluation in finite terms has hitherto been an unsolved problem.

32. Boole, Forsyth, Lord Rayleigh and others have already shown how to employ operational symbols for solving ordinary linear differential equations with constant coefficients, or systems of these equations, and their proof of the validity of the method consists in verifying that the result satisfies the equation.

For the sake of comparison, I am now going to obtain the same solution and to explain how it is directly justified in the light of my theorems, without the necessity of testing the solution obtained.

To take a simple case, suppose that we have to solve the equation

$$a_0+a_1\frac{dy}{dt}+\ldots+a_n\frac{d^ny}{dt^n}=V(t)$$

where the a's are constant and $V(t)$ is a given function of t. We write it in the form

$$(a_0+a_1\Delta+\ldots+a_n\Delta^n)y=V(t),$$

therefore

$$y=\frac{1}{a_0+a_1\Delta+\ldots+a_n\Delta^n}V(t).$$

The operator is an expression composed of Δ and constant coefficients only; therefore it is permissible to expand it by the rules of ordinary algebra. For the sake of simplicity we suppose that all roots of the polynomial $a_0+a_1\Delta+\ldots a_n\Delta^n$

*See these formulae in HAGEN's *Synopsis der Hoeheren Mathematik*. See also PINCHERLE's writings on the same subject.

are different, then the fraction, by a known method, may be expanded into a sum of terms like $\frac{A}{\Delta-\rho}$ which we have to evaluate. By the theorem of transposition

$$\frac{A}{\Delta-\rho}=e^{\rho t}\frac{A}{\Delta}e^{-\rho t}.$$

Therefore the general evaluation is:

$$\frac{A}{\Delta-\rho}V(t)=Ae^{\rho t}\int^{t}\frac{V(\tau)}{e^{\rho\tau}}\,d\tau,$$

and a sum of terms similar to this composes the general solution of the equation.

In the application to physical problems where the conditions of succession and of extinction are satisfied, the general evaluation is split into the fundamental evaluation and the free variation as follows:

$$\frac{A}{\Delta-\rho}V(t)=Ae^{\rho t}\int_{-\infty}^{t}\frac{V(\tau)}{e^{\rho\tau}}\,d\tau+Ce^{\rho t}, \tag{35}$$

and ρ in these cases is always a quantity with a negative abscissa, the absolute value of which abscissa measures one of the degrees of stability of the system. For further details see my memoir: *Il metodo simbolico*, etc., Nos. 22 to 31, and go back to the very suggestive theory given by Lord Rayleigh in the *Theory of Sound*, vol. I.

Part 11. How to deal with problems of a higher order.

33. Coming now to higher problems, suppose we want to solve a partial differential equation relating to the propagation of heat, as for instance the system:

$$A^2\frac{d^2u}{dx^2}+Bu=\frac{du}{dt};\qquad u=\frac{dv}{dx}, \tag{36}$$

where A, B are constant coefficients.

Putting $\frac{d}{dt}=\Delta$, and substituting Ax instead of x, Av instead of v, the second equation remains unchanged, and the first one may be written

$$\frac{d^2u}{dx^2}+(B-\Delta)u=0. \tag{37}$$

In view of our general theorem, Δ may be dealt with here as if it were the symbol of a constant quantity. Therefore the equation is now an ordinary differential equation, and by the preceding method we get the general solution in the form

$$u=e^{x\sqrt{\Delta-B}}C_1+e^{-x\sqrt{\Delta-B}}C_2,$$

and therefore

$$(38) \qquad v = \frac{1}{\sqrt{\Delta - B}} e^{x\sqrt{\Delta-B}} C_1 + \frac{1}{\sqrt{\Delta-B}} e^{-x\sqrt{\Delta-B}} C_2 + C_3,$$

where C_1, C_2 and C_3 are arbitrary constants with respect to x, and therefore arbitrary functions of t. Suppose the conditions of succession and extinction (as must be) satisfied, and suppose that B is negative and x positive. Then C_1 and C_3 must be zero, and writing $-V_2(t)$ instead of C_2, we have

$$v = \frac{1}{\sqrt{\Delta-B}} e^{-x\sqrt{\Delta-B}} V_2(t).$$

To evaluate this, first use the transposition theorem which gives

$$v = e^{Bt} \frac{e^{-x\sqrt{\Delta}}}{\sqrt{\Delta}} V(t),$$

where $V(t)$ stands for $e^{-Bt}V_2(t)$. The problem is now reduced to evaluating the operator $f(\Delta) = \frac{e^{-x\sqrt{\Delta}}}{\sqrt{\Delta}}$. On account of the 9th theorem of No. 29, this operator is diffusive: therefore

$$(39) \qquad \frac{e^{-x\sqrt{\Delta}}}{\sqrt{\Delta}} V(t) = \int_{-\infty}^{t} G(t-\tau) V(\tau) d\tau,$$

where the generating function $G(t)$ is a non-impulsive function whose expression may be obtained in terms of known functions by an application of the theorem of residues. I hope to have some opportunity to develop the full treatment of this case elsewhere.

34. Considering the general case of electric systems, a far-reaching theorem may be deduced from our theory. Suppose that any finite or infinite combination of constant resistances, inductances and capacities is given; then any problem whatever in regard to variable currents may be reduced to a steady-current problem, by simply imagining that for each inductance l is substituted a resistance $l\Delta$, and for each capacity k a conductance $k\Delta$; the solution obtained by the methods which are valid for steady currents will then be composed of terms of the form $f(\Delta) V(t)$, and it is only necessary to evaluate these terms by the general rules which we have given here.

The utility of this theorem is two-fold, because it is not only equivalent to a rule for solving the differential equations, but also it avoids the necessity of writing them.

35. To illustrate what has been said in the preceding paragraph, let a network with a finite number of conductors be given, however connected, each conductor having a resistance, an inductance, and a capacity, and suppose that any number of arbitrarily variable electromotive forces are acting between the several points of the network: it is required to find the current flowing through a given section of the system. Consider the generalized conduct-

ance of any conductor of the network as given by $G_s = k_s\Delta + \dfrac{1}{r_s + l_s\Delta}$. If the G_s were ohmic conductances, the result would be given by a formula of this kind:

$$I = \sum_i \frac{P_i(G_1, G_2, \ldots, G_n)}{Q_i(G_1, G_2, \ldots, G_n)} V_i, \tag{40}$$

where the P's and the Q's are polynomials.* Substitute for each G its expression, and remember that the V's are given functions of the time. Then we get

$$I_i = \sum_i R_i(\Delta) V_i(t) \tag{41}$$

where the R_i's are rational algebraic functions of Δ involving constant coefficients.

Each of these $R(\Delta)$ may be developed, by algebraic methods, into a polynomial $M(\Delta)$ and a sum of terms like $\dfrac{a}{\Delta - \rho}$ or $\dfrac{a}{(\Delta-\rho)^2}$ or similar ones, which we know how to evaluate.

The problem may be further complicated by the inclusion of mutual inductances. See the results in my memoir: *Il metodo simbolico*, etc., Nos. 61 to 65.

Part 12. The propagation of variable currents in cables.

36. We shall deal now with a problem of propagation along circuits composed of a continuous infinity of elements, which is not dealt with in my previous memoirs.

Consider a cable or a line, containing a longitudinal resistance r, a transversal conductance g, a longitudinal inductance l, and a transversal capacity k, all distributed uniformly and reckoned per unit length of the cable. Let the cable be of infinite length, and let a voltage $V_0(t)$ be impressed at its origin. We want to know the voltage $V(x, t)$ at a point x and at the time t, in terms of the values of $V_0(t)$ which are supposed to be given for all previous times, from $-\infty$ to t.

If the currents were steady, the solution would be

$$V = e^{-x\sqrt{gr}} V_0,$$

the hypothesis that V does not become infinite for $x = +\infty$ being implied.

Here, instead of g put $g + k\Delta$; and instead of r put $r + l\Delta$; and the solution for variable currents will be given by:

$$V(x, t) = e^{-x\sqrt{(g+k\Delta)(r+l\Delta)}} V_0(t). \tag{42}$$

It is only necessary to find the fundamental evaluation of this expression, because the conditions of succession and of extinction are satisfied. The expression may be written

$$V(x, t) = e^{-ax\sqrt{(\Delta-\rho)^2-\sigma^2}} V_0(t), \tag{43}$$

*See full details in my memoir *Il metodo simbolico*, etc., Nos. 54 to 60.

where:

$$(44)\qquad \begin{cases} a=\sqrt{kl}, \\ \rho=-\dfrac{1}{2}\left(\dfrac{g}{k}+\dfrac{r}{l}\right), \\ \sigma=\dfrac{1}{2}\left(\dfrac{g}{k}-\dfrac{r}{l}\right). \end{cases}$$

37. In the particular case $\sigma=0$, and only in this case, the formula reduces to

$$(45)\qquad V(x,t)=e^{-ax(\triangle-\rho)}V_0(t),$$

which is immediately evaluated by our formulae (25) and (31), giving

$$(46)\qquad V(x,t)=e^{\rho ax}V_0(t-ax),$$

that is, the function V_0 travels along the cable with the velocity $\dfrac{1}{a}=\sqrt{\dfrac{1}{kl}}$, and with a coefficient of attenuation $e^{\rho ax}$, otherwise keeping its form unchanged. It is the *distortionless* case discovered by Heaviside, and our theory not only confirms his results, but proves at once that no other combination of g, r, k, l permits a distortionless transmission.

38. In the general case $\sigma \lessgtr 0$, the transposition theorem gives

$$V(x,t)=e^{\rho t}e^{-ax\sqrt{\triangle^2-\sigma^2}}e^{-\rho t}V_0(t).$$

In this case we have to evaluate the operator $e^{-ax\sqrt{\triangle^2-\sigma^2}}$. To do it, we may notice that in the proximity of the infinity point of the complex plane, this function behaves in the same way as the distortionless operator $e^{-ax\triangle}$; it is thus useful to write

$$V(x,t)=e^{\rho t}e^{-ax\triangle}.e^{-\rho t}V_0(t)+e^{\rho t}[e^{-ax\sqrt{\triangle^2-\sigma^2}}-e^{-ax\triangle}]e^{-\rho t}V_0(t).$$

The first part is therefore the same as in the former case, and the second part shows the additional term due to the distortion factor σ. Thus we have

$$(47)\qquad V(x,t)=e^{\rho ax}V_0(t-ax)+e^{\rho t}[e^{-ax\sqrt{\triangle^2-\sigma^2}}-e^{-ax\triangle}].e^{-\rho t}V_0(t).$$

Now it may be remarked that the operator within the square brackets satisfies the conditions for diffusivity, given in our 9th theorem of No. 29. The difference of results corresponding to the first and the second terms of this formula is easily seen if we consider the generating functions, that is, if we reduce the $V_0(t)$ to an impulsive function: the first term represents a wave front, the same as obtained in the distortionless case; the second term gives the tail of the wave, that is, it exhibits the phenomenon whereby each impulsion gives rise to a diffused voltage which spreads along the cable. This second term, on account of its diffusivity, may be expressed by

$$(48)\qquad e^{\rho t}\int_{-\infty}^{t}S(t-\tau)e^{-\rho\tau}V_0(\tau)d\tau,$$

where $S(t)$ is the generating function of the operator within the brackets, and it is a proper function, which may be obtained by the use of Cauchy's residues according to our 10th theorem of No. 29, or by expanding it into an infinite series, following the 7th and 8th theorems of No. 29. When the calculation is performed (it would take too long to develop it here), an expression $S(t)$ is obtained, containing exponential functions and Bessel's non-oscillating functions.

The result for the particular operand $V(t)=1(t)$, has been foreseen by Heaviside, and by other analysts in similar cases; but I think this is the first attempt to prove it by a rigorous theory and to give the general formula.

The theory of the propagation of electromagnetic or light waves in the ether or in an absorbing or in a diffusing medium may be given by the same equations.

39. There is no difficulty in solving in the same way still more complicated problems: suppose that instead of the voltage $V(x, t)$ at a point x of a cable, we wish to know the current $I(x, t)$ flowing through the cable. In the case of steady currents, it is known that $I(x)=\sqrt{\frac{g}{r}}\,V(x)$; therefore by putting $g+k\Delta$ and $r+l\Delta$ instead of g and r, we get

$$I(x, t)=\sqrt{\frac{g+k\Delta}{r+l\Delta}}\,V(x, t), \tag{49}$$

where $V(x, t)$ is found as above; and the action of the operator $\sqrt{\frac{g+k\Delta}{r+l\Delta}}$ may be evaluated by the same methods.

If a system of concentrated resistances, inductances and capacities is connected to the origin of the cable, this system gives rise to another factor which is a rational function of Δ, and we have seen how to evaluate it. Without the use of the $f(\Delta)$ method, it would scarcely be possible to solve any problem concerning such a system.

---

40. The most fruitful results of the $f(\Delta)$ method are however obtained when we learn how to foresee in each case the properties of the solutions by the analytical form of the $f(\Delta)$ expression; this is done mainly by considering the behaviour of $f(\Delta)$ at the infinity point of the complex plane, and the position and the character of the singularities at finite distance.

It is also possible to employ the method for solving an inverse problem which is of very special interest, *i.e.*, how to arrange a combination of component physical elements in order that the resulting system may possess certain properties. For instance, in electrical engineering, we have a cable which gives rise to distortion: we want to arrange at the origin or at the end of the cable a system of inductances, capacities and resistances, in order that the final effect is improved: the problem is solved by finding a rational function $R(\Delta)$ which when attached as a factor to the $f(\Delta)$ of the cable, suitably modifies its behaviour at the infinity point, so as to approach the $f(\Delta)$ of a distortionless cable; then, on the ground of the theory already given for a network of conductors, it is found how to compose a system which gives rise to the $R(\Delta)$.

I shall not dwell any further on the subject. I hope that I have sufficiently awakened the interest of my fellow-workers in physical mathematics and in electrical engineering to the study of functional dependences by the use of $f(\Delta)$ and similar operators, and that I have shown them how a rigorous theory may be built up which leads to rules for the safe use of operational symbols, and how this theory opens the way to an almost untouched field of research, destined to be very fruitful.

UEBER RELATIVE MAXIMA DES NEWTONSCHEN POTENTIALS

VON A. BEZIKOWITSCH,
Professor an der Universität Leningrad, Leningrad, Russland.

Das Problem der Bestimmung der Gleichgewichtsform einer ruhenden Flüssigkeitsmasse lässt sich bekanntlich auf dasjenige der Bestimmung des Maximums des Newtonschen Potentials einer homogenen Masse zurückführen. Die erste Lösung dieses Problems hat Liapounow* gegeben, indem er gezeigt hat, dass ein Potentialmaximum, wenn es überhaupt existiere, nur bei kugelförmigen Massenverteilung stattfinden kann. Die Frage nach der Existenz dieses Maximums blieb jedoch unaufgeklärt.

Im Jahre 1920 hat H. Rosenblatt† die Existenz eines absoluten Maximums bewiesen, welches eben der kugelförmigen Massenverteilung entspricht. Damit war aber noch nichts über die Existenz der relativen Maxima gesagt. In der vorliegenden Notiz wird ein Beweis der Nichtexistenz der relativen Maxima des Newtonschen Potentials gegeben, womit das Problem zu einem gewissen Abschluss gebracht zu sein scheint.

Wir wollen annehmen, dass der Begriff der Massenverteilung in einer räumlichen Punktmenge nur für (im Lebesgueschen Sinne) messbare Mengen definiert ist; die Gesamtmasse setzen wir dem Masse der Menge gleich (die Dichte wird dabei gleich Eins angenommen).

Es sei V die von der gegebenen Masse erfüllte Punktmenge und A die Orthogonalprojektion derselben auf eine Ebene P. Bekanntlich ist die Menge der Fusspunkte derjenigen Lote zur Ebene P, auf welchen die Punkte von V eine nicht messbare Menge bilden, eine Menge vom Masse Null. Es seien ferner: M_1, M_2 zwei willkürliche Punkte von V; L_1, L_2 die durch die beiden Punkte auf die Ebene P gefällten Lote (es wird angenommen dass die Punkte von V die auf L_1 und L_2 liegen messbare Mengen bilden); dl_1, dl_2 die Differentiale des Masses der auf den Geraden L_1, und L_2 liegenden Punkte von V; r_{12} die Entfernung von dl_1 und dl_2; $d\sigma_1$ bezw. $d\sigma_2$ das den Punkte M_1 bzw. M_2 enthaltende Flächenelement von A. Dann drückt sich das Potential von V folgendermassen aus:

$$\int\limits_{A\,(\sigma_1)}\!\!\int \int\limits_{A\,(\sigma_2)}\!\!\int d\sigma_1\, d\sigma_2 \int\limits_{L_1}\int\limits_{L_2} \frac{dl_1\, dl_2}{r_{12}}.$$

*Liapounow. Communications de la Soc. Math. de Kharkow. 1e Série. 1886 (en russe.)
†Rosenblatt. Bull. Sci. Math. 1920.

2—5

Man beweist ohne Mühe die folgende Eigenschaft des Integrals:

$$\int_{L_1}\int_{L_2}\frac{dl_1 dl_2}{r_{12}}.$$

Falls die Punkte von V auf der Geraden L_1, L_2 eine einzige Strecke bilden, nimmt dieses Integral zu, wenn man jene Strecken längs den Geraden L_1, L_2 auf eine solche Weise verschiebt, dass die Entfernung ihrer Mittelpunkte abnimmt.*

Wir wollen nun, auch für den Fall einer allgemeinen Verteilung der Punkte von V auf den Geraden L_1, L_2, eine Klasse von analogen Operationen definieren; jede Operation wird dabei durch eine Zahl $a > 1$ charakterisiert und mit dem Namen "a-Verschiebung" bezeichnet.

Wir nehmen an, dass die Punkte der Menge V auf den Geraden L_1, L_2 eine endliche Anzahl von Strecken bilden. Wir verschieben jede Strecke längs der entsprechenden Geraden derart, dass die Entfernung von der Ebene P des Mittelpunktes der Strecke in ihrer Endlage sich auf den a_1-ten Teil ($a_1 > 1$) der ursprünglichen Entfernung reduziert; dabei darf aber kein Streckenpaar in seiner neuen Lage einen gemeinsamen Teil haben; dies wird bei genügend kleinem a_1 immer der Fall sein. Eine solche Lagenänderung nennen wir "elementare a_1-Verschiebung".

Es ist leicht einzusehen, dass durch eine derartige Verschiebung der Wert des Integrals

$$\int_{L_1}\int_{L_2}\frac{dl_1 dl_2}{r_{12}}$$

vergrössert wird, denn es gibt mindestens zwei Strecken, die auf den Geraden L_1 resp. L_2 liegen und für welche die Entfernung ihrer Mittelpunkte dadurch verkleinert wird.

Die grösstmögliche elementare a_1—Verschiebung bringt mindestens ein Paar von Strecken zum Zusammenstossen, und die Anzahl der Strecken wird dadurch mindestens um eine Einheit vermindert.

Gestattet ein gegebenes Streckensystem eine elementare a-Verschiebung, so wollen wir eben diese als "a-Verschiebung" schlechtweg bezeichnen. Wenn dagegen a dazu zu gross ist, so bringen wir zunächst die grösstmögliche elementare a_1-Verschiebung zustande; dann unterwerfen wir das soeben entstandene Streckensystem einer neuen a-Verchiebung u. s. f. Fahren wir auf diese Weise fort, so werden wir schliesslich imstande sein, eine solche a_k-Verschiebung durchzuführen, dass das Produkt $a_1 a_2 \dots a_k$ gleich der gegebenen Zahl a wird. In der Tat wird durch jede grösstmögliche Elementarverschiebung die Anzahl der Strecken vermindert; wenn aber auf jeder der Geraden L_1, L_2 eine einzige Strecke übrig bleibt, wird man offenbar eine beliebige Elementarverchiebung vollziehen können. Die aus diesen k Elementarverschiebungen zusammengesetzte Lagenänderung wollen wir ebenfalls "a-Verschiebung" nennen. Offenbar ist die

*Der oben angeführte Ausdruck für das Potential bildet den Ausgangspunkt des Beweises von H. Rosenblatt; der Beweis stützt sich auf die soeben zitierte Eigenschaft des Integrals.

a-Verschiebung auf einer der beiden Geraden L_1, L_2 von der Lage der Strecken auf der anderen Geraden unabhängig und die gleichzeitige Betrachtung der beiden Geraden hatte allein den Zweck, die Ausdehnung der bereits ausgesprochenen Eigenschaft des Integrals

$$\int_{L_1}\int_{L_2}\frac{dl_1\,dl_2}{r_{12}}$$

auf den Fall der a-Verschiebung zu ermöglichen. Unsere Betrachtungen zeigen nämlich, dass dieses Integral bei einer a-Verchiebung zunimmt. Wenn auf einer Geraden L die Punkte von V unendlich viele Strecken bilden, so kann man die a-Verschiebung mit Hilfe eines geeigneten Grenzüberganges definieren. Man unterwerfe nämlich der a-Verschiebung die n grössten Strecken und lasse sodann n über alle Grenzen wachsen. Man überzeugt sich leicht, dass dieser Vorgang alle Strecken von L wirklich in eine gewisse Grenzlage bringt.

Wenn die Punkte von V auf der Geraden L eine willkürliche messbare Menge E vom Masse m bilden, konstruiere man eine Folge von Streckensystemen

$$m+\epsilon_1,\ m+\epsilon_2,\ \ldots,\ m+\epsilon_n,\ \ldots\ ;\epsilon_n\rightarrow 0$$

deren jede alle Punkte von E bedecken möge. Auch in diesem Falle überzeugt man sich ohne Mühe, dass die Punkte E im Laufe der Verschiebungen einer Grenzverteilung E' zustreben; dabei wird das Mass von E' den ursprünglichen Wert m behalten. Auch in diesem allgemeinen Falle sieht man leicht ein, dass die a-Verschiebung den Wert des Integrals

$$\int_{L_1}\int_{L_2}\frac{dl_1dl_2}{r_{12}}$$

vergrössert.

Nur in dem Falle, dass die Punkte von V auf jeder der Geraden L_1, L_2 eine einzige (abgesehen von einer Menge vom Masse Null*) Strecke bilden, wobei die Mittelpunkte beider Strecken von der Ebene P gleich entfernt sind, ändert das Integral

$$\int_{L_1}\int_{L_2}\frac{dl_1dl_2}{r_{12}}$$

bei einer a-Verschiebung seinen Wert nicht.

Man sieht daraus, dass das Potential der Menge V seinen Wert nur in dem Falle nicht ändert, wenn die Menge V (abgesehen von einer Menge vom Masse Null) in Bezug auf eine der Ebene P parallele Ebene symmetrisch liegt, und wenn überdies alle zu P senkrechten Geraden mit der Menge V je eine einzige Strecke (stets von einer Menge vom Masse Null abgesehen) gemein haben ($\pm$ eine Menge vom Masse Null). Ist dies nicht der Fall, so nimmt das Potential zu.

*d. h. die Strecke$\pm$eine Menge vom Masse Null.

Ist $a=1+\epsilon$, wo ϵ eine unendlichkleine Grösse bedeutet, so kann man noch zeigen, dass der Zuwachs des Potentials in Bezug auf ϵ von der ersten Ordnung ist.

Man sieht also, dass die a-Verschiebungen in allen möglichen Richtungen nur dann zu keiner Vergrösserung des Potentials führen, wenn V Symmetrieebenen von allen möglichen Lagen besitzt und wenn überdies eine Gerade von beliebiger Richtung mit V eine einzige Strecke (abgesehen von einer Menge vom Masse Null) gemein hat.

Die einzige Menge, welche diesen Bedingungen genügt, ist offenbar (abgesehen von einer Menge vom Masse Null) eine Kugel.

Ist die Menge V in einem begrenzten Gebiete eingeschlossen, so entspricht einem unendlich kleinem Werte von ϵ stets eine unendlich kleine $a=(1+\epsilon)$-Verschiebung der Menge V.

Dies beweist eben die Nichtexistenz der relativen Maxima des Newtonschen Potentials, und es ist also die Kugel die einzig mögliche Gleichgewichtsform einer ruhenden von äusseren Kräften nicht angegriffenen Flüssigkeit.

Ich glaube, dass der oben angeführte Hinweis auf die Grössenordnung des Potentialzuwachses für die vollständige Lösung des Problems von Wichtigkeit ist. Es ist überdies folgender Umstand hervorzuheben. Für das eigentliche Problem des Gleichgewichts einer Flüssigkeitsmasse kommen vielleicht nur diejenigen Massenverteilungen in Betracht, welche gewissen Bedingungen analytischen Characters genügen, wie z. B. die Stetigkeit der Massenverteilung, das Vorhandensein der Berührungsebenen der Oberfläche u. s. w.

Man sieht leicht ein, dass eine Verschiebung $a=1+\epsilon$ die Stetigkeit derartiger Massenverteilungen nicht stört. Wenn aber die Frage nach den Berührungsebenen, nach der Krümmung u. s. w. ausser Betracht bleibt, so kann man, nachdem man die a-Verschiebung bereits vollzogen hat, mit Hilfe von beliebig kleinen neuen Lagenänderungen der Massenverteilung die gewünschten analytischen Züge erteilen, ohne damit den Hauptteil des Potentialzuwachses zu ändern.

Wenn dagegen V nicht in einem begrenzten Gebiete eingeschlossen ist, unterwirft die $a=(1+\epsilon)$-Verschiebung, bei unendlich kleinem ϵ, unendlich ferne Punkte unendlich grossen Lagenänderungen.

Es ist in diesem Falle nicht mehr erlaubt, die Lagenänderung von V als unendlich klein zu betrachten.

Es bietet aber keine Schwierigkeit von der zur obigen Beweisführung angewandten Idee ausgehend, auch in diesem Falle eine zu einer Vergrösserung des Potentials hinleitende unendlich kleine Verschiebung zu konstruieren.

AU SUJET DES ONDES D'ÉMERSION

PAR M. R. RISSER,
Chef du Service de l'Actuariat du Ministère du Travail, Paris, France.

Une masse liquide, d'une profondeur finie ou infinie est au repos; on la met en mouvement par l'action de forces impulsives appliquées à sa surface (effet d'un coup de vent) ou par l'émersion d'un corps solide; dans le premier cas, il se produit des ondes dites par impulsion, dans le second cas, des ondes par émersion.

Prenons pour étudier le mouvement des ondes, trois axes de coordonnées rectangulaires; le plan xoy est confondu avec le plan de la surface libre au repos, et l'axe oz dirigé vers le bas. Nous désignerons par ϕ le potentiel des vitesses, par u, v, w les composantes de la vitesse d'une particule, et par g l'accélération de la pesanteur.
On sait que l'on a:

$$(u,\ v,\ w)=\frac{\partial\phi}{\partial(x,y,z)}.$$

Le fluide étant incompressible, l'équation de conservation des volumes (ou équation de continuité) donne

$$\frac{\partial^2\phi}{\partial x^2}+\frac{\partial^2\phi}{\partial y^2}+\frac{\partial^2\phi}{\partial z^2}=0 \quad \text{ou} \quad \Delta_2\phi=0.$$

Les équations d'Euler conduisent, de plus, à la relation:

$$\frac{p}{\rho}=gz-\frac{\partial\phi}{\partial t}-\tfrac{1}{2}\left[\left(\frac{\partial\phi}{\partial x}\right)^2+\left(\frac{\partial\phi}{\partial y}\right)^2+\left(\frac{\partial\phi}{\partial z}\right)^2\right]$$

qui est également vérifiée en tous les points du fluide, relation où p représente la pression au point $(x,\ y,\ z)$ et ρ la densité du fluide. Si l'on choisit, avec M. Boussinesq, les unités de façon à rendre égales à 1 la densité et l'accélération de la pesanteur, et, de plus, si l'on admet que les carrés des composantes de la vitesse sont négligeables, l'équation précédente devient:

$$p=z-\frac{\partial\phi}{\partial t}. \tag{1}$$

Comme l'action d'une pression constante s'exerçant en tous les points de la surface d'un fluide n'a aucune action sur les mouvements, nous pouvons donc supposer nulle la pression qui règne au-dessus du fluide. Dans ces conditions, on a, à la surface libre:

(2) $$z - \frac{\partial \phi}{\partial t} = 0.$$

Soit $h = F(x, y, t)$ la dénivellation de la surface libre au temps t; l'équation (2) s'écrit

$$h = \frac{\partial \phi}{\partial t},$$

équation dans laquelle on remplace z par 0, en raison de la petitesse de h et de la continuité de ϕ.

Nous avons admis que u, v, h étaient des quantités très petites; admettons qu'il en soit de même pour $\frac{\partial h}{\partial x}$ et $\frac{\partial h}{\partial y}$; nous en déduisons que $w = \frac{dh}{dt}$ et l'équation de condition à la surface

$$\left(\frac{\partial \phi}{\partial t}\right)_{x=0} = \left(\frac{\partial^2 \phi}{\partial t^2}\right)_{z=0}.$$

Les données initiales caractéristiques des ondes d'émersion, exprimant l'état du fluide à l'époque $t=0$, consisteront à poser pour $t=0$, $\phi_0 = 0$ et $h = f(x, y)$, ϕ_0 étant la valeur de ϕ à la surface, et $f(x, y)$ désignant les petites ordonnées primitives connues de la surface.

Dans le cas où la profondeur du liquide, ainsi que les dimensions horizontales, peuvent être regardées comme infiniment grandes, on substitue à la condition $\frac{\partial \phi}{\partial n} = 0$ (dérivée normale de ϕ à la paroi du vase), la condition pour ϕ de s'annuler asymptotiquement lorsque l'on s'éloigne indéfiniment de l'origine des coordonnées.

Si l'on pose $\tau = \frac{\partial \phi}{\partial z} - \frac{\partial^2 \phi}{\partial t^2}$, on remarque que la fonction τ est harmonique, et que $\frac{\partial \phi}{\partial z} - \frac{\partial^2 \phi}{\partial t^2} = 0$ est une équation indéfinie. M. Boussinesq a d'ailleurs montré que $\tau = 0$ et $\frac{\partial \tau}{\partial z} + \frac{\partial^2 \tau}{dt^2} = 0$ sont équivalentes.

La dernière de ces équations devient, en tenant compte de la valeur de τ:

$$\frac{\partial^2 \phi}{\partial z^2} + \frac{\partial^4 \phi}{\partial t^4} = 0$$

qui est l'équation de Cauchy.

[Je prie le lecteur de se reporter, au sujet de cette dernière équation, à deux très intéressantes communications de M. Hadamard aux Comptes Rendus de l'Académie des Sciences (7 et 21 mars 1910), ainsi qu'à sa note du 4 juin 1916 aux C.R. Reale Academia Lincei].

Il y a lieu d'observer qu'à l'époque $t=0$, où ϕ s'annule partout, ainsi que $\frac{\partial \phi}{\partial z}$, la relation $\tau = 0$ se réduit à $\frac{\partial^2 \phi}{\partial t^2} = 0$ pour $t = 0$.

Ces préliminaires exposés, nous allons rappeler les formules donnant les valeurs du potentiel ϕ et de la dénivellation $\frac{\partial \phi}{\partial t}$ dans le cas des ondes cylindriques en supposant le canal de profondeur infinie, et dans le cas d'un milieu indéfini.

Nous mettrons en lumière quelques points qui nous semblent nouveaux, se rapportant au problème des ondes cylindriques et à celui des ondes d'émersion dans un milieu indéfini et dans un canal de profondeur finie ou infinie, à fond horizontal.

Cas des ondes cylindriques.

Poisson, Cauchy, et M. Boussinesq ont déterminé pour un instant donné les points le plus élevés et les plus abaissés de la surface fluide qui sont les sommets des ondes apparentes qui se propagent à cette surface, dans le cas où l'on étudie les ondes produites par l'émersion dans un canal, d'un cylindre dont les génératrices sont perpendiculaires à l'axe du canal et en occupent toute la largeur.

Nous nous proposons ici de rechercher les maxima et les minima de la dénivellation, pour des particules situées dans le plan $z = z_0$, en nous plaçant dans le cas des ondes cylindriques.

Rappelons, à ce propos, que M. Lamb, dans son traité d'Hydrodynamique (3[e] édition, 1906) donne une méthode fort intéressante pour la détermination du potentiel des vitesses (voir formule 13, page 364):

$$\Phi = \frac{gt}{\pi}\left\{\cos\theta - \tfrac{1}{3}\left(\tfrac{1}{2}gt^2\right)\frac{\cos 2\theta}{r^2} + \frac{1}{1.3.5}\left(\tfrac{1}{2}gt^2\right)^2\frac{\cos 3\theta}{r^3} - \dots\right\},$$

qui n'est autre que celle donnée par M. Rousier dans sa thèse, et obtenue par une méthode essentiellement différente, et où r est le rayon vecteur partant de l'origine à la particule, et θ l'angle, formé par la verticale avec le rayon vecteur.

M. Rousier, partant de l'expression de $\frac{\partial \phi}{\partial t}$ donnée par M. Boussinesq, définit l'expression de la dénivellation

$$(3) \qquad h = \frac{S}{\pi r\cos\theta} - \frac{S}{\pi}\left\{\sum_{n=0}^{\infty}\left[\frac{t^{4n+2}\cos(2n+2)\theta}{2^{2n+1}\,r^{2n+2}\,\overline{4n+1}} - \frac{t^{4n+4}\cos(2n+3)\theta}{2^{2n+2}\,r^{2n+3}\,\overline{4n+3}}\right]\right\};$$

r désigne ici la distance d'un point m de la zone d'ébranlement à la particule fluide M (x, z), θ l'angle formé par mM et la verticale, S l'aire de la section droite du cylindre immergé, et le symbole $\overline{4n+1}$ représente le produit $1.3.5\dots(4n+1)$; $\zeta = f(\xi)$ désigne l'équation du cylindre, et l'on suppose ξ négligeable devant x, ou z.

Les sommets et les creux des ondes apparentes correspondent à $\frac{\partial h}{\partial x} = 0$ pour $z = z_0$, et sont définis par l'équation

$$(4) \quad \sin 2\theta - \sum_{n=0}^{\infty}\left[\frac{(2n+2)\sin(2n+3)\theta}{\gamma^{2n+1}\,\overline{4n+1}} - \frac{(2n+3)\sin(2n+4)\theta}{\gamma^{2n+2}\,\overline{4n+3}}\right] = 0 \text{ avec } \frac{1}{\gamma} = \frac{t^2}{2r}.$$

En développant le premier membre de (4) suivant les puissances successives de z, et ensuite suivant les puissances de x, nous trouvons:

$$(5)\qquad 2\left(\frac{t^2x}{2r^2}\right)\frac{x^2}{r^2}f_1+z\frac{x}{r^2}f_2-z^2\left(\frac{t^2x}{2r^2}\right)\frac{1}{r^2}f_3+z^3\left(\frac{t^2x}{2r^2}\right)\frac{1}{xr^2}f_4-\ldots=0,$$

$$(5')\qquad \left(\frac{xz}{r^2}\right)g_1+\frac{x^3}{3!zr^2}g_2-\frac{x^5}{5!z^3r^2}g_3+\frac{x^7}{7!z^5r^2}g_4-\ldots=0;$$

dans (5), les f_i sont des séries entières alternées en $\frac{t^2x}{2r^2}$ et dans (5′), les g_j des séries alternées en $\frac{t^2z}{2r^2}$. Les f_i et g_j ont toutes une infinité de racines positives et distinctes.

Étude du phénomène dans le cas où $\frac{t^2}{2r}$ n'a pas une grande valeur.

Si, après avoir posé, conformément aux notations de Poisson $\left(\frac{t^2x}{2r^2}\right)^2=p$, on fait $z=0$ dans (5), on retrouve $f_1=0$, ce qui donne les points les plus hauts et les plus bas de la surface fluide, en rejetant toutefois les grandes valeurs de p.

Si la particule fluide envisagée correspond à une valeur de z finie, et à une valeur de x qui, tout en étant petite, est supérieure aux valeurs des abscisses de la zone d'ébranlement, il suffit de se reporter à (5)′ et l'on constate que l'équation caractéristique des maxima et des minima est définie par $g_1=0$.

Supposons maintenant x et z finis, mais $\frac{t^2}{2r}$ petit; les maxima et les minima sont donnés par $f_2=0$; on trouve que le sommet de l'onde apparente qui se produira au bout du temps t au point (x, z_0) correspond à:

$$h=\frac{x}{r^2}\left(\frac{p^{1/2}}{1}-\frac{p^{3/2}}{5}+\frac{p^{5/2}}{9}-\frac{p^{7/2}}{13}+\ldots\right)+\frac{z_0}{r^2}\left[1+\sum_{n=1}^{\infty}C_{2n+1}^{2n}\frac{p^n(-1)^n}{4n+1}\right]$$
$$-\frac{t^2z_0^2}{2r^4}\left[1+\sum_{n=1}^{\infty}C_{2n+2}^{2n}\frac{p^n(-1)^n}{4n+1}\right]+\ldots,$$

où C_m^n représente le nombre des combinaisons de m objets n à n, et où $p=\left(\frac{t^2x}{2r^2}\right)^2$ est une racine de $f_2=0$. Supposons z_0 fini, x très grand, toutengar dant à $\frac{t^2}{2r}$ une valeur petite; dans ce cas, les maxima et les minima sont définis par $f_1=0$.

Étude du phénomène dans le cas où $\frac{t^2}{2r}$ a une grande valeur.

On suppose que les petites valeurs de x correspondent à des points situés en dehors de la zone d'ébranlement initial; elles peuvent être telles qu'elles rendent l'expression $\frac{t^2x}{2r^2}$ petite ou finie; le problème ne présente pas de difficultés spéciales.

Lorsque x est fini et inférieur à z, on a recours à une méthode d'approximations successives en partant de

$$g_1+\frac{x^2}{3!z^2}g_2=0.$$

Si, au contraire, $z<x$ on emploie le même procédé en utilisant l'équation $2\lambda f_1-\frac{z}{x^2}f_2-\frac{z^2}{x^2}\lambda f_3-\ldots=0$ avec $\lambda=\frac{t^2}{2r^2}$ et en prenant pour premières valeurs approchées de λ les racines de $f_1=0$.

Ondes par émersion en milieu indéfini

L'examen du potentiel des vitesses de Poisson*, où l'on fait $g=1$,

$$\phi=-\frac{t}{\pi^2}\iint_{-\infty}^{+\infty}\left(\frac{\partial Z}{\partial z}+\frac{t^2}{3!}\frac{\partial^2 Z}{dz^2}+\frac{t^4}{5!}\frac{\partial^3 Z}{\partial z^3}+\ldots\right)F(\xi,\eta)d\xi d\eta,$$

(ξ, η représentant les coordonnées horizontales de la zone d'émersion, et $F(\xi,\eta)$ les coordonnées primitivement connues de la surface) avec

$$Z=\frac{\pi}{2\sqrt{z^2+(x-\xi)^2+(y-\eta)^2}},$$

montre que la série entre parenthèses sous le signe d'intégration sera d'autant plus convergente que le rapport $\frac{t^2}{\sqrt{z^2+(x-\xi)^2+(y-\eta)^2}}$ sera plus petit, et que cette série sera en défaut, quelque petit que soit le temps t, si l'on considère un point de la surface fluide pris dans l'étendue de l'ébranlement primitif.

Il y a donc lieu de chercher ce qui se passe en réalité aux points de la surface dont les coordonnées x et y sont respectivement comparables à ξ et η.

Nous guidant sur la méthode préconisée par M. Boussinesq dans le cas de deux dimensions (cylindre immergé dont les génératrices sont perpendiculaires à l'axe du canal), nous adopterons pour l'expression de la dénivellation

$$h=\frac{4}{\pi^2 t}\frac{d}{dt}\int_0^{\frac{\pi}{2}}d\mu\int_0^{2\pi}d\theta\int_0^{\infty}\left(\frac{t^2\cos\mu}{4r}\right)^{3/2}\psi'\left(\frac{t^2\cos\mu}{4r}\right)F(x+r\cos\theta,\ y+r\sin\theta)dr^*,$$

où ψ est définie par

$$\psi(\gamma)=\int_0^{\sqrt{\gamma}}\sin(\gamma-\mu^2)d\mu=\tfrac{1}{2}\sqrt{\frac{\pi}{2}}(\sin\gamma-\cos\gamma)+\int_0^{\infty}e^{-2m\sqrt{\gamma}}\cos m^2dm.$$

*Voir, p. 144 de son *Mémoire sur la Théorie des Ondes.*

Nous remplacerons le corps par son paraboloïde osculateur et substituerons alors $H\left[1-\left(\frac{x^2}{a^2}+\frac{y^2}{b^2}\right)\right]$ à $F(x, y)$.

Dans l'expression de la dénivellation, qui peut s'écrire

$$h=\frac{4}{\pi^2 t}\frac{d}{dt}\int_0^{\frac{\pi}{2}}d\mu\int_0^{\frac{\pi}{2}}d\theta\int_0^{\infty}\left(\frac{t^2\cos\mu}{4r}\right)^{3/2}\psi'\left(\frac{t^2\cos\mu}{4r}\right)F(x\pm r\cos\theta, y\pm\gamma\sin\theta)dr,$$

la fonction F n'a de valeur que si

$$-a<x\pm r\cos\theta<a,\quad -b<y\pm r\sin\theta<b.$$

A titre indicatif, nous n' étudierons ici que le cas $x<a$, $y<b$, et remarquons qu'en la circonstance:

et
$$h=\frac{4}{\pi^2 t}\frac{d}{dt}\int_0^{\frac{\pi}{2}}d\mu\int_0^{\frac{\pi}{2}}d\theta\left(\sum\int_0^{r_i'}\right)\text{ avec } (i=1, 2, 3, 4),$$

$$r_1'=\sqrt{(a-x)^2+(b-y)^2},\qquad r_2'=\sqrt{(a+x)^2+(b-y)^2},$$
$$r_3'=\sqrt{(a-x)^2+(b+y)^2},\qquad r_4'=\sqrt{(a+x)^2+(b+y)^2}.$$

Les fonctions F entrant dans les deux premiers éléments de Σ sont:

$$F(x+r\cos\theta, y+r\sin\theta)\text{ et } F(x-r\cos\theta, y+r\sin\theta),$$

et celles entrant dans les autres sont:

$$F(x+r\cos\theta, y-r\sin\theta)\text{ et } F(x-r\cos\theta, y-r\sin\theta).$$

Si l'on s'en tient à une première approximation, et si l'on prend pour valeur approchée de F

$$(F)=\frac{\text{volume du corps immergé}}{\pi ab},$$

on voit que, dans le cas envisagé, on a

$$h=\frac{1}{\sqrt{2}\,\pi t}(F)\frac{d}{dt}\int_0^{\frac{\pi}{2}}t^{2'}\cos\mu d\mu\left(\sum\int_{\alpha_i}^{\infty}\right)\text{ avec } \alpha_i^2=\frac{t^2\cos\mu}{2r_i}\text{ et }\int_{\alpha_i}^{\infty}=\int_{\alpha_i}^{\infty}\psi'\left(\frac{\alpha^2}{2}\right)d\alpha.$$

Les α_i étant des nombres positifs considérables, on peut, ainsi que l'a fait M. Boussinesq, substituer $-\frac{1}{\alpha}\psi\left(\frac{\alpha^2}{2}\right)$ à $\int_\alpha^\infty$ et par suite, donner à h une forme relativement simple. L'examen détaillé des divers cas sera effectué dans un mémoire spécial.

En tous cas, ces résultats peuvent être résumés ainsi qu'il suit:

Suivant que la particule fluide M occupe dans l'angle yox des axes de la section à fleur d'eau les positions:

$$(x>a, y>b),\ (x>a, y<b)\text{ ou }(x<a, y>b),\ (x<a, y<b),$$

*Voir formule (238), p. 644. *Traité des Potentiels* de M. Boussinesq.

elle peut être considérée, suivant les cas, comme soumise à 1, 2 ou 4 trains absolus d'ondes.

Si la particule fluide est dans l'une des positions $(x>a,\ y<b)$, $(x<a,\ y>b)$ les deux trains sont distincts.

Si la particule fluide est à l'intérieur de la section, les quatre trains sont distincts toutes les fois qu'elle n'est pas sur l'un des axes; si elle est sur l'un des axes, les 4 trains se réduisent à deux trains distincts, et enfin, si elle est au centre, les 4 trains d'ondes sont identiques.

Si la section à fleur d'eau est circulaire, toutes les fois que la particule est à l'intérieur de la section, les 4 trains se réduisent à deux trains doubles distincts.

Cette remarque particulière incite donc à penser que les calculs doivent se simplifier d'une manière considérable toutes les fois que le corps immergé sera de révolution; elle a été pleinement justifiée par la suite.

Introduction de termes secondaires dans le potentiel des vitesses.

L'étude des ondes d'émersion dans un canal de profondeur infinie se trouvant, grâce à l'emploi des images de Kelvin* ramenée à celle des ondes dans un milieu indéfini, il est facile d'examiner pour un tel milieu l'influence perturbatrice motivée par l'intervention d'une coordonnée η de la zone d'émersion, non négligeable par rapport à la coordonnée y du point pour lequel on veut, à un instant donné, évaluer la grandeur de la dénivellation.

En supposant ξ négligeable devant x, et η non négligeable devant y, on est conduit à calculer l'intégrale

$$Z_1 = \iint \frac{F(\xi,\ \eta)d\xi d\eta}{\sqrt{z^2+x^2+(y-\eta)^2}} = \iint \frac{F d\xi d\eta}{\sqrt{R^2-2y\eta+\eta^2}}\,.$$

En substituant au dénominateur son développement en série, et en assimilant la partie du corps immergé à son paraboloïde osculateur, de telle façon que $F(\xi,\eta) = H\left(1-\frac{\xi^2}{l^2}-\frac{\eta^2}{l'^2}\right)$ on trouve que l'intégrale ci-dessus, après le changement de variable $\xi = ls\cos\psi$, $\eta = l's\sin\psi$ devient:

$$Z_1 = Hll'\int_0^1 \frac{(1-s^2)sds}{R}\int_0^{2\pi} d\psi\left[1+\frac{\eta}{2}\left(\frac{2y-\eta}{R^2}\right)+\ldots\right].$$

Si l'on fait abstraction des termes du développement en série apparaissant dans la parenthèse, où figurent $\frac{1}{R^4}$, $\frac{1}{R^6}$, . . .; on trouve l'expression Z_2 approchée de Z_1:

$$Z_2 = Hll'\left(\frac{1}{2R}-\frac{l'^2}{6R^3}\right) = \pi V\left(\frac{1}{2R}-\frac{l'^2}{6R^3}\right),\ \text{avec } \pi V = Hll'.$$

*Voir le Mémoire de M. Boussinesq: *Sur une importante simplification de la Théorie des Ondes que produisent à la surface d'un liquide lémersion d'un solide ou l'impulsion d'un coup de vent.*—Annales de l'École Normale Supérieure (1910).

Le potentiel des vitesses ayant pour valeur

$$\phi=-\frac{V}{\pi^2}\left[t\frac{\partial Z_2}{\partial z}+\frac{t^3}{3!}\frac{\partial^2 Z_2}{\partial z^2}+\ldots+\frac{t^{2n+1}}{(2n+1)!}\frac{\partial^{n+1}Z_2}{\partial z^{n+1}}+\ldots\right],$$

on voit de suite apparaître les termes provenant de $\frac{d^n}{dz^n}\left(\frac{1}{R^3}\right)$ pour lesquels on a

$$\frac{d^n}{dz^n}\left(\frac{1}{R^3}\right)=\frac{(n+1)(n+2)P_n+2(n+2)\cos\theta P_n'+\cos^2\theta P_n''}{R^{n+3}}=\frac{Q_n}{R^{n+3}},$$

avec $\cos\theta=\frac{z}{R}$ et P_n étant le polynome de Legendre d'ordre n.

L'équation $Q_n=0$ est de degré n en $\cos\theta$ et ne renferme que des termes ayant la parité de n; elle a, comme $P_n=0$, toutes ses racines réelles et distinctes comprises entre -1 et $+1$.

L'influence du terme en $\frac{1}{R^{2p+1}}$ conduirait à une expression $({}_1Q_n)$ analogue à Q_n:

$${}_1Q_n=AP_n+BxP_n'+\ldots+Lx^{2p}P_n^{(2p)};$$

l'équation ${}_1Q_n=0$ a aussi toutes ses racines réelles et distinctes.

Si l'on fait abstraction des termes en $\frac{1}{R^3}$, on retrouve le développement classique de ϕ.

Sur les ondes d'émersion dans un canal de longueur donnée et de profondeur infinie.

Le potentiel des vitesses peut être représenté par la série de MacLaurin:

$$\phi=th_0+\frac{t^3}{3!}\frac{\partial h_0}{\partial z}+\ldots\ldots+\frac{t^{2n+1}}{(2n+1)!}\frac{\partial^n h_0}{\partial z^n}+\ldots\ldots, \tag{1}$$

où h_0 qui désigne la valeur initiale au point (x, y, z) de $\frac{\partial\phi}{\partial t}$ ou de la dénivellation répond aux conditions:

$$\Delta_2 h_0=0,\ \frac{dh_0}{dn}=0$$

sur les parois, et est telle que pour $z=z_0$ elle se réduise à $f(x, y)$.

Si l'on se trouve dans le cas limite d'un canal de largeur a, on peut évaluer la dénivellation, en introduisant le corps réel et ses $2k$ images, et adoptant comme aire totale d'émersion l'ensemble de l'aire vraie et de ses $2k$ images, à condition de faire croître k indéfiniment.

Alors que l'on ne faisait apparaître que la fonction

$$Z_1=\iint_{-\infty}^{+\infty}\frac{f(\xi,\eta)\,d\xi d\eta}{r}\ \text{avec}\ r=\sqrt{(x\quad\xi)^2+(y-\eta)^2+z^2},$$

pour un seul corps dans l'hypothèse d'un canal indéfini, il faut maintenant considérer

$$Z_1+\sum_{j=-k}^{+k}{}' \iint \frac{f[\xi,\ (-1)^j\eta+ja]d\xi d\eta}{\sqrt{(x-\xi)^2+[y-ja-(-1)^j\eta]^2+z^2}},$$

l'expression Σ' s'étendant à toutes les valeurs de j de $-k$ à $+k$, sauf zéro.

En supposant tout d'abord que, dans le cas d'un milieu indéfini, l'on s'en tienne à un seul élément dq immergé à l'origine des coordonnées, et que ξ et η soient négligeables devant (x, y, z), le potentiel est défini par

$$\phi_0=\frac{dq}{2\pi}\sum_{n=0}^{\infty}\frac{(-1)^n t^{2n+1}}{(n+2)(n+3)\dots(2n+1)}\,P_{n+1}(\cos\theta);$$

dans le cas d'un canal de largeur a, on associera à ϕ_0 les potentiels

$$\phi_{-k},\ \dots,\ \phi_{-1},\ \phi_1,\ \dots,\ \phi_k.$$

La dénivellation aura pour valeur

$$h=\sum_{j=-k}^{+k} h_j,$$

avec:

$$h_j=\frac{dq}{2\pi r_j^2}\left[P_1(\cos\theta_j)+\dots+\frac{(-1)^n t^{2n}}{(n+2)\dots 2nr_j^n}\,P_{n+1}(\cos\theta_j)+\dots\right],$$

à condition de faire croître indéfiniment le nombre des images.

En définitive, tout se passe comme si la particule fluide située en $M(x, y, z)$ subissait l'influence du corps, puis celle de chacune des images, et l'on voit que le mouvement est uniformément accéléré sur toute droite joignant M au corps et à ses images.

Dénivellation à la surface.

Cette dénivellation, dans le cas d'un milieu indéfini, de profondeur infinie, est représentée par l'expression

$$\frac{dq}{2\pi}\sum_{m=0}^{\infty}(-1)^{3m+2}\frac{\left(\frac{t^2}{2}\right)^{2m+1}}{(2.4\dots 2m)(2m+3)\dots(4m+1)r^{2m+3}};$$

la dénivellation de la particule fluide (x, y) située dans un canal à parois parallèles distantes de a, aura pour valeur une série dont le terme de rang m sera défini par l'expression

$$\frac{dq}{2\pi}(-1)^{3m+2}\frac{\left(\frac{t^2}{2}\right)^{2m+1}}{(2.4.6\dots 2m)(2m+3)\dots(4m+1)}\sum_{j=-k}^{+k}\frac{1}{r_j^{2m+3}}.$$

En faisant un calcul simple, on voit qu'en première approximation, le m^e terme de h a pour valeur:

$$\frac{dq}{\pi a R_1}(-1)^{3m+2}\left(\frac{t^2}{2R_1}\right)^{2m+1}\frac{1}{1.3.5\ldots(4m+1)},$$

avec $R_1=\sqrt{x^2+z^2}$; il en résulte que la formule représentative de h se ramène à celle trouvée par Poisson, Cauchy et M. Boussinesq, dans le cas de 2 dimensions, par la substitution de $\left(\frac{S}{\pi R_1}\right)$ à $\left(\frac{dq}{\pi R_1 a}\right)$.

Dénivellation au point (x, y, z).

Cette dénivellation a pour valeur:

$$h=\sum_{j=-k}^{+k} h_j, \text{ avec } h_j=\frac{dq}{2\pi}\sum_{n=0}^{\infty}(-1)^n\frac{t^{2n}}{r_j^{n+2}}\frac{(n+1)!}{2n!}P_{n+1}\cos(\theta_j).$$

On trouve, en première approximation, que h a pour valeur:

$$h=\frac{dq}{2\pi}\sum_{n=0}^{\infty}(-1)^n\left(\frac{t^2}{R}\right)^n\frac{2}{aR}\frac{n+1}{1.3.5\ldots(2n+1)}\sum_{p=0}^{p\leq\frac{n+1}{2}}\frac{(-1)^p}{2^{2p}}\frac{1}{p!}\frac{(n-p)!}{(n+1-2p)!}\cos^{n+1-2p}\theta$$

avec $R=\sqrt{x^2+z^2}$ et $\cos\theta=\frac{z}{R}$.

Nous avons vu, précédemment, que le potentiel des vitesses peut être représenté par:

$$\phi=th_0+\sum_{n=1}^{\infty}\frac{t^{2n+1}}{(2n+1)!}\frac{\partial^n h_0}{\partial z^n},$$

où h_0 a le sens défini plus haut.

Faisant état de suggestions tirées de l'étude d'un premier mémoire de M. Appell (Acta Mathematica T. VIII) et d'un deuxième mémoire du même auteur *Développements en séries trigonométriques de certaines fonctions périodiques vérifiant l'équation* $\Delta F=0$, Journal de Mathématiques pures et appliquées, T. 3, 4^e série, 1887), il a été possible de donner une expression rigoureuse de ϕ, et, par suite, de la dénivellation, en supposant le corps réduit à un élément dm.

En effet, dans ce cas, l'expression définie ci-dessous

$$h_0=dm\left[\frac{1}{r}+\sum_{k=1}^{\infty}\left(\frac{1}{\sqrt{x^2+z^2+(y-ka)^2}}-\frac{1}{ka}\right)+\sum_{k=1}^{\infty}\left(\frac{1}{\sqrt{x^2+z^2+(y+ka)^2}}-\frac{1}{ka}\right)\right]$$

répond aux conditions exigées

$$\Delta_2 h_0=0,$$

$$\left(\frac{\partial h_0}{\partial n}\right)=0,$$

aux parois.

Or, l'expression

$$C_1(x, y, z, a)=\frac{1}{\sqrt{x^2+y^2+z^2}}+\sum_{k=1}^{\infty}\left[\frac{1}{\sqrt{(y-ka)^2+(x^2+z^2)}}-\frac{1}{\sqrt{k^2a^2}}\right],$$

qui intervient dans l'expression de h_0, est susceptible d'être développée en série trigonométrique ainsi qu'il suit:

$$C_1=A_0+\sum_{\nu=1}^{\infty}A_\nu \cos\frac{2\nu\pi y}{a},$$

où $A_0=\frac{-1}{a}\log u+B_0$, avec $u=x^2+z^2$, et B_0 étant une constante indépendante de u,

$$A_\nu=\frac{2}{a}\int_0^\infty e^{-tu-\frac{\pi^2\nu^2}{a^2t}}\frac{dt}{t}.$$

Cas d'un canal de largeur et de longueur limitées, et de profondeur infinie.

Si l'on se trouve dans ce cas, l'on est conduit à représenter le potentiel h_0 par l'expression $d_mC_2(x, y, z; a, b)$, avec

$$C_2(x+a, y, z; a, b)=C_2(x, y+b, z; a, b)=C_2(x, y, z);$$

de plus, cette fonction C_2 doit être telle que:

$$\left(\frac{\partial C_2}{\partial x}\right)_{x=\pm\frac{a}{2}}=0, \qquad \left(\frac{\partial C_2}{\partial y}\right)_{y=\pm\frac{b}{2}}=0.$$

Or, le développement

$$C_2=\sum_{\mu=0}^{+\infty}\sum_{\nu=0}^{+\infty}A_{\mu,\nu}\cos\frac{2\mu\pi x}{a}\cos\frac{2\nu\pi y}{b},$$

avec:

$$A_{\mu,\nu}=\frac{4e^{-2\pi z\sqrt{\frac{\mu^2}{a^2}+\frac{\nu^2}{b^2}}}}{ab\sqrt{\frac{\mu^2}{a^2}+\frac{\nu^2}{b^2}}}+B_{\mu,\nu},$$

répond à la question, à la condition de prendre $B_{\mu,\nu}$ nulle pour tout système de valeurs de μ, ν associées, non nulles à la fois et

$$B_{0,0}=C_2(0, 0, z; a, b)+\frac{2\pi z}{ab}-\sum_{\mu,\nu}{}'\frac{4e^{-2\pi z\sqrt{\frac{\mu^2}{a^2}+\frac{\nu^2}{b^2}}}}{ab\sqrt{\frac{\mu^2}{a^2}+\frac{\nu^2}{b^2}}}.$$

Si le corps est immergé dans un milieu limité par les parois,

$$x=\pm\frac{a}{2},\quad y=\pm\frac{b}{2},\quad z=c,$$

il suffit de prendre pour expression de h_0

$$dm[C_2(x,\ y,\ z;\ a,\ b)+C_2(x,\ y,\ 2c-z;\ a,\ b)]$$

car la quantité entre crochets est une fonction harmonique dont les dérivées $\dfrac{\partial h_0}{\partial(x, y, z)}$ sont nulles pour $x=\pm\dfrac{a}{2}$, $y=\pm\dfrac{b}{2}$, $z=c$.

On voit donc que l'on peut, grâce à la théorie des images, trouver rigoureusement la valeur du potentiel des vitesses dans un bassin parallélipipédique, de profondeur finie ou infinie, *mais il me semble bon de rappeler que l'on peut aboutir par une voie différente.* En effet, choisissons avec Poincaré une fonction $\phi\ (x, y, z, t)$ se mettant sous la forme $\phi_1(x,\ y,\ t)\phi_2(z)$, vérifiant les conditions:

$$\Delta_2\phi=0,\ \phi=0,$$

pour $t=0$,

$$\frac{\partial\phi}{\partial x}=0$$

sur les parois verticales ($x=\pm a,\ y=\pm b$), et sur le fond $z=c$:

$$\left(\frac{\partial\phi}{\partial t}\right)_{z=t=0}=f(x,\ y),\quad \left(\frac{\partial\phi}{\partial z}\right)_{z=0}=\left(\frac{\partial^2\phi}{\partial t^2}\right)_{z=0}.$$

On trouve ainsi que ϕ est constitué par 4 séries, dont la première ϕ_1 est définie par l'expression:

$$\phi_1=\sum_{\mu=0}^{\infty}\sum_{\nu=0}^{\infty}{}_1A_{\mu,\,\nu}\sin(m_1t)\cos\frac{\mu\pi x}{a}\cos\frac{\nu\pi y}{b}\,\frac{e^{\frac{z}{k_1}}+e^{\frac{2c-z}{k_1}}}{m_1(1+e^{\frac{2c}{k_1}})},$$

en ayant soin de choisir pour k_1 la valeur:

$$\frac{-1}{\pi\sqrt{\dfrac{\mu^2}{a^2}+\dfrac{\nu^2}{b^2}}},\ \text{avec}\ m_1^2=\frac{e^{\frac{2c}{k_1}}-1}{k_1(1+e^{\frac{2c}{k_1}})},$$

et:

$${}_1A_{\mu,\,\nu}=\frac{1}{ab}\int_{-a}^{+a}\int_{-b}^{+b}f(\xi,\ \eta)\cos\frac{\mu\pi\xi}{a}\cos\frac{\nu\pi\eta}{b}\,d\xi d\eta.$$

Quant aux trois autres séries, on les formerait facilement, eu égard aux considérations précédentes, et en se rappelant que les ${}_2A_{\mu,\nu}$, ${}_3A_{\mu,\nu}$, ${}_4A_{\mu,\nu}$ correspondent aux associations suivantes:

$$\begin{bmatrix}\cos\dfrac{\mu\pi\xi}{a}\\[2ex] \sin\dfrac{(2\nu+1)}{2b}\pi\eta\end{bmatrix},\quad \begin{bmatrix}\sin\dfrac{(2\mu+1)\pi\xi}{2a}\\[2ex] \cos\dfrac{\nu\pi y}{b}\end{bmatrix},\quad \begin{bmatrix}\sin\dfrac{2\mu+1}{2a}\pi\xi\\[2ex] \sin\dfrac{2\nu+1}{2b}\pi\eta\end{bmatrix}.$$

On peut, dans la fonction ϕ ainsi formée, faire c infini, c et b infinis, et l'on retrouve les formes classiques du potentiel des vitesses données par Poisson.

LES LOIS DE L'ÉLASTICITÉ EN COORDONNÉES QUELCONQUES

PAR M. LÉON BRILLOUIN,
Collège de France, Paris, France.

1. *Introduction.*

Il est très important de savoir formuler les lois de la physique sous une forme qui soit indépendante du type de coordonnées employées, c'est ce qu'a permis de faire le calcul différentiel absolu, dont la fécondité et la facilité d'emploi sont illustrées par le rôle important qu'il joue dans la Relativité genéralisée.

Celle-ci fait usage de quatre dimensions, les trois coordonnées d'espace et le temps y étant introduits très symétriquement. Mais, avant d'écrire les équations de la physique dans cet univers quadridimensionnel, il est indispensable de les avoir formulées complètement dans l'espace ordinaire. Ceci a été fait rapidement par les fondateurs de la Relativité, mais bien des chapitres sont à reprendre et à préciser. Tel est le point de vue auquel s'était placé M. P. Langevin, dans ses leçons au Collège de France (1922-1923) et ses Conférences-Rapports. Je suivrai les mêmes méthodes, pour en faire l'application aux problèmes de l'élasticité.

Ce chapitre est, d'ailleurs, un des plus délicats; c'est à ce propos que les physiciens découvrirent ces grandeurs que l'on nomme *tenseurs*; et cette appellation même garde la trace de son origine, le premier tenseur reconnu ayant été celui des tensions élastiques. Il est aussi curieux de remarquer que Lamé avait déjà pensé pouvoir simplifier l'écriture des lois de l'élasticité par l'emploi de surfaces courbes de coordonnées; et cette idée est à l'origine de ses travaux fondamentaux sur les coordonnées curvilignes. Malgré des recherches nombreuses, nous n'avons pas, actuellement, l'énoncé des lois de l'élasticité sous une forme tensorielle valable en coordonnées quelconques. J'ai essayé de combler cette lacune, en indiquant nettement les propriétés géométriques qu'il convient d'attribuer aux diverses grandeurs physiques. Les lois de l'élasticité ont été exposées sous une forme excellente dans un mémoire de E. et F. Cosserat* pour le cas des coordonnées rectilignes rectangulaires. Ces auteurs ont mis au point et très bien précisé les raisonnements classiques sans faire d'hypothèses simplificatrices qui en réduisent la portée; leur travail me servira de guide, et j'indiquerai partout les correspondances de notations et de formules. A la fin de ce même article, E. et F. Cosserat ont établi les formules d'élasticité en coordonnées quelconques, mais ils emploient une méthode très différente de celle des tenseurs; c'est en se servant d'un trièdre mobile de référence qu'ils ont

*E. et F. Cosserat. *Sur la théorie de l'élasticité.* Annales de la Faculté des Sciences de Toulouse, t. 10 (1896), p. I_1-I_{116}.

écrit les relations valables pour des coordonnées différentes; en se reportant à leurs formules, on sera frappé de leur complexité; l'écriture tensorielle permet, au contraire, de garder aux équations exactement la même forme, quel que soit le système de coordonnées, et de présenter les formules sous un aspect très bref et ramassé.

2. *Notations: Tenseurs et pseudo-tenseurs.*

Soient $x^i (i=1, 2, 3)$ les coordonnées qui nous servent à fixer la position d'un point, et

(1) $$ds^2 = g_{ik} dx^i dx^k, \qquad (i, k=1, 2, 3),$$

le carré de la distance de deux points voisins x^i et $x^i + dx^i$. La notation tensorielle* explicite nettement, à propos de chaque grandeur, la manière dont varie celle-ci lors d'un changement de coordonnées. Soient $\bar{x}^r$ de nouvelles coordonnées, définies par les fonctions $x^i = f^i(\bar{x}^r)$, nous avons entre les différentielles des x et des $\bar{x}$, les relations:

(2) $$dx^i = a_r^i \, d\bar{x}^r, \quad a_r^i = \frac{\partial x^i}{\partial \bar{x}^r},$$

d'où nous tirons, par inversion,

(2^{bis}) $$d\bar{x}^r = \beta_i^r \, dx^i.$$

Si l'on prend le déterminant des a_r^i, qui n'est autre que le déterminant fonctionnel

$$\frac{D(x^1, x^2, x^3)}{D(\bar{x}^1, \bar{x}^2, \bar{x}^3)} = \Delta,$$

le coefficient β_i^r est égal au mineur de a_r^i, divisé par Δ. Un tenseur t_{ab}^{ef} donne, après le changement de coordonnées, un tenseur $\bar{t}_{pq}^{rs}$, et les formules de transformation sont:

(3) $$t_{ab}^{ef} = a_r^e a_s^f \beta_a^p \beta_b^q \bar{t}_{pq}^{rs}.$$

Ce tenseur est dit deux fois covariant (indices a et b) et deux fois contrevariant (indices e et f).

Considérons une grandeur physique représentée par un tenseur t; il est commode, suivant les cas, de prendre les composantes ayant une variance

*Suivant un usage établi, nous n'écrivons pas les signes de sommation; lorsque, dans une formule, un même indice j figure deux fois; une fois en haut (contrevariance) et une fois en bas (covariance) il est convenu que l'on doit prendre la somme de tous les termes obtenus en faisant $j=1$, 2 et 3.

Je renvoie pour de plus amples explications, aux traités classiques, ou aux exposés généraux tels que:

Eddington, *Espace, temps, gravitation.*

H. Weyl; *Temps, espace, matière*, trad. française. Blanchard, Paris.

Galbrun; *Introduction géométrique à la théorie de la relativité*, Gauthier-Villars, Paris.

Cartan; *Leçons sur les invariants intégraux*, Hermann, Paris.

G. Darmois; *Eléments de géométrie des espaces;* Ann. de Phys. 10° série t. 1 (1924), p. 1-88.

déterminée; on passe des unes aux autres par déplacement des indices, suivant les formules:

(4)
$$t_{abc}^{f}=g_{ec}t_{ab}^{ef}, \qquad t_{ab}^{ef}=g^{ce}t_{abc}^{f}.$$

Les g_{ec} sont les coefficients de la forme quadratique fondamentale (1). Les g^{ec} sont les mineurs de g_{ec} divisés par $|g|$, déterminant des g; on en déduit les relations suivantes:

(5)
$$g_{ik}g^{il}=\begin{cases}0 \text{ si } k\neq l\\ 1 \text{ si } k=l\end{cases}.$$

Les conditions de symétrie des tenseurs jouent un rôle important: un tenseur qui a deux indices de même variance est dit *symétrique* si l'on a

$$t^{ab}=t^{ba}.$$

Cette symétrie n'a pas d'expression simple si l'on prend les composantes mixtes t_e^a, mais se retrouve sur les composantes covariantes $t_{ef}=t_{fe}$; les g_{ik} sont un exemple de tenseur symétrique.

Un tenseur est dit *symétrique gauche*, si l'on a entre ses composantes des relations

$$t^{ab}=-t^{ba},$$

ce qui implique que les composantes diagonales t^{aa} soient nulles; cette propriété n'a pas de correspondance simple pour les composantes mixtes et se retrouve sur les composantes covariantes. On a un exemple simple de tenseur symétrique gauche, si l'on forme le *produit extérieur* (produit vectoriel)* de deux vecteurs u et v:

$$t^{ab}=[u^a.v^b]=u^av^b-u^bv^a.$$

Ces tenseurs symétriques gauches permettent de définir de nouvelles grandeurs, les *pseudo-tenseurs*, dont l'introduction systématique a permis à M. P. Langevin de classer très heureusement les grandeurs physiques. Cette conception me sera très utile, car je montrerai que le tenseur des efforts, en élasticité est en réalité un pseudo-tenseur; aussi insisterai-je un peu sur ce point. Dans l'espace à trois dimensions un tenseur symétrique gauche t^{ab} n'a que trois composantes distinctes; le tableau des composantes s'écrit:

$$\begin{matrix} 0 & t^{12} & t^{13} \\ -t^{12} & 0 & t^{23} \\ -t^{13} & -t^{23} & 0. \end{matrix}$$

Si l'on choisit un certain ordre pour les indices a, b, c, l'ordre 1, 2, 3, par exemple, on peut faire correspondre aux 2 indices a, b un unique indice c et écrire :

(7)
$$t^{ab}=\tau, \quad \begin{cases} t^{12}=\tau_3=-t^{21}, \\ t^{23}=\tau_1=-t^{32}, \\ t^{31}=\tau_3=-t^{13}. \end{cases}$$

*V. Cartan, *loc. cit.*

Les trois composantes τ_c forment un pseudo-tenseur covariant du premier ordre; on vérifie facilement qu'il jouit de propriétés analogues à celles d'un tenseur covariant, les formules de transformation étant

$$\tau_c = \Delta \beta_c^k \bar{\tau}_k. \tag{8}$$

On a en effet:

$$\tau_c = t^{ab} = \sum_i \sum_j \alpha_i^a \alpha_j^b \bar{t}^{ij} = \sum_k (\alpha_i^a \alpha_j^b - \alpha_j^a \alpha_i^b) \bar{\tau}_k,$$

k étant le troisième indice qui correspond à i et j. La parenthèse représente le mineur de α_k^c dans le déterminant $\Delta = |\alpha_k^c|$. D'après la définition des β, ce mineur a pour valeur $\Delta\beta_c^k$. L'apparition du déterminant fonctionnel Δ différencie ces formules de celles qui se rapportent à un vrai tenseur. Si l'on passe des coordonnées rectilignes rectangulaires à d'autres du même type, sans changer les unités de longueur, Δ est égal à $+1$, si les sens de rotations sont les mêmes pour les deux systèmes d'axes,et à -1 dans le cas contraire; on voit donc que le pseudo-tenseur à un indice correspond à un vecteur axial, le tenseur à un indice étant un vecteur polaire.

D'un tenseur deux fois covariant, et symétrique gauche s_{ab} on déduirait de même un pseudo-tenseur contrevariant $\Sigma^c = s_{ab}$ pour lequel les formules de transformation seraient:

$$\Sigma^c = \frac{1}{\Delta} \alpha_k^c \bar{\Sigma}^k. \tag{9}$$

Parmi ces pseudo-tenseurs, il est important de distinguer, autrement que par la place des indices, les deux types précédents; cette distinction est indispensable pour les pseudo-tenseurs à plusieurs indices; ceux qui se transforment comme Σ^c sont déjà appelés *densités tensorielles**; si l'on se rappelle la relation de transformation de g:

$$|\bar{g}| = \Delta^2 |g|, \tag{10}$$

on voit que l'on obtient un tenseur vrai en multipliant une densité tensorielle par $\frac{1}{\sqrt{g}}$.

Pour les grandeurs qui se transforment comme τ_c, il faudra adopter un nom différent, tel que celui de *capacité tensorielle*†. On obtient un tenseur en multipliant une capacité tensorielle par $\sqrt{g}$. Afin de distinguer ces deux catégories, nous adopterons les lettres grecques ou cursives majuscules pour les densités tensorielles et minuscules pour les capacités tensorielles; les lettres romaines seront réservées aux tenseurs.

Pour les tenseurs à trois indices, il est utile d'établir une correspondance du même genre. Soit t^{ijk} un tenseur symétrique gauche pour toutes les permutations des indices; les valeurs de ses composantes se réduisent à $+t$, 0 et $-t$;

*V. p. ex. H. Weyl; *Temps, espace, matière*; trad. française, p. 94, p. 100.

†Ces grandeurs ne semblent pas avoir été systématiquement classées jusqu'à présent; leur emploi est pourtant très utile.

il ne subsiste donc, numériquement, qu'une valeur $t=\tau$; pour déterminer τ, il suffit de se donner l'ordre $i, j, k=1, 2, 3$ par exemple, pour les indices de la composante spéciale que l'on choisit. Les formules de changement de coordonnées se réduisent à

$$\tau=\Delta\bar{\tau}. \tag{11}$$

C'est un pseudo-scalaire, ou mieux une *capacité scalaire*. Nous aurons un élément de ce type si nous considérons le produit extérieur de trois vecteurs $[a^i . b^j . c^k]$, qui représente l'élément de volume construit sur ces trois vecteurs. Les composantes sont:

$$\pm\begin{vmatrix} a^1 & a^2 & a^3 \\ b^1 & b^2 & b^3 \\ c^1 & c^2 & c^3 \end{vmatrix}, \text{ et zéro.} \tag{12}$$

Choisissons l'ordre 1, 2, 3, c'est-à-dire prenons le signe + devant ce déterminant et nous avons la capacité τ, qui correspond à ce volume.

Une densité de volume Σ, au sens physique ordinaire du mot, est une *densité scalaire*, ou pseudo-scalaire du second type. La densité se modifie, lors d'un changement d'axe, suivant la formule

$$\Sigma=\frac{1}{\Delta}\bar{\Sigma}, \tag{13}$$

l'expression $\sqrt{g}$ est une densité.

Le produit d'une densité d'un ordre quelconque par une capacité donne un vrai tenseur; c'est ainsi que $\Sigma\tau$ est un invariant, et représente la masse comprise dans la capacité τ; $\sqrt{g}\tau$ est un autre invariant, et représente le volume; en général une expression

$$A_{ijk}^{lm}=\tau_{ij}\Sigma_k^{lm} \tag{14}$$

est un vrai tenseur.

Le produit d'une densité par un tenseur est une densité tensorielle, et de même pour une capacité.

Toutes ces définitions, ainsi que le mode de formation des capacités ou densités tensorielles à partir des tenseurs symétriques gauches, se généralisent aisément dans un espace riemannien à n dimensions.

3. *Déplacement parallèle, dérivée covariante.*

Les notions de géométrie tensorielle, rappelées au paragraphe précédent, suffisent, ainsi que l'a montré M. P. Langevin, pour l'étude de tous les chapitres de la physique qui ne font pas intervenir les déformations: électro-magnétisme, chaleur, diffusion, etc.

Pour définir les déformations d'un corps solide, il faut faire intervenir la notion de déplacement parallèle, et celle de dérivée covariante qui s'en déduit. L'élasticité est le premier chapitre où ces problèmes se présentent, et son étude prépare celle des cas plus complexes, tels que l'électro ou la magnéto-striction, la double réfraction par déformation, la piézo-électricité, etc.

Il est indispensable, ayant un tenseur f^k en un point $M(x^i)$, de pouvoir transporter ce tenseur sans changement en un point voisin $M'(x^i+dx^i)$. Si l'on ne préjuge rien des propriétés de l'espace, le transport du vecteur f^k donnera des résultats différents suivant le chemin suivi pour aller de M en M'. On précise donc que l'on suivra une géodésique MM'; les composantes du tenseur transporté seront alors:

$$f^k+df^k \text{ avec } df^k = -\begin{Bmatrix} k \\ rs \end{Bmatrix} f^r dx^s. \tag{15}$$

La parenthèse représente le symbole de seconde espèce de Christoffel*; ce symbole s'écrit:

$$\begin{Bmatrix} k \\ rs \end{Bmatrix} = \frac{1}{2} g^{ik} \left[\frac{\partial g_{ri}}{\partial x^s} + \frac{\partial g_{si}}{\partial x^r} - \frac{\partial g_{rs}}{\partial x^i} \right]. \tag{16}$$

Il est bien entendu que la formule (15) représente une somme double étendue aux valeurs $r=1, 2, 3$ et $s=1, 2, 3$; la formule (16) représente une somme en $i=1, 2, 3$.

Le second symbole de Christoffel est symétrique par rapport aux deux indices inférieurs r et s.

On obtient la dérivée covariante du tenseur f^k si l'on compare la valeur de f au point M à la valeur du tenseur en M', cette dernière étant ramenée en M par transport parallèle; ceci nous donne la formule de dérivation

$$\frac{Df^k}{Dx^s} = \frac{\partial f^k}{\partial x^s} + \begin{Bmatrix} k \\ rs \end{Bmatrix} f^s. \tag{17}$$

La dérivée ordinaire $\dfrac{\partial f^k}{\partial x^s}$ ne forme pas un tenseur; on démontre que la dérivée covariante est un tenseur véritable. Pour un tenseur covariant h_i la formule de dérivation s'écrit

$$\frac{Dh_i}{Dx^s} = \frac{\partial h_i}{\partial x^s} - \begin{Bmatrix} r \\ is \end{Bmatrix} h_r. \tag{18}$$

On définit aussi la dérivée covariante d'un tenseur à plusieurs indices

$$\frac{Dt_{jk}^{lm}}{Dx^s} = \frac{\partial t_{jk}^{lm}}{\partial x^s} - \begin{Bmatrix} r \\ ks \end{Bmatrix} t_{jr}^{lm} - \begin{Bmatrix} r \\ js \end{Bmatrix} t_{rk}^{lm} + \begin{Bmatrix} m \\ rs \end{Bmatrix} t_{jk}^{lr} + \begin{Bmatrix} l \\ rs \end{Bmatrix} t_{jk}^{rm}; \tag{19}$$

on ajoute à la dérivée ordinaire autant de termes correctifs qu'il y a d'indices; le signe de ces termes est $+$ si l'indice muet (r) figure comme indice contrevariant du tenseur t; le signe est $-$, si l'indice muet est un indice covariant du tenseur t.

*J'ai interverti les positions des indices, et écrit $\begin{Bmatrix} k \\ rs \end{Bmatrix}$ au lieu de $\begin{Bmatrix} rs \\ k \end{Bmatrix}$; ceci a l'avantage de ne pas modifier les conventions relatives aux rôles des indices haut et bas; Weyl écrit Γ_{rs}^k.

On vérifie enfin aisément que l'on a la relation suivante:

$$\text{(20)} \qquad \frac{D\, a_\alpha^k b^\alpha}{Dx^s} = b^\alpha \frac{Da_\alpha^k}{Dx^s} + a_\alpha^k \frac{Db^\alpha}{Dx^s}\,;$$

l'expression $a_\alpha^k b^\alpha$ sur laquelle porte la dérivation du premier membre est un produit contracté, équivalent à un tenseur f^k. La formule (20) est l'analogue de la formule de dérivation d'un produit.

Nous avons vu, au paragraphe précédent, comment s'introduisaient les densités ou capacités tensorielles; il nous faut établir les formules de dérivation pour ces pseudo-tenseurs. Nous pouvons, tout d'abord, partir du tenseur primitif symétrique gauche t^{ij} et appliquer les formules précédentes:

$$\text{(21)} \qquad \frac{Dt^{ij}}{Dx^s} = \frac{\partial t^{ij}}{\partial x^s} + \begin{Bmatrix} j \\ rs \end{Bmatrix} t^{ir} - \begin{Bmatrix} i \\ rs \end{Bmatrix} t^{jr}$$

en remplaçant t^{rj} par $-t^{jr}$.

On voit aussitôt que cette dérivée covariante est un tenseur symétrique gauche pour i et j; à t^{ij} nous faisions correspondre la capacité tensorielle τ_k; de sa dérivée nous déduirons la dérivée de τ_k, qui sera une nouvelle capacité tensorielle:

$$\text{(22)} \qquad \frac{D\tau_k}{Dx^s} = \frac{Dt^{ij}}{Dx^s} = \frac{\partial \tau_k}{\partial x^s} - \begin{Bmatrix} r \\ ks \end{Bmatrix} \tau_r + \begin{Bmatrix} r \\ sr \end{Bmatrix} \tau_k;$$

le développement que nous écrivons ici se déduit de la formule (21) par un groupement convenable des termes*. On retrouve exactement le même développement en se rappelant que $\sqrt{g}\tau_k$ est un tenseur vrai, dont nous pouvons former la dérivée covariante; pour retrouver une capacité tensorielle, nous écrirons:

$$\text{(23)} \qquad \frac{D\tau_k}{Dx^s} = \frac{1}{\sqrt{g}} \frac{D\sqrt{g}\tau_k}{Dx^s} = \frac{\partial \tau_k}{\partial x^s} - \begin{Bmatrix} r \\ ks \end{Bmatrix} \tau_r + \tau_k \frac{\partial \log\sqrt{g}}{\partial x^s}.$$

On connait la valeur de $\dfrac{\partial \log \sqrt{g}}{\partial x^s} = \begin{Bmatrix} r \\ sr \end{Bmatrix}$, et l'on voit que les deux développements sont égaux.

Pour une densité tensorielle, les raisonnements sont tout-à-fait semblables, et donnent:

$$\text{(24)} \qquad \frac{D\Sigma^k}{Dx^s} = \sqrt{g} \frac{D \frac{1}{\sqrt{g}} \Sigma^k}{Dx^s} = \frac{\partial \Sigma^k}{\partial x^s} + \begin{Bmatrix} k \\ rs \end{Bmatrix} \Sigma^r - \begin{Bmatrix} r \\ sr \end{Bmatrix} \Sigma^k.$$

*Prenons, par exemple t^{12} et formons la dérivée (21)

$$\frac{Dt^{12}}{Dx^s} = \frac{\partial t^{12}}{\partial x^s} + \begin{Bmatrix} 2 \\ 1s \end{Bmatrix} t^{11} + \begin{Bmatrix} 2 \\ 2s \end{Bmatrix} t^{12} + \begin{Bmatrix} 2 \\ 3s \end{Bmatrix} t^{13} - \begin{Bmatrix} 1 \\ 1s \end{Bmatrix} t^{21} - \begin{Bmatrix} 1 \\ 2s \end{Bmatrix} t^{22} - \begin{Bmatrix} 1 \\ 3s \end{Bmatrix} t^{23};$$

les termes t^{11} et t^{22} sont nuls, puisque $t^{ij} = -t^{ji}$

$$\frac{Dt^{12}}{Dx^s} = \frac{\partial t^{12}}{\partial x^s} + \left[\begin{Bmatrix} 1 \\ 1s \end{Bmatrix} + \begin{Bmatrix} 2 \\ 2s \end{Bmatrix} + \begin{Bmatrix} 3 \\ 3s \end{Bmatrix} \right] t^{12} - \begin{Bmatrix} 1 \\ 3s \end{Bmatrix} t^{23} - \begin{Bmatrix} 2 \\ 3s \end{Bmatrix} t^{31} - \begin{Bmatrix} 3 \\ 3s \end{Bmatrix} t^{12},$$

ce qui n'est autre que le développement (22).

On en déduit la règle suivante: pour former la dérivée covariante d'un pseudo-tenseur par rapport à x^s, on prend d'abord les termes qui correspondraient à un vrai tenseur de même variance et l'on ajoute $+\left\{ {r \atop sr} \right\}\tau$ pour une capacité tensorielle τ, ou $-\left\{ {r \atop sr} \right\}\Sigma$ pour une densité tensorielle Σ.

4. *Étude des déformations.*

Nous allons pouvoir, avec les notions de géométrie introduites dans les précédents paragraphes, aborder maintenant l'étude des déformations d'un corps solide. Je rappellerai d'abord la méthode suivie par E. et F. Cosserat* et qu'il me faudra généraliser: Un corps continu étant rapporté à un système de coordonnées rectangulaires x, y, z, et u, v, w représentant les projections du déplacement d'un point (x, y, z), le corps possède, après sa déformation par un système de forces extérieures, un élément linéaire dont le carré

$$ds^2 = (dx+du)^2 + (dy+dv)^2 + (dz+dw)^2$$

est de la forme:

$$ds^2 = (1+2\epsilon_1)dx^2 + (1+2\epsilon_2)dy^2 + (1+2\epsilon_3)dz^2 + 2\gamma_1 dydz + 2\gamma_2 dzdx + 2\gamma_3 dxdy.$$

Les six éléments ϵ et γ sont ceux par lesquels on a l'habitude de caractériser la déformation dans la théorie de l'élasticité.

Cette définition est tout-à-fait générale et s'applique à une déformation finie quelconque; il est évident que l'on doit réserver le mot de déformation pour les genres de déplacements qui modifient la forme du corps, c'est-à-dire changent la distance de deux points voisins.

Les notations de Cosserat manquent de symétrie, et nous verrons que, pour former un tenseur, il faut prendre comme composantes $2\epsilon_1$, $2\epsilon_2$, $2\epsilon_3$, γ_1, γ_2, γ_3; on rétablit aussitôt l'homogénéité des formules; Cosserat trouve, en axes rectangulaires, les expressions suivantes: (*loc. cit.*, p. 10, formules 3):

$$(25)\qquad \left\{ \begin{aligned} 2\epsilon_1 &= 2\frac{\partial u}{\partial x} + \left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial v}{\partial x}\right)^2 + \left(\frac{\partial w}{\partial x}\right)^2, \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ \gamma_1 &= \frac{\partial w}{\partial y} + \frac{\partial v}{\partial z} + \frac{\partial u}{\partial y}\frac{\partial u}{\partial z} + \frac{\partial v}{\partial y}\frac{\partial v}{\partial z} + \frac{\partial w}{\partial y}\frac{\partial w}{\partial z}, \\ &\cdots\cdots\cdots\cdots\cdots\cdots \end{aligned} \right.$$

Il est indispensable de garder ces expressions au complet, sans se limiter, comme on le fait souvent, aux premiers termes.

Ce raisonnement se transcrit sans trop de difficultés en coordonnées quelconques, ainsi que je vais le montrer. Considérons un corps solide, rapporté à un système k de coordonnées; soient x^1, x^2, x^3 les coordonnées d'un point M de ce corps, dans l'état initial non déformé; considérons deux points voisins M

*E. et F. Cosserat. Annales de la Faculté des Sciences de Toulouse, t. 10 (1896), I, p. 2.

et $M'(x^1+dx^1,\ x^2+dx^2,\ x^3+dx^3)$. Le carré de leur distance est donné par la forme quadratique

$$ds^2 = g_{ij} dx^i dx^j. \tag{26}$$

Supposons maintenant le corps soumis à une déformation; les points M et M' du solide viennent occuper des positions différentes, M_1 et M_1'; prenons les coordonnées X^i du point M_1 et X^i+dX^i de M_1', par rapport au système d'axes k, dont je suppose essentiellement qu'il est resté invariable pendant la déformation. Le carré de la distance M_1M_1' est maintenant:

$$dS^2 = g_{RS} dX^R dX^S, \tag{27}$$

la notation g_{RS} indiquant qu'il s'agit des valeurs des g prises au point $M_1(X^i)$ et non pas en $M(x^i)$.

Pour définir la déformation en tout point, je dois me donner X^1, X^2, X^3 comme fonctions des x^1, x^2, x^3; je puis alors développer la formule (27) et écrire:

$$dS^2 = g_{RS}\, a_i^R a_j^S dx^i dx^j = h_{ij} dx^i dx^j, \tag{28}$$

en posant

$$a_i^R = \frac{\partial X^R}{\partial x^i}, \qquad dX^R = a_i^R dx^i. \tag{28$^{\text{bis}}$}$$

Pour éviter toute confusion, il faut se rappeler que les formules (28^{bis}) ne représentent pas un changement de coordonnées, mais un déplacement du corps solide par rapport à un système de coordonnées rigides.

Par la manière même dont je les ai formés, je vois que les coefficients h_{ij} forment un tenseur deux fois covariant; j'aurai le tenseur de la déformation en prenant la différence

$$e_{ij} = h_{ij} - g_{ij}. \tag{29}$$

Pour le cas des coordonnées rectilignes rectangulaires, la correspondance entre mes notations et celles de Cosserat, est la suivante:

$$2\epsilon_1 = e_{11}, \ldots, \quad \gamma_1 = e_{23}, \ldots .$$

Je dois maintenant chercher à expliciter les expressions générales (29) qui définissent la déformation; pour cela j'introduirai le déplacement u^i d'un point du corps solide, en posant:

$$X^i = x^i + u^i. \tag{30}$$

Je noterai aussitôt que u^i n'est un tenseur covariant que si la déformation est infiniment petite; les a_i^R ne forment pas un tenseur. J'aurai, pour ces derniers coefficients, les valeurs suivantes:

$$a_i^r = \begin{cases} 1 + \dfrac{\partial u^r}{\partial x^i} & (r = i), \\[2ex] \dfrac{\partial u^r}{\partial x^i} & (r \neq i). \end{cases}$$

Dans l'expression des e_{ij}, il faut donc distinguer les termes particuliers pour lesquels les deux indices d'un coefficient a deviennent égaux; ceci permet d'écrire:

$$e_{ij} = a_i^r a_j^s g_{RS} - g_{ij} = g_{IJ} - g_{ij} + \frac{\partial u^s}{\partial x^j} g_{IS} + \frac{\partial u^r}{\partial x^i} g_{JR} + \frac{\partial u^r}{\partial x^i}\frac{\partial u^s}{\partial x^j} g_{RS}; \tag{31}$$

les termes exceptionnels écrits en premier lieu proviennent des cas $r=i$ et $s=j$. Nous allons développer cette expression, en gardant les termes du premier et du second ordre; nous aurons, par un développement en série de Taylor:

$$g_{RS} = g_{rs} + \frac{\partial g_{rs}}{\partial x^p} u^p + \tfrac{1}{2} \frac{\partial^2 g_{rs}}{\partial x^p \partial x^q} u^p u^q,$$

d'où nous tirons:

$$\begin{aligned} e_{ij} = {} & \frac{\partial g_{ij}}{\partial x^p} u^p + \frac{\partial u^s}{\partial x^j} g_{is} + \frac{\partial u^r}{\partial x^i} g_{jr} \\ & + \frac{\partial u^r}{\partial x^i}\frac{\partial u^s}{\partial x^j} g_{rs} + \frac{\partial u^r}{\partial x^i} u^p \frac{\partial g_{rj}}{\partial x^p} + \frac{\partial u^s}{\partial x^j} u^p \frac{\partial g_{si}}{\partial x^p} + \tfrac{1}{2} \frac{\partial^2 g_{ij}}{\partial x^p \partial x^q} u^p u^q + \dots \end{aligned}$$

Les termes du premier degré, groupés sur la première ligne se transcrivent sans peine sous la forme:

$$\frac{Du_i}{Dx^j} + \frac{Du_j}{Dx^i}.$$

Nous pouvons ainsi, après quelques calculs simples, arriver à la forme suivante:

$$\begin{aligned} e_{ij} = {} & \frac{Du_i}{Dx^j} + \frac{Du_j}{Dx^i} + g_{rs} \frac{Du^r}{Dx^i}\frac{Du^s}{Dx^j} \\ & + u^p \left\{ {r \atop ps} \right\} \left[g_{ir} \frac{\partial u^s}{\partial x^j} + g_{jr} \frac{\partial u^s}{\partial x^i} \right] + u^p u^q \left[\tfrac{1}{2} \frac{\partial^2 g_{ij}}{\partial x^p \partial x^q} - g_{rs} \left\{ {r \atop ip} \right\} \left\{ {s \atop jq} \right\} \right] + \dots \end{aligned} \tag{32}$$

La seconde ligne, et tous les termes d'ordre supérieur (que nous n'avons pas écrits) disparaissent si l'espace est euclidien et si l'on prend des axes rectilignes, car toutes les dérivées des g sont alors nulles. Ces formules (32), réduites à la première ligne, sont identiques à celles de Cosserat*, si l'on se place dans le cas particulier des axes orthogonaux; la symétrie $e_{ij}=e_{ji}$ est évidente sur la formule.

5. *Formules complémentaires sur les déformations.*

Considérons un corps qui a subi une première déformation; deux points, voisins, de coordonnées initiales x^i et $x^i + dx^i$ sont venus occuper de nouvelles positions X^i et $X^i + dX^i$. Le carré de leur distance a pour valeur, après cette déformation:

$$dS^2 = g_{RK} dx^R \, dx^K,$$

**loc. cit.*, éq. 3, p. 10, transcrites ici sous le N° 25, paragraphe 4.

suivant notre formule (27); supposons que nous effectuions une très petite variation de la déformation: les coordonnées X^i sont augmentées de $\delta X^i = \delta u^i$; celles du point $X^i + dX^i$ deviennent $X^i + dX^i + d\delta u^i$; le dS^2 subit alors une variation:

$$\delta dS^2 = g_{RK} dX^r d\delta u^k + g_{RK} dX^k d\delta u^r + \delta g_{RK} dX^r dX^k$$

ou, en changeant la dénomination des indices muets des deux derniers termes ($r \longrightarrow l$, $K \longrightarrow s$ au 2^e terme; $K \longrightarrow s$ au 3^e terme):

$$\delta dS^2 = \left\{ g_{RK} \frac{\partial \delta u^k}{dX^s} + g_{SL} \frac{\partial \delta u^l}{\partial X^r} + \frac{\partial g_{RS}}{\partial X^p} \delta u^p \right\} dX^r dX^s;$$

ceci s'écrit encore, moyennant quelques transformations simples:

(33) $$\delta dS^2 = \left\{ \frac{D\delta u_r}{DX^s} + \frac{D\delta u_s}{DX^r} \right\} dX^r dX^s.$$

Mais nous pouvons aussi exprimer le dS^2 au moyen des dx^i en utilisant la formule (28); lors d'une variation de la déformation, les dx^i restent inchangés, et nous obtenons:

(34) $$\delta dS^2 = \delta h_{ij} dx^i dx^j = \delta e_{ij} dx^i dx^j.$$

Comparons les expressions (33) et (34), et nous en tirons aussitôt:

(35) $$\delta e_{ij} = \left\{ \frac{D\delta u_r}{DX^s} + \frac{D\delta u_s}{DX^r} \right\} a_i^r a_j^s,$$

les coefficients a_i^r sont, comme précédemment, les dérivées $\dfrac{\partial X^r}{\partial x^i}$. Cette formule sera très utile par la suite; nous allons voir comment elle peut s'interpréter.

Nous avons, jusqu'à présent, toujours envisagé les coordonnées X^i d'un point du corps déformé, rapportées aux axes k invariables. Nous pouvons procéder autrement et définir un point du corps déformé par les trois nombres x^1, x^2, x^3, qui donnaient sa position à l'état initial; nous pouvons choisir, dans le corps déformé, de nouvelles surfaces coordonnées, formant un système d'axes $\overline{K}$, et par rapport auxquelles les coordonnées d'un point sont numériquement égales à x^1, x^2, x^3; ce système d'axes $\overline{K}$ est celui qui se déduit des axes initiaux **k** en les supposant entraînés par le corps et déformés avec lui. Pour ces axes $\overline{K}$, nous aurons donc;

(36) $$\overline{X}^i = x^i, \quad d\overline{S}^2 = h_{ij}\, dx^i\, dx^j = h_{ij}\, d\overline{X}^i\, d\overline{X}^j.$$

Le changement de coordonnées se fait alors suivant les formules:

(37) $$dX^r = a_i^r\, dx^i = a_i^r\, d\overline{X}^i,$$

et un tenseur t_{ij}, se transforme ainsi:

(38) $$\overline{t}_{ij} = a_i^r a_j^s t^{rs}.$$

Comparons les formules (37) et (35) et nous arrivons à la relation:

(39) $$\delta e_{ij} = \frac{D\delta u_i}{D\overline{X}^j} + \frac{D\delta u_j}{D\overline{X}^i}.$$

Ce résultat peut s'énoncer ainsi: les petits déplacements supplémentaires δu étant mesurés, chaque fois, par rapport au système d'axes $\overline{K}$ entraînés et déformés par toute la déformation précédente, les variations des composantes du tenseur e_{ij} de la déformation sont données par la formule très simple (39).

6. *Corps isotrope initialement, invariants de la déformation.*

Lorsqu'on considère un corps initialement homogène et isotrope, il est important de chercher quelles sont les combinaisons invariantes que l'on peut former au moyen des composantes du tenseur de déformation; ces invariants s'écrivent au moyen des composantes mixtes $e_i^k = g^{kj}e_{ij}$.
Ce sont:

$$\text{(40)}\qquad \begin{aligned} &\text{l'invariant du premier degré,} && I_1 = e_i^i = e_1^1 + e_2^2 + e_3^3, \\ &\text{l'invariant du second degré,} && I_2 = e_i^k\ e_k^i, \\ &\text{et celui du troisième degré,} && I_3 = e_i^k e_k^l e_l^i. \end{aligned}$$

Comparons ces résultats aux formules élémentaires: nous nous reporterons aux notations de Cosserat, relatives aux axes rectilignes orthogonaux; dans un tel système d'axes, on sait que les composantes mixtes sont numériquement égales aux composantes covariantes, et l'on a:

$$(40^{\text{bis}})\qquad \begin{cases} I_1 = 2(\epsilon_1 + \epsilon_2 + \epsilon_3), \\ I_2 = 2(\gamma_1^2 + \gamma_2^2 + \gamma_3^2 - 4\epsilon_1\epsilon_2 - 4\epsilon_2\epsilon_3 - 4\epsilon_3\epsilon_1) + I_1^2; \end{cases}$$

les deux parenthèses sont les combinaisons invariantes choisies par Cosserat*.

On peut rapporter la déformation, pour le voisinage immédiat d'un point M, à ses axes principaux, c'est-à-dire à un système d'axes orthogonaux pour lequel les glissements γ s'annulent; soient alors α_1, α_2, α_3 les trois dilatations principales, on a:

$$\begin{cases} I_1 = 2(\alpha_1 + \alpha_2 + \alpha_3), \\ I_2 = 4(\alpha_1^2 + \alpha_2^2 + \alpha_3^2). \end{cases}$$

Il ne faut pas oublier que I_1 n'est pas égal à la dilatation cubique; ces deux expressions n'ont que leurs termes du premier ordre qui soient semblables, et les termes du 2^e au 3^e ordre, par rapport aux dérivées de la déformation sont différents†. Un élément de volume, défini par trois déplacements dx^1, dx^2, dx^3 avait pour volume, avant la déformation, $\sqrt{g}\,dx$ et après la déformation $\sqrt{h}\,dx$, en désignant par dx le produit $dx^1dx^2dx^3$; la dilatation cubique est donc:

$$\text{(41)}\qquad \theta = \frac{\sqrt{h}}{\sqrt{g}} - 1 = |\Delta| - 1.$$

Dans ces formules, g représente le déterminant des g_{RS} pris au point X^r, h celui des h^{ik}, et Δ le déterminant des a_i^R, dérivées partielles des nouvelles coordonnées X^r par rapport aux coordonnées initiales x^i.

loc. cit., p. 26.
†E. F. Cosserat, *loc. cit.*, p. 23.

7. *Les efforts dans un corps déformé.*

Après l'étude des déformations, il nous faut maintenant considérer la nature des efforts dans un corps solide. Nous chercherons quelles sont les forces réciproques qu'exercent l'une sur l'autre deux parties du solide, à travers une surface infiniment petite ds qui les sépare, et nous nous limiterons d'abord au solide *au repos*.

Il faut définir un élément de surface ds, ce qui se fait de la manière suivante: en un point M (coordonnées X^i) nous considérerons deux déplacements infiniments petits b^1 (composantes b^{1i}) et b^2 (composantes b^{2i}). Le produit extérieur de ces deux vecteurs b^1 et b^2 nous donne l'expression de l'élément de surface constitué par la parallélogramme construit sur b^1 et b^2:

$$ds = [b^1 . b^2].$$

Les composantes de l'élément de surface sont:

$$ds^{ij} = [b^{1i} . b^{2j}] = b^{1i}b^{2j} - b^{1j}b^{2i}; \tag{42}$$

nous obtenons ainsi un tenseur deux fois covariant et symétrique gauche.

Les forces exercées par les portions du corps solide à travers la surface ds seront proportionnelles à cette surface; nous écrirons donc*:

$$f_l = T_{l.ij}\, ds^{ij}. \tag{43}$$

Les coefficients $T_{l.ij}$ ainsi introduits forment un tenseur trois fois covariant. Mais ds^{ij} étant symétrique gauche, nous ne diminuons en rien la généralité† si nous supposons que $T_{l.ij}$ est symétrique gauche par rapport aux deux indices i et j.

Nous avons vu, au paragraphe 2, que nous pouvions représenter l'élément de surface ds^{ij} au moyen de trois composantes $d\sigma_k$ formant une *capacité tensorielle*. Ceci nous oblige à choisir un sens de rotation favorisé i, j, k sur les axes de coordonnées; il revient au même de dire que, les vecteurs b^1 et b^2 étant donnés, nous pouvons définir un sens positif sur la normale à l'élément de surface $[b^1 . b^2]$.

Le tenseur $T_{l.ij}$ se représente alors par une densité tensorielle:

$$\Theta_l^k = 2T_{l.ij} = -2T_{l.ji}, \tag{44}$$

le facteur 2 est nécessaire si nous voulons garder pour les efforts f_l l'expression simple:

$$f_l = \Theta_l^k d\sigma_k \tag{45}$$

où figurent moitié moins de termes que dans la formule (43) (3 termes au lieu de 6). Cette force, avec les définitions usuelles, représente l'effort exercé, sur la partie solide située du côté de la normale négative, par la portion du solide

*Une force est définie ordinairement par ses composantes covariantes de telle sorte que le travail $f_l \delta x^l$ soit un invariant.

†Supposons que nous ayons pris un tenseur $t_{l.ij}$ quelconque; la somme (43) porte sur les indices i et j; groupons les termes ij et ji; nous pouvons écrire $f_l = \frac{1}{2}(t_{l.ij} - t_{l.ji})ds^{ij}$ ce qui nous donne le tenseur symétrique gauche $T_{l.ij} = \frac{1}{2}(t_{l.ij} - t_{l.ji})$ la partie symétrique de ce tenseur ne joue donc aucun rôle et peut être supprimée.

située vers la normale positive. La somme (45), représente le produit contracté d'une capacité tensorielle $d\sigma_k$ par une densité tensorielle Θ_l^k, ce qui donne bien un vrai tenseur covariant f_l.

Les formules (44) ou (45) nous donnent les forces exercées à travers un élément de surface quelconque; c'est le résultat que l'on obtient ordinairement au moyen d'un raisonnement basé sur l'équilibre d'un tétraèdre.

8. *Symétrie du tenseur des efforts.*

On introduit toujours, en élasticité*, une hypothèse particulière qui donne une condition de symétrie pour les tensions; on suppose que les efforts élastiques n'ont pas pour résultat de créer un moment de rotation sur un petit volume élémentaire. Nous allons voir que cette condition nous permet d'écrire les relations:

$$\Theta^{lk} = \Theta^{kl}, \tag{46}$$

qui indiquent la symétrie du pseudo-tenseur Θ^{lk}, lorsqu'on prend ses composantes deux fois contrevariantes. Nous pouvons chercher quelle est la condition correspondante pour le tenseur T à trois indices, dont nous prendrons les composantes mixtes T_{ij}^l; nous savons que ce tenseur étant symétrique gauche, les composantes, pour lesquelles les deux indices i et j sont égaux, sont nulles. Prenons, par exemple, le cas $\Theta^{12} = \Theta^{21}$; cette condition se transcrit ainsi:

$$T_{31}^1 = T_{23}^2 \text{ ou } T_{31}^1 + T_{32}^2 = 0,$$

ou encore, puisque T_{33}^3 est nul:

$$T_{31}^1 + T_{32}^2 + T_{33}^3 = 0,$$

ce qui s'écrit:

$$T_{3l}^l = 0.$$

Le tenseur mixte satisfait donc aux 3 relations suivantes:

$$T_{il}^l = 0. \tag{47}$$

Il nous reste maintenant à donner, sous une forme très générale, la démonstration de la relation (46).

Considérons trois déplacements infiniment petits b^{1i}, b^{2i} et b^{3i} autour d'un point M. Ces trois déplacements définissent un parallélogramme élémentaire:

$$[b^{1i} . b^{2i} . b^{3i}] = \tau,$$

et différentes surfaces:

$$ds^{\alpha\beta . ij} = b^{\alpha i} b^{\beta j} - b^{\alpha j} b^{\beta i} .$$

Prenons un sens de rotation 1, 2, 3 pour les indices α, β, γ et i, j, k. Nous pourrons écrire (Fig. 1):

$$ds^{\alpha\beta . ij} = d\sigma_{\gamma . k}.$$

*Une telle hypothèse serait inadmissible dans d'autres problèmes physiques (questions de magnéto ou électro-élasticité).

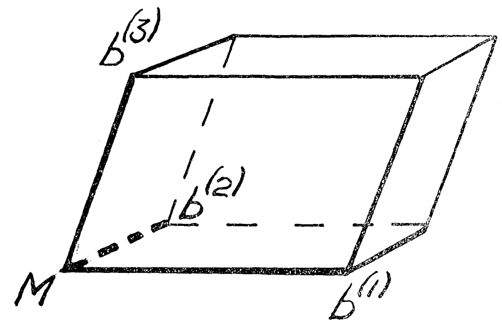


Fig. 1

L'élément de volume τ est égal au déterminant des $b^{\alpha i}$ et l'on vérifie facilement les relations suivantes:

(48) $$\sum_{\gamma} b^{(\gamma)i} d\sigma_{(\gamma)k} = \begin{cases} 0 & si\ i \neq k \\ \tau & si\ i = k. \end{cases}$$

Je considère la force qui s'exerce à travers une face $\alpha\beta$; je prends ses composantes contrevariantes $f^{\,r}_{(\gamma)}$:

$$f^{\,r}_{(\gamma)} = \Theta^{rk}\, d\sigma_{(\gamma).k}\,.$$

Le point d'application de cette force est à l'extrémité du vecteur $b^{(\gamma)s}$; je puis prendre le moment de cette force par rapport à l'origine (M); ce moment s'écrit:

(49) $$M^{rs}_{(\gamma)} = [f^{\,r}_{(\gamma)} . b^{(\gamma)s}] = f^{\,r}_{(\gamma)} b^{(\gamma)s} - f^{\,s}_{(\gamma)} b^{(\gamma)r} = \Theta^{rk} b^{(\gamma)s} d\sigma_{(\gamma)k} - \Theta^{sk} b^{(\gamma)r} d\sigma_{(\gamma)k}.$$

J'ai mis les indices γ entre parenthèses, car jusqu'à présent je supposais α, β et γ fixés et n'envisageais que les grandeurs relatives à une face γ donnée. Si je veux chercher le moment résultant pour les différentes faces, il me suffit, dans la formule (48), de convenir de faire la somme par rapport à l'indice γ.

En tenant compte des relations (48) ceci se réduit à:

(50) $$M^{rs} = [\Theta^{rs} - \Theta^{sr}]\tau.$$

En annulant le moment résultant, on trouve donc bien la condition de symétrie du pseudo-tenseur Θ^{rs}.

Ces divers raisonnements nous ont donné l'occasion de nous servir des simplifications d'écriture qu'entraîne l'emploi des pseudo-tenseurs; il serait très mal commode de garder constamment le tenseur $T_{l.ij}$ à 3 indices, et ses 27 composantes; les conditions de symétrie gauche réduisent à 9 le nombre des composantes indépendantes; les 3 conditions $\Theta^{sr} = \Theta^{rs}$ ramènent ce nombre à 6; il eut été très peu pratique de garder des notations qui ne laissent pas apparaître ces simplifications.

9. *Forces résultantes sur un élément de volume.*

Considérons, de nouveau, le parallélépipède bâti sur trois déplacements infinitésimaux b^{1i}, b^{2i}, b^{3i}; nous supposerons que l'ordre de ces trois vecteurs correspond au sens de rotation des axes, de telle sorte que le vecteur b^3, par exemple, se trouve par rapport à l'élément de surface $[b^1, b^2]$ du coté de la normale positive; dans ces conditions, cherchons les forces qui s'exercent à travers deux faces opposées du parallélépipède; nous prendrons, par exemple la face $[b^\alpha, b^\beta]$ qui passe au sommet M, et la face opposée, qui se déduit de celle-ci par une translation parallèle le long du vecteur b^γ.

Considérons d'abord la face $[b^\alpha, b^\beta]$; la partie du solide située à l'extérieur de cette face se trouve du coté de la normale négative; les forces exercées sur le solide intérieur au parallépipède s'écrivent donc, avec nos conventions antérieures:

$$f^r_{(\gamma)} = -\Theta^{rk} d\sigma_{(\gamma)k}. \tag{51}$$

Nous avons écrit les composantes contrevariantes de la force, mais le calcul se ferait exactement de la même manière au moyen des composantes covariantes. Pour la face opposée, le sens sera inversé le solide intérieur au parallélipipède se trouvant du côté de la normale négative; quant à la grandeur de la force, elle sera:

$$f^r_{(\gamma)} + \frac{Df^r_{(\gamma)}}{DX^i} b^{(\gamma)i}:$$

la notation D correspond à la dérivée covariante, qu'il est indispensable d'introduire pour obtenir la valeur de la force en un point distant de $b^{(\gamma)}$ de l'origine M; l'indice $(\gamma)_{\text{est}}$ mis entre parenthèses afin de rappeler qu'il n'y a pas lieu, tout d'abord, de sommer par rapport à γ. La force résultante des efforts sur les deux faces γ opposées s'écrit donc:

$$F^r_{(\gamma)} = \frac{Df^r_{(\gamma)}}{DX^i} b^{(\gamma)i}, \tag{52}$$

et l'on aura la force totale, due aux efforts sur les trois groupes de faces, en faisant maintenant la somme par rapport à l'indice γ.

Nous savons d'autre part, d'après les règles de dérivation covariante, que nous pouvons écrire [cf. formule (20), §3)]:

$$\frac{Df^r_\gamma}{DX^i} = \frac{D\Theta^{rk}.d\sigma_{\gamma k}}{DX^i} = \frac{D\Theta^{rk}}{DX^i} d\sigma_{\gamma k} + \Theta^{rk} \frac{Dd\sigma_{\gamma k}}{DX^i}, \tag{53}$$

le dernier terme est nul; les deux faces opposées γ se déduisent en effet l'une de l'autre par transport parallèle; la dérivée covariante $\frac{Dd\sigma_{\gamma k}}{DX^i}$ est donc nulle; il nous reste alors:

$$F^r = \frac{D\Theta^{rk}}{DX^i} b^{\gamma i} d\sigma_{\gamma k}.$$

Si nous nous reportons aux formules (48) nous voyons que, dans cette somme, les termes non nuls sont donnés par le cas $i = k$ et s'écrivent:

$$F^r = \frac{D\Theta^{rk}}{DX^k} . \tau. \tag{54}$$

Il existe donc, dans le corps solide déformé, une densité de forces, dont les composantes contrevariantes sont:

(55) $$\Phi^r = \frac{D\Theta^{rk}}{DX^k} = \frac{\partial\Theta^{rk}}{\partial X^k} + \left\{ \begin{matrix} r \\ kl \end{matrix} \right\} \Theta^{kl} \, ;$$

les composantes covariantes s'écrivent:

(55bis) $$\Phi_r = \frac{D\Theta_r^k}{DX^k} = \frac{\partial\Theta_r^k}{\partial X^k} - \left\{ \begin{matrix} l \\ rk \end{matrix} \right\} \Theta_l^k .$$

Les formules (55) représentent l'opération bien connue de la *divergence* appliquée à une densité tensorielle.

10. *Le milieu solide en mouvement.*

Nous avons supposé, dans les paragraphes précédents, que nous étudions les efforts dans un solide déformé, mais au repos. Nous avons utilisé un système de coordonnées rigides, en définissant chaque point M du corps solide par les nombres X^i qui fixent la position actuelle de ce point par rapport à des axes indéformables.

Il nous est facile de passer maintenant au cas d'un solide en mouvement, rapporté aux axes fixes que nous venons de définir; les actions qui se transmettent à travers un élément de surface ds immobile sont alors de deux sortes: 1° les efforts élastiques statiques représentées par la densité tensorielle Θ_l^k introduite plus haut; 2° le flux de quantité de mouvement, qui correspond à l'écoulement matériel au travers de la surface ds. La densité de quantité de mouvement s'écrit ρv_l, en appelant v_l les composantes covariantes de la vitesse*; le flux de quantité de mouvement n'est autre que la densité tensorielle

$$\rho v_l v^k .$$

Dans toutes les formules précédentes, la densité tensorielle Θ_l^k sera donc à remplacer par $\Theta_l^k + \rho v_l v^k$; la symétrie des composantes contrevariantes est conservée, puisque $\rho v^l v^i = \rho v^i v^l$ et les forces résultantes sur l'élément de volume représentent une densité tensorielle:

(56) $$\Phi_r = \frac{\partial\Theta_r^k}{\partial X^k} + \frac{\partial \rho v_r v^k}{\partial X^k} - \left\{ \begin{matrix} l \\ rk \end{matrix} \right\} (\Theta_l^k + \rho v_l v^k).$$

Ces formules sont utiles pour l'étude des mouvements vibratoires d'un corps solide comme pour la propagation des ondes élastiques. Les équations du mouvement s'écrivent:

(56bis) $$\Phi_r = \frac{\partial \rho v_r}{\partial t} ,$$

*Il faut prendre les composantes covariantes pour garder sa forme simple à l'équation:

$$f_l = \frac{m\partial v_l}{\partial t} .$$

la dérivée étant prise au point considéré. Ces formules correspondent aux transformations classiques entre coordonnées de Lagrange et d'Euler.

II. *Relation entre les forces et les déformations.*

Lorsqu'on produit une petite variation de la déformation, le travail correspondant doit pouvoir s'exprimer en fonctions des variations du tenseur de déformation e_{ij}. Le travail prendra la forme de l'intégrale de volume d'une densité:

$$\delta \mathfrak{T}_e = \int \mathfrak{A}^{ij} \delta e_{ij} dx, \tag{57}$$

dx représentant l'élément de volume initial, et l'intégrale étant étendue à tout le solide considéré *dans son état primitif avant déformation.*

Il nous faut établir les relations qui existent entre les coefficients $\mathfrak{A}^{ij}$ (densité tensorielle deux fois contrevariante) et le tenseur des efforts, que nous avons introduit aux paragraphes précédents.

Admettons que les forces extérieures se traduisent par: 1° une densité de forces Ψ_r, que l'on pourra écrire, suivant une habitude courante, sous la forme ρF_r, en mettant en évidence la densité matérielle ρ. Ces forces s'exercent dans tout le volume du solide; si le corps est en mouvement, les forces d'accélération $\dfrac{\partial \rho v_r}{\partial t}$ rentreront dans ce terme général; 2° une densité superficielle de forces Ξ_r^k agissant sur la surface limite du solide; la force extérieure sur un élément de surface $d\sigma_k$ s'écrira $\Xi_r^k \, d\sigma_k$. La résultante des forces agissant sur un élément de volume $d\tau$ s'écrira alors, (formule 55^{bis}):

$$(\Psi_r + \Phi_r) d\tau,$$

tandis qu'avec nos conventions de signe, la résultante des forces sur un élément $d\sigma$ de la surface sera (51):

$$(\Xi_r^k - \Theta_r^k) d\sigma_k.$$

Le volume $d\tau$ et la surface $d\sigma$ sont pris dans les coordonnées finales X, qui représentent le corps déformé.

Supposons maintenant que l'équilibre soit réalisé, les deux termes ci-dessus étant nuls, je puis écrire qu'un déplacement arbitraire infiniment petit δu^r produit un travail nul:

$$\int_V (\Psi_r + \Phi_r) \, \delta u^r d\tau + \int_S (\Xi_r^k - \Theta_r^k) \delta u^r d\sigma_k = 0, \tag{58}$$

ce qui me donne:

$$\int_V \Psi_r \delta u^r d\tau + \int_S \Xi_r^k \, \delta u^r \, d\sigma_k + \int_V \frac{D\Theta_r^k}{DX^k} \delta u^r d\tau - \int_S \Theta_r^k \delta u^r d\sigma_k = 0.$$

Les deux derniers termes se laissent aisément transformer au moyen d'une intégration par parties; j'ai en effet:

$$\int_S \Theta_r^k \delta u^r d\sigma_k = \int_V \frac{D\Theta_r^k \delta u^r}{DX^k} \, d\tau,$$

de sorte que je puis écrire:

$$(59) \qquad \delta\mathfrak{T}_e = \int_V \Psi_r \delta u^r d\tau + \int_S \Xi_r^k \delta u^r d\sigma_k = \int_V \Theta_r^k \frac{D\delta u^r}{DX^k} d\tau,$$

$\delta\mathfrak{T}_e$ représentant le travail des forces extérieures Ψ et Ξ pour les déplacements arbitraires δu^r que nous avons imposés*. La dernière intégrale peut aussi bien s'écrire au moyen des composantes mixtes Θ^{lk}, car j'ai:

$$\Theta_r^k \frac{D\delta u^r}{DX^k} = \Theta^{lk} g_{lr} \frac{D\delta u^r}{DX^k} = \Theta^{lk} \frac{D\delta u_l}{DX^k},$$

puisque les dérivées covariantes sont de vrais tenseurs.

En vertu de la symétrie du pseudo-tenseur Θ^{lk} je puis alors grouper les termes lk et kl et prendre:

$$(60) \qquad \delta\mathfrak{T}_e = \int_V \Theta^{lk} \frac{D\delta u_l}{DX^k} d\tau = \tfrac{1}{2} \int_V \Theta^{lk} \left(\frac{D\delta u_l}{DX^k} + \frac{D\delta u_k}{DX^l}\right) d\tau.$$

L'expression entre parenthèses s'exprime au moyen des variations du tenseur de déformation. Nous avons, en effet, établi la relation suivante (§5, formule 35):

$$\delta e_{ij} = \left(\frac{D\delta u_l}{DX^k} + \frac{D\delta u_k}{DX^l}\right) a_i^k a_j^l,$$

qui se résout aussitôt ainsi:

$$\frac{D\delta u_l}{DX^k} + \frac{D\delta u_k}{DX^l} = \beta_k^i \beta_l^j \delta e_{ij},$$

$\Delta\beta_k^i$ étant le mineur de a_i^k dans le déterminant $\Delta = |a_i^k|$. D'autre part, pour comparer nos formules (60) et (57), il nous faut encore passer d'une intégrale portant sur les coordonnées X (après déformation) à une intégrale relative aux coordonnées initiales x; ceci se fait aisément grâce à la relation:

$$d\tau_x = \Delta . dx, \qquad \Delta = \frac{D(X^1, X^2, X^3)}{D(x^1, x^2, x^3)}.$$

Nous aboutissons donc au résultat suivant:

$$(61) \qquad \delta\mathfrak{T}_e = \tfrac{1}{2} \int \Delta . \beta_k^i \beta_l^j \Theta^{kl} \delta e_{ij} dx.$$

Nous trouvons alors, par comparaison des formules (61) et (57), les relations:

$$\mathfrak{A}^{ij} = \tfrac{1}{2} \Delta . \beta_k^i \beta_l^j \Theta^{kl},$$

ou inversement:

$$(62) \qquad \Theta^{kl} = \frac{2}{\Delta} a_i^k a_j^l \mathfrak{A}^{ij}.$$

*Pour un déplacement réel δu^r et un mouvement à vitesse fini, on obtiendrait ici $d\mathfrak{T}_l - dE_{cin}$, dE_{cin} représentant l'énergie cinétique totale du corps solide; ceci est évident, puisque nous avons fait rentrer les forces d'inertie dans le terme Ψ_r.

Ce résultat correspond exactement aux formules établies par Cosserat pour le cas des axes orthogonaux (*loc. cit.*, p. 44 éq. 31); on ne peut manquer d'être frappé de la simplification d'écriture apportée par les notations tensorielles; les formules développées de Cosserat tiennent une page entière! Le facteur 2 manque chez Cosserat; cela tient à ce qu'il a pris des définitions différentes pour les déformations; notre e_{11} correspond à $2\epsilon_1$, de Cosserat; notre e_{23} correspond à γ_1, chez Cosserat, mais dans la formule qui remplace notre développement 57, Cosserat compte une fois seulement un terme tel que $\delta\gamma_1$ tandis que nous le comptons deux fois en considérant séparément δe_{23} et δe_{32}; ces deux différences de notation expliquent l'absence du facteur 2 dans les formules de Cosserat.

Le résultat que nous venons d'obtenir est susceptible d'une interprétation simple; reportons-nous à la remarque faite à la fin du paragraphe 5, et considérons, dans le corps déformé, les surfaces coordonnées $\overline{K}(\overline{X})$ qui se déduisent par déformation à partir des surfaces initiales; c'est le système de coordonnées *entraîné* par le corps pendant sa déformation; dans ce système, la densité tensorielle $\overline{\mathfrak{S}}^{rk}$ des efforts s'exprime en fonction des coefficients $\mathfrak{A}^{ij}$ par les relations très simples:

$$\overline{\mathfrak{S}}^{ij} = 2\mathfrak{A}^{ij}. \tag{63}$$

La formule (62) se réduit en effet à la formule de transformation de la densité tensorielle $\mathfrak{S}^{ij}$ lorsqu'on passe des axes rigides X aux axes entraînés $\overline{X}$.

12. *Définition d'une densité d'énergie.*

Les principes de la thermodynamique permettent de définir l'énergie d'un corps donné, comme fonction des variables mécaniques (déformations) et de la température; appelons, en effet, $d\mathfrak{T}_e$ le travail des forces extérieures, dQ la chaleur fournie, dE_{cin} la variation de l'énergie cinétique totale du corps lors d'une transformation infiniment petite.
Nous avons:

$$dQ + d\mathfrak{T}_e = dE_{cin} + dU,$$

U étant l'énergie interne du solide considéré.

D'autre part, le second principe nous donne, pour une transformation *réversible*:

$$dQ - TdS = 0,$$

S étant l'entropie du corps étudié; nous en déduisons les relations suivantes:

$$dU - TdS = d\mathfrak{T}_e - dE_{cin}, \tag{64}$$

$$d(U - TS) + SdT = d\mathfrak{T}_e - dE_{cin}. \tag{65}$$

Nous voyons que nous pouvons éliminer la variable température dans les deux cas suivants:

1° *Transformations adiabatiques.*

Le corps est supposé isolé thermiquement, de telle sorte qu'il ne puisse échanger d'énergie avec l'extérieur. Ce cas se réalise pratiquement si l'on étudie des mouvements vibratoires rapides, tels que les échanges de chaleur n'aient pas le temps de produire et d'égaliser les températures. On emploiera la relation (64), avec $dS=0$; on voit alors que la fonction U joue le rôle d'une énergie potentielle pour le corps solide; en décomposant le corps en petits éléments de volume, on pourra définir la *densité d'énergie* ϵ qui satisfera à la condition

$$U=\int_V \epsilon \, dx, \tag{66}$$

les coordonnées utilisées ici sont naturellement les *coordonnées entraînées* par le corps solide dans ses mouvements de telle sorte que les limites d'intégration, dans la formule (66) soient fixes et ne dépendent pas des déformations.

2° *Transformations isothermes.*

Le corps est supposé plongé dans une enceinte à température constante; les mouvements doivent être assez lents pour que les échanges de chaleur entre le corps et le thermostat aient toujours le temps d'intervenir pour uniformiser la température. On utilise alors la relation (65) avec $dT=0$, et l'on voit que la fonction $U-TS$ (potentiel thermodynamique) joue le rôle d'énergie potentielle; on définira, une *densité d'énergie* η par la condition:

$$U-TS=\int_V \eta \, dx, \tag{67}$$

en coordonnées entraînées.

On voit que, dans l'un et l'autre de ces deux cas, il est possible de définir une densité d'énergie ϵ ou η; ces deux grandeurs jouant, pour la suite des calculs, exactement de même rôle, nous ne les distinguerons pas plus loin, et nous parlerons d'une *densité d'énergie* ϵ.

La formule (57) du paragraphe précédent nous avait permis de définir une densité tensorielle $\mathfrak{A}^{ij}$; en comparant les formules (57) et (66) on voit aussitôt que, s'il existe une densité d'énergie ϵ, la densité tensorielle $\mathfrak{A}^{ij}$ a pour expression:

$$\mathfrak{A}^{ij}=\frac{\partial \epsilon}{\partial e_{ij}}. \tag{68}$$

Les formules (62) ou (63) nous donnent alors l'expression de la densité tensorielle Θ^{kl} qui représente les efforts internes, en fonction des dérivées de la densité d'énergie; par rapport aux axes entraînés, on a:

$$\overline{\Theta}^{ij}=2\frac{\partial \epsilon}{\partial e_{ij}}=\frac{\partial \epsilon}{\partial e_{ij}}+\frac{\partial \epsilon}{\partial e_{ji}}. \tag{63$^{\text{bis}}$}$$

En axes fixes $X^1X^2X^3$ on obtient les efforts par les formules:

$$\Theta^{kl}=\frac{1}{\Delta}a_i^k a_j^l\left(\frac{\partial \epsilon}{\partial e_{ij}}+\frac{\partial \epsilon}{\partial e_{ji}}\right). \tag{62$^{\text{bis}}$}$$

13. *Expression de la densité d'énergie en fonction de la déformation; loi de Hooke.*

Pour obtenir les lois de l'élasticité, il nous reste une dernière étape à franchir, en indiquant la manière dont la densité d'énergie s'exprime en fonction des déformations. Nous ne pourrons, sur ce point, que procéder par approximations, et former un développement en série, valable pour de petites déformations à partir d'un état initial donné.

Le choix de cet état initial est souvent trop complètement précisé, et on fait l'hypothèse que l'état initial est l'état naturel du corps solide, lorsqu'il n'est soumis à aucun effort extérieur.

Cette restriction n'est pas heureuse, car elle limite beaucoup les applications des formules que l'on établit ainsi. Nous préférons supposer, comme l'a fait H. Poincaré*, que l'état initial est un état quelconque, naturel ou déformé; nous obtiendrons alors pour la densité d'énergie un développement de la forme suivante:

$$\epsilon = \epsilon_0 + \Omega^{ij} e_{ij} + \Lambda^{ij.hk} e_{ij} e_{hk} + \ldots \tag{69}$$

le terme constant ϵ_0 ne joue aucun rôle; voyons ce que signifient les autres termes.

Les termes linéaires, qui sont groupés sous la forme $\Omega^{ij} e_{ij}$ n'existent que si, dans l'état initial, le corps est déjà soumis à des efforts extérieurs; ce groupe de termes disparait si l'on part de l'état naturel, sous efforts extérieurs nuls. Les formules (63^{bis}) nous montrent qu'au point $M(x^1, x^2, x^3)$, où la densité d'énergie est représentée par le développement (69) les efforts sont, dans l'état initial, donnés par la densité tensorielle:

$$\Theta^{ij} = \Omega^{ij} + \Omega^{ji} \tag{70}$$

les formules (63^{bis}) ou (62^{bis}) sont alors identiques, puisque, dans l'état pris comme état initial, les axes rigides ou les axes entraînés sont en coïncidence.

Les coefficients Ω^{ij} forment une *densité tensorielle*, et jouissent de la propriété de *symétrie*; ceci ressort aussitôt de leur mode de définition.

Si les efforts extérieurs se réduisent à une pression uniforme ϖ on obtient pour les Ω^{ij} les valeurs suivantes:

$$\Omega^{ij} = -\frac{\overline{\varpi}}{2} g^{ij},$$

$$\Omega^{ij} e_{ij} = -\frac{\overline{\varpi}}{2} g^{ij} e_{ij} = -\frac{\overline{\varpi}}{2} e_i^i;$$

l'ensemble des termes linéaires se réduit au produit de $-\frac{\overline{\varpi}}{2}$ par l'invariant linéaire e_i^i (voir § 6). Ce point se vérifie immédiatement, si l'on se rappelle que, en axes rectangulaires, le tableau des composantes du pseudo-tenseur des efforts doit se réduire, dans le cas d'une pression uniforme, à

$$\begin{matrix} -\overline{\varpi} & 0 & 0 \\ 0 & -\overline{\varpi} & 0 \\ 0 & 0 & -\overline{\varpi}, \end{matrix}$$

*H. Poincaré: *Leçons sur la théorie de l'élasticité*; G. Carré, Paris (1892) et Cosserat, *loc. cit.*

Il importe de compter combien de coefficients distincts s'introduisent dans le cas le plus général; leur nombre est celui des composantes d'un tenseur double symétrique, ce qui nous donne 6 composantes distinctes. Nous devons faire ici une remarque importante, qui justifie le mode de définition que nous avons choisi (§ 4) pour les déformations; nous insistions sur la nécessité de garder dans le développement du tenseur de déformation [formule (32)] tous les termes nécessaires, d'ordre supérieur au premier; il est facile de voir que, si nous développons la densité d'énergie jusqu'aux termes d'ordre n par rapport aux déformations [$n=2$ dans la formule (69)] il nous faudra, dans les termes linéaires, par rapport aux déformations, écrire les déformations en les développant jusqu'aux termes d'ordre n par rapport aux déplacements u ou à leurs dérivées. En particulier, si nous ne voulons pas négliger des termes du même ordre que ceux gardés autre part, il est indispensable, dans la formule (69) de prendre les expressions des déformations jusqu'aux termes du 2^{e} degré.

Les termes du second degré, par rapport aux déformations, groupés sous la forme $\Lambda^{ij,hk} e_{ij} e_{hk}$ correspondent à des efforts proportionnels aux déformations, ces dernières étant comptées à partir de l'état initial; c'est le cas des petites déformations et de la loi de Hooke. Pour des corps cristallisés, ou pour des corps soumis initialement à des efforts quelconques, il y aura un nombre élevé de coefficients distincts; *la densité tensorielle* $\Lambda^{ij,hk}$ ne se trouve alors soumise qu'aux conditions de *symétrie* pour les deux groupes d'indices, ij d'une part et hk de l'autre; ces conditions sont

$$\Lambda^{ij,hk}=\Lambda^{ji,hk}=\Lambda^{ij,kh}=\Lambda^{ji,kh}. \tag{71}$$

Nous pouvons compter aisément le nombre des composantes indépendantes: il y a 6 grandeurs e_{ij}, indépendantes; l'ensemble des termes que nous étudions forme un polynome homogène du 2^{e} degré par rapport à ces 6 déformations; le nombre des coefficients de ce polynome est:

$$\frac{6(6+1)}{2}=21.$$

Il est important de considérer les simplifications qui se produisent, si l'état initial est isotrope; ceci suppose un corps isotrope soumis, dans l'état initial, à une pression uniforme. Dans ces conditions, on ne doit voir apparaître, dans les termes du second degré, que les combinaisons invariantes. Nous avons vu (§ 6) que les combinaisons invariantes du 2^{e} degré sont:

$$(e_i^i)^2 \text{ et } e_i^j e_j^i;$$

le développement (69) se réduira alors à:

$$\epsilon=\epsilon_0-\frac{\varpi}{2}e_i^i+\frac{\lambda}{8}(e_i^i)^2+\frac{\mu}{4}e_i^j e_j^i+\dots \tag{72}$$

Dans l'invariant e_i^i on aura soin, comme nous le rappelons plus haut, de garder les expressions des déformations développées jusqu'aux termes du 2^{e} ordre; les coefficients λ et μ ont leur sens habituel et représentent deux densités scalaires; comparons la formule ci-dessus aux formules de Cosserat, nous constatons les correspondances suivantes*:

L. Brillouin:	$-\frac{\varpi}{2}$	λ	μ
E. F. Cosserat:	2γ	4λ	4μ

le coefficient numérique 4 tient aux différences de définitions sur lesquelles nous avons déjà insisté [§4, formules (29); §6, formule (40^{bis}); § 11, formule (62)].

Il n'est peut-être pas inutile de montrer quelle est, dans le cas de l'isotropie, la forme de la densité tensorielle $\Lambda^{ij,hk}$. Pour comparer les formules (69) et (72), il nous faut passer aux composantes mixtes:

$$\Lambda^{i,h}_{r,s,} = g_{jr}\, g_{ks}\, \Lambda^{ij,hk}.$$

Nous voyons alors que ces composantes mixtes ne peuvent prendre que les valeurs distinctes suivantes:

$$\frac{\lambda}{8} + \frac{\mu}{4}, \qquad \text{pour} \qquad i = r = h = s,$$

$$\frac{\lambda}{8}, \qquad \text{dans le cas} \qquad i = r \neq h = s,$$

$$\frac{\mu}{4}, \qquad \text{pour} \qquad i = s \neq r = h,$$

et 0 pour les autres cas. Ceci correspond évidemment à un cas de dégénérescence assez curieux; le tenseur à 4 indices a, en général, 81 composantes; les conditions de symétrie (71) avaient réduit ce nombre à 21, et l'isotropie nous ramène à 2 seulement.

14. *Conclusions.*

J'ai tenu, dans cette étude, à écrire complètement les équations de l'élasticité, sans faire aucune hypothèse simplificatrice dans le choix des axes de coordonnées; j'ai dû, tout d'abord (§ 2) préciser nettement la nature des grandeurs qui s'introduisent, et j'ai distingué, à coté des *tenseurs*, deux types de *pseudo-tenseurs*, qui sont les densités tensorielles, déjà connues, et les *capacités tensorielles*; ces dernières grandeurs n'avaient pas été, semble-t-il, nettement distinguées jusqu'à présent. Les densités ou capacités tensorielles se déduisent de tenseurs ordinaires symétriques gauches, si l'on choisit un ordre de succession des indices, c'est-à-dire un sens de rotation des axes; cette méthode de formation, indiquée par M. P. Langevin dans ses conférences, joue un rôle important dans les définitions de nombreuses grandeurs physiques. Pour l'étude de l'élasticité, et de tous les phénomènes où interviennent les déformations, on est obligé d'introduire la notion de dérivation covariante, appliquée aux tenseurs, aux densités et aux capacités tensorielles; j'ai rappelé (§ 3) et établi les diverses formules nécessaires.

L'étude des déformations (§ 4, 5 et 6) peut alors être abordée, suivant la méthode de Cosserat; le tenseur des déformations se définit par la variation du tenseur fondamental des g_{ik}; cette méthode permet d'obtenir les déformations même grandes, en gardant tous les termes nécessaires dans le développement;

*E. F. Cosserat, *loc. cit.*, p. 71, formules 69 et 70.

j'insiste, à propos des développements classiques de l'énergie (§ 13) sur la nécessité de se reporter à ces définitions tout à fait générales, faute de quoi le développement de la théorie n'est pas cohérent.

Après l'étude des déformations vient celle des efforts (§ 7 et 8); ceux-ci forment un tenseur symétrique gauche à 3 indices qui se laisse ramener, suivant la méthode de réduction rappelée plus haut, à une densité tensorielle; cette dernière jouit en outre de la propriété de symétrie, si l'on suppose que les forces extérieures n'exercent pas de couple sur un élément de volume infiniment petit; cette hypothèse est classique en élasticité, mais peut ne pas être valable pour d'autres problèmes de physique (actions magnétiques, par exemple).

Les forces sur un élément de volume (densité de forces, § 9) s'obtiennent en prenant la divergence du pseudo-tenseur des efforts, défini ci-dessus; si le solide déformé est, non pas en équilibre statique, mais en mouvement, ce qui se produit dans les phénomènes vibratoires, il faut ajouter au pseudo-tenseur des efforts celui du flux de quantité de mouvement (§ 10); ceci ne modifie pas la mise en équation générale.

Quelques formules permettent alors (§ 11) d'établir une relation entre les pseudo-tenseurs des efforts et le travail nécessaire pour une petite variation de la déformation; ces relations présentent une certaine complexité, inhérente au problème lui-même; l'écriture tensorielle a d'ailleurs le très grand avantage de permettre de grouper les résultats sous une forme très compacte, et d'en faire immédiatement ressortir le sens physique; la nature des relations trouvées correspond bien aux genres de définitions adoptées, et au classement des grandeurs en tenseurs, densités ou capacités tensorielles.

Ces formules très générales trouvent leur application, si l'on introduit la notion de *densité d'énergie* (§ 12); elles donnent l'expression du tenseur des efforts en fonction des dérivées partielles de la densité d'énergie; il faut seulement bien définir le problème à étudier, et distinguer le cas adiabatique de celui des transformations isothermes.

Pour compléter la théorie, il est d'usage d'écrire le développement de la densité d'énergie en fonction des déformations (§ 13); j'insiste, à ce sujet, sur la possibilité de partir d'un état initial qui ne soit pas l'état naturel sous efforts nuls; les développements s'écrivent aussi bien à partir d'un état initial déjà déformé; le cas d'un corps isotrope sous pression uniforme introduit, d'ailleurs, d'importantes simplifications.

Il n'était pas inutile, je crois, d'écrire en tenseurs les formules d'élasticité; ce travail a nécessité un classement précis* des diverses grandeurs physiques, et a résolu le problème que se posait déjà Lamé, d'appliquer à l'élasticité les coordonnées curvilignes les plus générales, de manière à pouvoir choisir, pour chaque corps cristallisé ou chaque nature de conditions aux limites les coordonnées les mieux adaptées et les plus favorables.

*Cette distinction n'est pas faite, par exemple, dans le mémento de Madelung (*Die mathematischen Hilfsmittel des Physikers*, J. Springer, Berlin, 1922).

L'auteur (p. 184) y introduit seulement des tenseurs, et ne s'aperçoit pas des différences qui caractérisent les pseudo-tenseurs; les formules écrites n'y sont d'ailleurs valables qu'en coordonnées rectilignes, et les définitions de la déformation s'arrêtent aux termes du 1er ordre.

SUR LA PROPAGATION DES ONDES PLANES DANS LES MILIEUX ÉLASTIQUES ANISOTROPES

PAR M. LE COMMANDANT BARRÉ,
Ministère de la Guerre, Paris, France.

I.

POSITION DE LA QUESTION.

Le présent travail a pour but d'étendre aux milieux homogènes anisotropes généraux, régis par la loi de Hooke, les résultats obtenus par de nombreux savants pour les mouvements de l'éther lumineux. Les principes fondamentaux de la présente étude ont été examinés, à notre connaissance par Blanchet et par Green*.

Rappelons que, dans les milieux de Green, tout mouvement par ondes planes peut être considéré comme le mouvement résultant de trois mouvements élémentaires régis par l'équation du son. Nous étudierons, ici, plus spécialement, ceux de ces mouvements qui se propagent dans le même sens pour chaque direction.

Les conclusions de Green s'étendent d'ailleurs à tous les milieux régis par la loi de Hooke, à cela près que les mouvements composants ne sont plus rectangulaires (et à l'exception de milieux spéciaux que nous laisserons de côté). Ces résultats sont résumés par les théorèmes I et II ci-après, que nous avons énoncés dans cette étude, bien que déjà connus, afin d'obtenir un ensemble cohérent.

II.

ÉQUATION GÉNÉRALE DU MOUVEMENT INTÉRIEUR D'UNE ONDE PLANE SIMPLE DANS UN MILIEU HOMOGÈNE ANISOTROPE QUELCONQUE. (Pas de force appliquée à l'élément de volume).

Nous sommes donc conduits, par application des considérations précédentes, à examiner la propagation des mouvements définis en coordonnées cartésiennes quelconques par les formules:

$$(1)\qquad \begin{cases} u = \phi_1(\alpha x+\beta y+\gamma z-pt), \\ v = \phi_2(\alpha x+\beta y+\gamma z-pt), \\ w = \phi_3(\alpha x+\beta y+\gamma z-pt). \end{cases}$$

*On peut rapprocher de ces études celles qu'a faites M. Boussinesq sur les *équations des petits mouvements des milieux isotropes comprimés* (C. R. de l'Académie des Sciences de Paris, 22 Juillet 1867—Journal de Mathématiques de 1868. T. XIII, p. 209 à 241).

Afin d'éviter des redites, nous supposerons uniformément les axes rectangulaires, ce qui ne restreint en rien la généralité, un changement de coordonnées cartésiennes conservant à u, v, w leur forme du type

$$F(\alpha x+\beta y+\gamma z-pt).$$

Cette remarque est utile pour éviter des confusions dans l'étude des questions métriques que nous aurons à examiner.

Les équations générales du mouvement, en l'absence de forces de volume, sont, avec les notations de Lamé, et en désignant par ρ la densité du milieu:

$$\rho\frac{\partial^2 u}{\partial t^2}=\frac{\partial N_1}{\partial x}+\frac{\partial T_3}{\partial y}+\frac{\partial T_2}{\partial z} \tag{1^{bis}}$$

et deux analogues. Les N et les T sont déterminés par les formules suivantes que nous donnons pour fixer les notations:

$$\begin{cases} N_i=A_1^i\epsilon_1+A_2^i\epsilon_2+A_3^i\epsilon_3+B_1^i\gamma_1+B_2^i\gamma_2+B_3^i\gamma_3, & (i=1,\ 2,\ 3), \\ T_i=a_1^i\epsilon_1+a_2^i\epsilon_2+a_3^i\epsilon_3+b_1^i\gamma_1+b_2^i\gamma_2+b_3^i\gamma_3, & (i=1,\ 2,\ 3), \end{cases} \tag{2}$$

avec :

$$\epsilon_1=\frac{\partial u}{\partial x},\quad \epsilon_2=\frac{\partial v}{\partial y},\quad \epsilon_3=\frac{\partial w}{\partial z},$$

$$\gamma_1=\frac{\partial w}{\partial y}+\frac{\partial v}{\partial z},\quad \gamma_2=\frac{\partial u}{\partial z}+\frac{\partial w}{\partial x},\quad \gamma_3=\frac{\partial u}{\partial y}+\frac{\partial v}{\partial x}.$$

Dans le cas des milieux de Green les 36 constantes des formules (2) se réduisent à 21, eu égard à la symétrie du déterminant des ϵ, γ. Nous les désignons par les notations:

$$\boldsymbol{A_i},\ B_j$$

en prenant:

$$\begin{cases} A_1^1=A_{(1)},\quad A_2^2=A_{(2)},\quad A_3^3=A_{(3)},\quad b_1^1=A_{(4)},\quad b_2^2=A_{(5)},\quad b_3^3=A_{(6)}, \\ A_1^2=A_2^1=B_{(15)},\quad A_3^2=A_2^3=B_{(10)},\quad B_1^3=a_3^1=B_{(6)},\quad B_2^1=a_1^2=B_{(12)}, \\ A_3^1=A_1^3=B_{(14)},\quad B_1^2=a_2^1=B_{(9)},\quad B_2^3=a_3^2=B_{(5)},\quad B_3^2=a_2^3=B_{(7)}, \\ B_1^1=a_1^1=B_{(12)},\quad B_2^2=a_2^2=B_{(8)},\quad B_3^3=a_3^3=B_{(4)},\quad B_3^1=a_1^3=B_{(11)}, \\ \qquad b_1^2=b_2^1=B_{(3)},\quad b_3^1=b_1^3=B_{(2)},\quad b_3^2=b_2^3=B_{(1)}. \end{cases} \tag{2^{bis}}$$

L'application des équations (1^{bis}) et (2) aux déplacements (1) conduit aux relations:

$$\begin{cases} \rho p^2\phi_1''=\mathfrak{A}_1^1\phi_1''+\mathfrak{A}_1^2\phi_2''+\mathfrak{A}_1^3\phi_3'', \\ \rho p^2\phi_2''=\mathfrak{A}_2^1\phi_1''+\mathfrak{A}_2^2\phi_2''+\mathfrak{A}_2^3\phi_3'', \\ \rho p^2\phi_3''=\mathfrak{A}_3^1\phi_1''+\mathfrak{A}_3^2\phi_2''+\mathfrak{A}_3^3\phi_3''. \end{cases} \tag{3}$$

avec :

$$
(4) \qquad \begin{cases}
\mathfrak{A}_1^1 = A_1^1\alpha^2 + b_3^3\beta^2 + b_2^2\gamma^2 + (B_2^1 + a_1^2)\alpha\gamma + (B_3^1 + a_1^3)\alpha\beta + (b_2^3 + b_3^2)\beta\gamma, \\
\mathfrak{A}_1^2 = B_3^1\alpha^2 + a_2^3\beta^2 + b_1^2\gamma^2 + (b_1^3 + a_2^2)\gamma\beta + (B_1^1 + b_3^2)\alpha\gamma + (A_2^1 + b_3^3)\alpha\beta, \\
\mathfrak{A}_1^3 = B_2^1\alpha^2 + b_1^3\beta^2 + a_3^2\gamma^2 + (a_3^3 + b_1^2)\beta\gamma + (A_3^1 + b_2^2)\alpha\gamma + (B_1^1 + b_2^3)\alpha\beta, \\
\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots
\end{cases}
$$

Si on laisse de côté les solutions $\phi_1''=0, \phi_2''=0, \phi_3''=0$, qui donnent pour u, v, w des expressions linéaires, donc susceptibles de s'accroître indéfiniment et rendant, par suite, illusoire l'application de la loi de Hooke, on est conduit à écrire la relation:

$$
(5) \qquad \| -\rho h^2\omega^2 + \mathfrak{A}_1^1 \quad \mathfrak{A}_1^2 \quad \mathfrak{A}_1^3 \| = 0
$$

dans le premier membre de laquelle figure un déterminant Δ représenté par sa première ligne et où l'on a posé:

$$
(6) \qquad p^2 = h^2\omega^2
$$

avec

$$
h^2 = \alpha^2 + \beta^2 + \gamma^2.
$$

La quantité ω ainsi introduite n'est autre que la vitesse de propagation de ces ondes.

Il résulte immédiatement de l'équation (5) la proposition suivante:

Théorème I. Dans un milieu anisotrope homogène quelconque, il existe en général, pour une direction de plan donnée, trois vitesses de propagation ω, distinctes, à chacune desquelles correspond une vibration polarisée se propageant suivant un mouvement ondulatoire plan parallèle au plan donné.

Cas des milieux de Green.—Un cas particulièrement intéressant est celui des *milieux de Green*. Le déterminant (5) est alors symétrique et si l'on pose $S=\rho h^2\omega^2$ S n'est autre que l'une des racines de l'équation en S d'une quadrique à centre (Σ) d'équation:

$$
(7) \qquad Ax^2 + A'y^2 + A''z^2 + 2Byz + 2B'zx + 2B''xy = 1
$$

que nous désignerons sous le nom de quadrique de polarisation.

Dans le cas où le corps admet en chaque point trois plans de symétrie rectangulaires, les expressions des A et des B sont:

$$
(8) \qquad \begin{cases}
A \ \ = A_1\alpha^2 + a_3\beta^2 + a_2\gamma^2, & B \ \ = (B_1 + a_1)\beta\gamma, \\
A' \ = a_3\alpha^2 + A_2\beta^2 + a_1\gamma^2, & B' \ = (B_2 + a_2)\gamma\alpha, \\
A'' = a_2\alpha^2 + a_1\beta^2 + A_3\gamma^2, & B'' = (B_3 + a_3)\alpha\beta,
\end{cases}
$$

équations dans lesquelles les A_i et les a_i sont les constantes de définition du milieu liées aux N et aux T par les relations:

$$
(9) \qquad \begin{cases}
N_1 = A_1\epsilon_1 + B_3\epsilon_2 + B_2\epsilon_3, & T_1 = a_1\gamma_1, \\
N_2 = B_3\epsilon_1 + A_2\epsilon_2 + B_1\epsilon_3, & T_2 = a_2\gamma_2, \\
N_3 = B_2\epsilon_1 + B_1\epsilon_2 + A_3\epsilon_3, & T_3 = a_3\gamma_3.
\end{cases}
$$

III.

Réduction canonique des expressions de u, v, w.

Lorsque la racine S annule le déterminant (5) mais non tous ses mineurs, on peut écrire, grâce à une double intégration:

$$\phi_1 = l.f(\alpha x+\beta y+\gamma z-pt)+\theta_1,$$
$$\phi_2 = m.f(\alpha x+\beta y+\gamma z-pt)+\theta_2,$$
$$\phi_3 = n.f(\alpha x+\beta y+\gamma z-pt)+\theta_3,$$

en désignant par θ_1, θ_2, θ_3 trois fonctions arbitraires de t du premier degré en t, par f une fonction quelconque et par l, m, n trois constantes, proportionnelles aux mineurs de Δ.

Mais en remarquant que ϕ_1, ϕ_2, ϕ_3 doivent rester assez faibles pour l'application de la loi de Hooke, on est conduit à ne retenir que les solutions où θ_1, θ_2, θ_3 sont des constantes, que l'on peut supposer nulles, leur intervention correspondant à une translation d'ensemble indépendante du temps. En définitive, dans ce cas, on peut, sans restreindre la généralité du problème traité, prendre:

$$(1') \qquad \begin{cases} u = l.f(\alpha x+\beta y+\gamma z-pt), \\ v = m.f(\alpha x+\beta y+\gamma z-pt), \\ w = n.f(\alpha x+\beta y+\gamma z-pt). \end{cases}$$

Cette réduction n'est plus possible pour une racine de (5) annulant les mineurs. Si elle n'annule pas les coefficients, tout ce que permettent de dire les considérations précédentes, c'est que l'on peut toujours considérer u, v, w, et non pas seulement leurs dérivées secondes, comme liées par la relation linéaire unique qui lie ces dérivées. Si la racine considérée annule tous les éléments les ϕ, comme les ϕ'' d'ailleurs, conservent toute liberté. Il va de soi que, même dans ces cas, les solutions du type (1) sont particulièrement intéressantes. Dans le cas général, l, m, n sont proportionnels à trois quantités données; dans le cas de l'annulation des mineurs, ces paramètres sont simplement liés par une relation linéaire et homogène et, dans le dernier cas, indépendants; dans tous les cas f est arbitraire.

IV.

Discussion de la nature des ondes dans les milieux de Green.

1° La quadrique Σ est un *ellipsoide imaginaire*. Les racines S sont négatives et ω est une imaginaire pure. *Lorsque les ondes sont pendulaires, ce sont les ondes évanescentes élémentaires.* C'est le cas où l'énergie de déformation est une forme définie négative.

2° La quadrique Σ est un *ellipsoïde réel*. Il y a trois vitesses réelles; à chacune desquelles correspond une direction de vibration: celle de l'axe de Σ correspondant à la valeur condsiérée de ω. C'est le cas où l'énergie de déformation est définie positive.

3° La quadrique Σ est un *hyperboloïde à une nappe*: deux systèmes d'ondes polarisées rectangulaires. La troisième est à vitesse imaginaire.

4° La quadrique Σ est un *hypreboloïde à deux nappes*: une seule direction d'ondes polarisées à vitesse réelle deux sont à vitesse imaginaire.

5° Dans le cas où la surface Σ est de révolution, sans être une sphère, il y a une direction de vibration polarisée parallèlement à l'axe. Les deux autres vibrations sont de direction indéterminée mais néanmoins perpendiculaires à l'axe. Si la quadrique est une sphère, il n'y a plus aucune polarisation.

6° La surface Σ est *un cylindre*. Une des vitesses ω est nulle; la direction qui serait polarisée parallèlement à l'axe du cylindre ne peut se transmettre. Les autres se comportent comme dans les cas précédents. Si toutefois le cylindre se réduit à *deux plans parallèles*, seule se propage une vibration polarisée perpendiculairement à leur direction.

Cas ou la quadrique de polarisation est indépendante de la direction de l'onde

En général la quadrique de polarisation dépend de la direction du plan d'onde. Supposons-la indépendante de cette direction; en rapportant le milieu à des axes parallèles à trois directions principales de la quadrique de polarisation, l'équation de (Σ) doit être de la forme

$$Ax^2+By^2+Cz^2=1, \tag{10}$$

où A, B, C sont des constantes indépendantes de α, β, γ. Mais, en fait, les coefficients de cette quadrique sont donnés, dans le cas d'un milieu de Green, par les relations:

$$\tag{11} \begin{cases} \mathfrak{A}_1^1=A_{(1)}\alpha^2+A_{(6)}\beta^2+A_{(5)}\gamma^2+2B_{(1)}\beta\gamma+2B_{(12)}\gamma\alpha+2B_{(11)}\alpha\beta, \\ \mathfrak{A}_2^1=\mathfrak{A}_1^2=B_{(11)}\alpha^2+B_{(7)}\beta^2+B_{(3)}\gamma^2+[B_{(8)}+B_{(2)}]\beta\gamma+[B_{(1)}+B_{(13)}]\gamma\alpha \\ \qquad\qquad +[B_{(15)}+A_{(6)}]\alpha\beta, \\ \mathfrak{A}_1^3=\mathfrak{A}_3^1=B_{(12)}\alpha^2+B_{(2)}\beta^2+B_{(5)}\gamma^2+[B_{(4)}+B_{(3)}]\beta\gamma+[B_{(14)}+A_{(5)}]\gamma\alpha \\ \qquad\qquad +[B_{(13)}+B_{(1)}]\alpha\beta, \\ \mathfrak{A}_2^2=A_{(6)}\alpha^2+A_{(2)}\beta^2+A_{(4)}\gamma^2+2B_{(9)}\beta\gamma+2B_{(2)}\gamma\alpha+2B_{(7)}\alpha\beta, \\ \mathfrak{A}_2^3=\mathfrak{A}_3^2=B_{(1)}\alpha^2+B_{(9)}\beta^2+B_{(6)}\gamma^2+[B_{(10)}+A_{(4)}]\beta\gamma+[B_{(4)}+B_{(3)}]\alpha\gamma \\ \qquad\qquad +[B_{(2)}+B_{(8)}]\alpha\beta, \\ \mathfrak{A}_3^3=A_{(5)}\alpha^2+A_{(4)}\beta^2+A_{(3)}\gamma^2+2B_{(6)}\beta\gamma+2B_{(5)}\gamma\alpha+2B_{(3)}\alpha\beta, \end{cases}$$

ou $A_{(1)}, \ldots, A_{(6)}, B_{(1)}, \ldots, B_{(15)}$ sont les 21 constantes du milieu de Green.

Pour que l'équation de la quadrique (Σ) puisse être ramenée à la forme (10) il faut que les $A_i^j (i \neq j)$ soient nuls et que les A_i^i soient indépendants de α, β, γ, ce qui exige l'annulation des coefficients rectangles et l'égalité* des autres coefficients dans A_1^1, A_2^2 et A_3^3. On déduit de là que *la surface* Σ *se réduit alors nécessairement à une sphère*†, car:

$$\mathfrak{A}_1^1=\mathfrak{A}_2^2=\mathfrak{A}_3^3=A(\alpha^2+\beta^2+\gamma^2),$$

formule dans laquelle nous désignons par A la valeur commune des $\mathfrak{A}_i$. Si on tient compte de la nullité des B qui figurent dans $\mathfrak{A}_1^1$, $\mathfrak{A}_2^2$ et $\mathfrak{A}_3^3$, on voit que les autres coefficients se réduisent à:

$$(12)\qquad \begin{cases} \mathfrak{A}_1^2=\mathfrak{A}_2^1=B_{(8)}\beta\gamma+B_{(13)}\gamma\alpha+[B_{(15)}+A]\alpha\beta, \\ \mathfrak{A}_1^3=\mathfrak{A}_3^1=B_{(4)}\beta\gamma+[B_{(14)}+A]\alpha\gamma+B_{(13)}\alpha\beta, \\ \mathfrak{A}_2^3=\mathfrak{A}_3^2=[B_{(10)}+A]\gamma\beta+B_{(4)}\gamma\alpha+B_{(8)}\alpha\beta. \end{cases}$$

En les annulant, on trouve:

$$(13)\qquad B_{(10)}=B_{(14)}=B_{(15)}=-A$$

et tous les autres nuls. Ces coefficients correspondent au milieu isotrope très spécial pour lequel les coefficients λ et μ de Lamé sont liés par la relation:

$$\lambda+\mu=0,$$

et pour lequel, comme on sait, l'onde longitudinale a même vitesse de propagation que l'onde transversale.

Donc:

Sauf pour les milieux isotropes spéciaux tels que $\lambda+\mu=0$, *la quadrique de la polarisation varie toujours avec la direction du plan d'onde.*

V.

Enveloppe du plan d'onde a un instant donné t_0.

Si l'on pose:

$$h\omega t_0=pt_0=-r,$$

on obtient, en partant de (5) l'équation tangentielle

$$(14)\qquad \| -\rho r^2+\mathfrak{A}_1^1 t_0^2 \quad \mathfrak{A}_1^2 \quad \mathfrak{A}_1^3 \| =0$$

qui represénte une surface de sixième classe.

VI.

Angle de la vibration et du plan d'onde.

Cet angle est égal à celui que fait le plan (α, β, γ) avec l'axe de (Σ) correspondant à la vibration considérée.

Donc:

*Ce dernier point devient évident si on remarque que, lorsque α, β, γ sont les cosinus directeurs eux-mêmes, $\alpha^2+\beta^2+\gamma^2=1$.

†Ce qui est à peu près évident a priori.

Théorème II.—Sauf le cas où le plan α, β, γ est parallèle à l'un des axes de la quadrique qui lui correspond, la vibration n'est ni transversale ni longitudinale.*

VII.

Recherche des directions de plans d'onde donnant des vibrations transversales.

La condition de transversalité est:

$$\alpha u+\beta v+\gamma w=0$$

d'où l'on déduit:

(15) $$\alpha\phi_1''+\beta\phi_2''+\gamma\phi_3''=0,$$

relation qui, pour les solutions du type (1′), peut remplacer la précédente.

Cela posé, si on multiplie les équations (3) respectivement par α, β, γ et qu'on les ajoute, on trouve, compte tenu de (15),

(16) $$\mathfrak{A}\phi_1''+\mathfrak{B}\phi_2''+\mathfrak{C}\phi_3''=0$$

en posant:

(17) $$\begin{cases}\mathfrak{A}=\alpha\mathfrak{A}_1^1+\beta\mathfrak{A}_2^1+\gamma\mathfrak{A}_3^1,\\ \mathfrak{B}=\alpha\mathfrak{A}_1^2+\beta\mathfrak{A}_2^2+\gamma\mathfrak{A}_3^2,\\ \mathfrak{C}=\alpha\mathfrak{A}_1^3+\beta\mathfrak{A}_2^3+\gamma\mathfrak{A}_3^3.\end{cases}$$

D'ailleurs une combinaison évidente des premières équations (3) donne:

(18) $$\mathfrak{A}_2^1(\phi_1'')^2-\mathfrak{A}_1^2(\phi_2'')^2+(\mathfrak{A}_2^2-\mathfrak{A}_1^1)(\phi_1'')(\phi_2'')+\mathfrak{A}_2^3(\phi_1'')(\phi_3'')-\mathfrak{A}_1^3(\phi_2'')(\phi_3'')=0$$

et des équations (15) et (16) on peut tirer, sauf cas exceptionnels, que nous n'examinerons pas pour ne pas allonger cette étude,

(19) $$\frac{\phi_1''}{\beta\mathfrak{C}-\gamma\mathfrak{B}}=\frac{\phi_2''}{\gamma\mathfrak{A}-\alpha\mathfrak{C}}=\frac{\phi_3''}{\alpha\mathfrak{B}-\beta\mathfrak{A}}.$$

En introduisant dans la relation (18) les dénominateurs des rapports (19) à la place des ϕ'' correspondants, on trouve une équation homogène et du dixième degré en α, β, γ.

Donc:

Théorème III.—En général, les plans menés par l'origine parallèlement aux directions des plans d'onde pour lesquels il existe une vibration transversale, enveloppent un cône de dixième classe.

*Nous considérons ici comme transversale une onde dont la vibration est dans le plan d'onde, comme longitudinale une onde dont la vibration est normale à l'onde.

Il y a identité avec la définition fondée sur la nullité de la dilatation cubique pour les vibrations transversales, ou du rotationnel pour les vibrations longitudinales, lorsque *u*, *v* et *w* sont du type (1′), ce qui est le cas le plus intéressant, mais non dans les autres cas.

VIII.

Recherches des directions de plans d'onde donnant des vibrations longitudinales.

Pour qu'il y ait une vibration longitudinale, il faut et il suffit évidemment que les équations (3) soient satisfaites quand on y remplace ϕ_1'', ϕ_2'' et ϕ_3'' respectivement par α, β et γ.

1° $\alpha\beta\gamma \neq 0$. En éliminant ρp^2 entre les équations (3) on trouve deux équations homogènes du quatrième degré, donnant seize directions de plans d'onde satisfaisant à la question. Dans le cas des milieux de Green, les deux autres vibrations trouvées sont transversales.

2° $\alpha\beta\gamma = 0$, soit, par exemple, $\alpha = 0$.

La première équation (3) se réduit à la suivante, homogène et du troisième degré en β et γ:

(20) $$\mathfrak{A}_1^2\beta + \mathfrak{A}_1^3\gamma = 0.$$

Si l'on suppose $\beta\gamma \neq 0$ on peut éliminer $\rho\, p^2$ entre les deux autres et l'on trouve la relation homogène et du quatrième degré en β et γ:

(21) $$(\mathfrak{A}_2^2\beta + \mathfrak{A}_2^3\gamma)\gamma = (\mathfrak{A}_3^2\beta + \mathfrak{A}_3^3\gamma)\gamma.$$

Les relations (20) et (21) qui définінissent toutes deux le rapport $\frac{\beta}{\gamma}$ seront en général incompatibles. Nous n'étudierons pas, dans le cas général, les conditions de compatibilité. Même conclusions si deux des coefficients α, β ou γ sont nuls.

De la discussion qui précède on déduit la proposition suivante:

Théorème IV.—Dans un milieu anisotrope quelconque, il y a, en général, seize directions de plans d'onde pour lesquelles il existe une vibration longitudinale. Dans le cas des milieux de Green les autres sont alors transversales.

IX.

Cas des milieux de Green présentant trois plans de symétrie de contexture.

Dans le cas des milieux de Green à trois plans de symétrie de contexture et si l'on rapporte le milieu à ses plans de symétrie (cf. n° II), les équations générales (3) deviennent:

(22) $$\begin{cases} (-\rho p^2 + A_1\alpha^2 + a_3\beta^2 + a_2\gamma^2)\phi_1'' + (B_3 + a_3)\alpha\beta\phi_2'' + (B_2 + a_2)\alpha\gamma\phi_3'' = 0, \\ (B_3 + a_3)\alpha\beta\phi_1'' + (-\rho p^2 + a_3\alpha^2 + A_2\beta^2 + a_1\gamma^2)\phi_2'' + (B_1 + a_1)\beta\gamma\phi_3'' = 0, \\ (B_2 + a_2)\alpha\gamma\phi_1'' + (B_1 + a_1)\beta\gamma\phi_2'' + (-\rho p^2 + a_2\alpha^2 + a_1\beta^2 + A_3\gamma^2)\phi_3'' = 0. \end{cases}$$

A. *Directions de plans auxquels correspondent une vibration longitudinale et par suite des vibrations transversales.* (Cf. N° VIII—Cas des milieux de Green).

1° Une première série de solutions est donnée par les équations:

$$(23) \qquad A_1\alpha^2+(B_3+2a_3)\beta^2+(B_2+2a_2)\gamma^2=(B_3+2a_3)\alpha^2+A_2\beta^2+(B_1+2a_1)\gamma^2$$
$$=(B_2+2a_2)\alpha^2+(B_1+2a_1)\beta^2+A_3\gamma^2.$$

Ces équations fournissent seulement quatre directions de plans d'ailleurs symétriques par rapport aux éléments de symétrie du système.

Ces équations du premier degré en α^2, β^2, γ^2 sont donc immédiatement résolubles et d'une discussion facile. Sauf le cas d'indétermination étudié ci-après (Rem. 2) il y aura toujours une solution et une seule en α^2, β^2, γ^2. Elle donnera quatre directions, réelles si les déterminants auxquels α^2, β^2 et γ^2 sont proportionnels sont de même signe (un ou deux pouvant être nuls)*, imaginaires dans le cas contraire. Si tous ces mineurs sont nuls c'est l'indétermination signalée.

2° Une deuxième série de solutions s'obtient en supposant $\alpha\beta\gamma$ nul, soit, par exemple, $\alpha=0$

(*a*) Si on suppose $\beta\gamma\neq 0$, tandis que la première équation (22) se réduit à une identité ($\phi_1''=\alpha=0$) les deux autres donnent, par élimination de ρp^2 et après remplacement de ϕ_2'' et de ϕ_3'' par les quantités proportionnelles β et γ, la relation suivante, homogène et du second degré en β et γ:

$$(24) \qquad (B_1+2a_1)(\beta^2-\gamma^2)=A_2\beta^2-A_3\gamma^2,$$

qui, jointe à $\alpha=0$,définit deux directions du plan yoz qui sont les directions de deux normales à deux ondes, solutions de la question.

En tenant compte des trois plans de symétrie, on trouve ainsi six directions nouvelles

(*b*) Si $\alpha=\beta=0$, avec $\phi_1''=\phi_2''=0$ les équations (22) sont satisfaites. Les trois directions d'axes sont donc encore trois solutions.

En résumé, on trouve donc en tout treize solutions dont trois, celles des axes, toujours réelles et distinctes. Une discussion assez longue montre que les autres, en général distinctes, peuvent être réelles ou imaginaires:

Les directions situées dans les plans coordonnés sont réelles lorsque le binome B_i+2a_i correspondant est extérieur à l'intervalle A_j, A_k $(j, k\neq i)$; les conditions de réalité pour les autres ont été données ci-dessus, mais non développées; elles sont d'une écriture un peu longue†.

On est ainsi conduit au théorème suivant:

Théorème V.—Dans les milieux de Green à trois plans de symétrie de contexture, il existe, en général, treize directions de plans pour lesquels les trois ondes planes correspondantes sont, l'une longitudinale, les autres transversales. Les directions des normales à ces plans se répartissent ainsi qu'il suit:

*Auquel cas un (ou deux) des paramètres α, β, γ est nul et les quatre solutions déduites du système unique de valeurs de α^2, β^2. γ^2 ne sont plus toutes distinctes. Ces solutions rentrent alors dans la deuxième série.

†Ces conditions sont les suivantes: les trois fonctions

$$(B_j+2a_j)[A_j-(B_i+2a_i)]+(B_k+2a_k)[A_k-(B_i+2a_i)]-A_jA_k+(B_i+2a_i)^2,$$
$$(i, j, k=1, 2, 3;\ i\neq j\neq k),$$

doivent avoir le même signe.

1° *Quatre directions non situées dans les plans de symétrie mais symétriques par rapport à ceux-ci;*

2° *Dans chaque plan de symétrie, deux directions symétriques par rapport aux autres plans de symétrie;*

3° *Les trois directions des axes de symétrie.*

REMARQUES.—1. Trois de ces directions, celles des axes, sont toujours réelles et distinctes; les autres sont en général distinctes mais peuvent être réelles ou imaginaires de telle sorte qu'on ait en tout 3, 5, 7, 9, 11 ou 13 directions réelles.

2. Exceptionnellement, il peut y avoir une infinité de directions répondant à la question; ce fait se produira dans les deux cas suivants:

1° Si les équations (23) ne sont pas distinctes, ou tout au moins si, en désignant pas R, S, T, leurs 3 membres, les polynomes $R-S$ et $T-S$ ont un facteur commun du premier degré.

2° Si l'équation (24) est une identité.

1[er] CAS.—(*a*) Les équations (23) ne sont pas distinctes. On trouve sans aucune difficulté que les conditions nécessaires et suffisantes pour qu'il en soit ainsi peuvent s'écrire*:

$$\frac{B_2+2a_2-B_3-2a_3}{A_1-B_3-2a_3}=\frac{B_1+2a_1-A_2}{B_3+2a_3-A_2}=\frac{A_3-B_1-2a_1}{B_2-B_1+2a_2-2a_1}. \tag{25}$$

Les milieux doués d'isotropie transversale autour d'un axe doivent évidemment répondre à la question.

Ces milieux sont, au point de vue des réactions élastiques, caractérisés par les équations à cinq constantes élastiques: $\mathfrak{L}$, $\mathfrak{M}$, $\mathfrak{N}$, $\mathfrak{P}$ (l'axe d'isotropie est pris pour *oz* et deux directions rectangulaires quelconques du plan d'isotropie pour *ox* et *oy*)

$$\begin{cases} N_1=\mathfrak{L}\theta+2\mathfrak{M}\epsilon_1+2\mathfrak{N}\epsilon_3, & T_1=a\gamma_1,\\ N_2=\mathfrak{L}\theta+2\mathfrak{M}\epsilon_2+2\mathfrak{N}\epsilon_3, & T_2=a\gamma_2,\\ N_3=(\mathfrak{L}+2\mathfrak{N})\theta+2\mathfrak{P}\epsilon_3, & T_3=\mathfrak{M}\gamma_3. \end{cases} \tag{26}$$

On vérifierait que, pour ces milieux, les équations (25) sont satisfaites mais, bien plus simplement, on voit directement que les équations (23) se réduisent à une seule,

$$(\alpha^2+\beta^2)(a-\mathfrak{M}+\mathfrak{N})=(a-\mathfrak{P})\gamma^2. \tag{27}$$

Si $a=\mathfrak{P}=\mathfrak{M}-\mathfrak{N}$, cette équation s'évanouit.

Pour ces milieux spéciaux, à tout plan d'onde correspond une vibration longitudinale et deux transversales.

Les milieux isotropes, en particulier, rentrent dans ce cas.

*Sous réserve des modifications faciles à introduire si l'un des dénominateurs est nul.

(*b*) Les binomes $R-S$ et $T-S$ ont un facteur commun du premier degré. Etant donné leur forme ils en ont nécessairement un second et cette solution ne diffère pas de la précédente.

2[e] CAS.—L'équation (24) se réduit à une identité. Ceci exige:

$$A_2=A_3=B_1+2a_1. \tag{28}$$

On trouverait évidemment, par permutation circulaire, deux autres séries de solutions. Il est à remarquer que, si deux de ces conditions sont vérifiées à la fois, les trois coefficients A_i, sont égaux. Si, en outre, les trois conditions déduites de (28) par permutation circulaire sont vérifiées, les trois binomes B_i+2a_i sont égaux entre eux et à la valeur commune A des A_i.

Mais nous verrons ci-après (N° X) que, dans ce cas, la propriété recherchée ici est réalisée pour toute direction du plan d'onde.

La double condition (28) *est vérifiée, comme il fallait s'y attendre, pour le plan d'isotropie d'un milieu à isotropie transversale.*

Voyons pour quels milieux à isotropie transversale cette condition sera vérifiée pour les plans méridiens: il suffira évidemment qu'elle le soit pour l'un d'eux. On trouve sans peine que les relations (28) correspondant à l'un des plans xoz ou xoy conduisent aux conditions:

$$\mathfrak{P}=a,\ \mathfrak{M}=\mathfrak{N}+a. \tag{29}$$

Il est évident que, pour un tel milieu, à tout plan d'onde correspondent une vibration longitudinale et deux transversales. Ces conditions sont réalisées par les milieux isotropes. Mais ce ne sont pas les seuls. On se trouve en présence d'un cas particulier du problème suivant. (Voir N° X).

Dans ces milieux les N et les T sont liés aux déformations par les relations:

$$\begin{cases} N_1=\mathfrak{L}\theta+2\mathfrak{M}\epsilon_1+2\mathfrak{N}\epsilon_3, & T_1=(\mathfrak{M}-\mathfrak{N})\gamma_1, \\ N_2=\mathfrak{L}\theta+2\mathfrak{M}\epsilon_2+2\mathfrak{N}\epsilon_3, & T_2=(\mathfrak{M}-\mathfrak{N})\gamma_2, \\ N_3=(\mathfrak{L}+2\mathfrak{N})\theta+2(\mathfrak{M}-\mathfrak{N})\epsilon_3, & T_3=\mathfrak{M}\gamma_3. \end{cases} \tag{30}$$

X.

Milieux de Green pour lesquels deux vibrations sont transversales et la troisième longitudinale.

Les conditions à remplir par les coefficients d'élasticité ont été trouvées en 1839 par Green lui-même. Il a montré, en outre, que, dans ce cas, la surface d'onde se décomposait en deux nappes: une sphère correspondant à la vibration longitudinale et la surface d'onde de Fresnel, correspondant aux vibrations transversales.

Ces milieux sont, au point de vue élastique, définis par les relations:

$$(31)\qquad \begin{cases} N_i = A\theta - 2(b_i\gamma_i + a_j\epsilon_k + a_k\epsilon_j), \\ T_i = -2b_i\epsilon_i + a_i\gamma_i + b_j\gamma_k + b_k\gamma_j, \\ \quad (i=1,\,2,\,3;\; i\neq j\neq k). \end{cases}$$

Ces milieux, comme on le voit, ne dépendent que de *sept* coefficients élastiques, A, a_1, a_2, a_3, b_1, b_2, b_3.

Si, en outre, le milieu admet trois plans de symétrie de constitution, les b sont nuls et l'on a:

$$(31^{\text{bis}})\qquad \begin{cases} N_1 = A\theta - 2a_3\epsilon_2 - 2a_2\epsilon_3,\; N_2 = A\theta - 2a_3\epsilon_1 - 2a_1\epsilon_3,\; N_3 = A\theta - 2a_2\epsilon_1 - 2a_1\epsilon_2, \\ T_1 = a_1\gamma_1,\; T_2 = a_2\gamma_2,\; T_3 = a_3\gamma_2. \end{cases}$$

Si l'on se reporte aux notations du N° II (*in fine*) relatives aux milieu en question, on voit qu'ici on aura:

$$B_1 + 2a_1 = B_2 + 2a_2 = B_3 + 2a_3 = A_1 = A_2 = A_3;$$

ce qui justifie une remarque faite ci-dessus (N° 9, 2^e^ cas 3^e^ alinéa).

XI.

Nous avons vu que, d'une facon générale, un milieu anisotrope n'était susceptible de propager, pour une direction d'onde donnée, que des vibrations polarisées suivant certaines directions. Il y a exception, en se limitant au milieux de Green, si la quadrique de polarisation est de révolution. Nous nous proposons, dans ce paragraphe, de rechercher toutes les directions de plans pour lesquelles la quadrique de polarisation est de révolution ou, au point de vue physique, *de rechercher toutes les directions de plans d'onde pour lesquelles, dans un milieu de Green, l'une au moins des vibrations ne soit pas polarisée.* Il résulte d'ailleurs de ce qui suit que dans ce cas, une seule des trois vibrations composantes reste polarisée*.

DÉFINITION.—*Nous appellerons axe de polarisation la parallèle menée par le centre de la quadrique de polarisation à la normale à une direction de plan d'onde telle que l'une au plus des vibrations correspondantes soit polarisée.*

Dans le cas *d'un milieu de Green quelconque,* on trouve *en général trente-six axes de polarisation*; mais nous étudierons d'une manière plus complète le cas du milieu à trois plans rectangulaires de symétrie de constitution. Ceux-ci étant pris comme plans de référence, la quadrique de polarisation a pour équation:

$$(32)\qquad \begin{aligned} &(A_1\alpha^2 + a_3\beta^2 + a_2\gamma^2)x^2 + (a_3\alpha^2 + A_2\beta^2 + a_1\gamma^2)y^2 + (a_2\alpha^2 + a_1\beta^2 + A_3\gamma^2)z^2 \\ &\quad + 2(B_3 + a_3)\alpha\beta xy + 2(B_2 + a_2)\gamma\alpha xz + 2(B_1 + a_1)\beta\gamma yz = 1. \end{aligned}$$

*Il peut même arriver que cette dernière, elle même ne le soit pas (Voir N° XIV).

Premier cas.—Milieux pour lesquels aucun des binomes B_i+a_i n'est nul.

1° Directions non contenues dans l'un des plans de symétrie ($\alpha\beta\gamma\neq 0$). Les conditions de révolution sont alors:

$$(33)\qquad (A_1\alpha^2+a_3\beta^2+a_2\gamma^2)-\frac{(B_2+a_2)(B_3+a_3)}{(B_1+a_1)}\alpha^2$$
$$=(a_3\alpha^2+A_2\beta^2+a_1\gamma^2)-\frac{(B_3+a_3)(B_1+a_1)}{B_2+a_2}\beta^2$$
$$=(a_2\alpha^2+a_1\beta^2+A_3\gamma^2)-\frac{(B_1+a_1)(B_2+a_2)}{(B_3+a_3)}\gamma^2,$$

qui définissent quatre directions, en général distinctes et symétriques par rapport aux plans de coordonnées. Elles peuvent être réelles ou imaginaires. Ces solutions donnent lieu à des remarques analogues à celles qui ont été faites à propos du système (25); nous ne nous y arrêterons pas*.

2° Directions situées dans les plans de symétrie: $\alpha\beta\gamma=0$. Soit, par exemple, $\gamma=0$. Supposons d'abord $\alpha\beta\neq 0$. Les coefficients de deux des termes rectangles de la quadrique Σ sont alors nuls et les conditions complémentaires de révolution sont, compte tenu de $\gamma=0$:

$$(34)\qquad [(A_1-a_2)\alpha^2+(a_3-a_1)\beta^2][(a_3-a_2)\alpha^2+(A_2-a_1)\beta^2]-(B_3+a_3)^2\alpha^2\beta^2=0,\ \gamma=0.$$

On trouve ainsi quatre directions réelles ou imaginaires, symétriques deux à deux par rapport aux plans coordonnées et qui peuvent être confondues soit hors de ces plans, soit dans ces plans; dans ce dernier cas elles coïncident avec un axe de symétrie et rentrent dans le cas examiné ci-après. Revenant au cas général et appliquant aux trois plans de coordonnées on trouve ainsi douze directions d'axes de polarisation.

Examinons si un axe peut répondre à la question; soit par exemple *ox* ($\beta=\gamma=0$). Les trois termes rectangles de Σ sont alors nuls. Il faut une condition supplémentaire, savoir que deux des coefficients des termes carrés soient égaux.

On trouve ainsi, α^2 étant alors nécessairement différent de zéro, l'une des trois conditions suivantes:

$$(35)\qquad A_1=a_3,\quad A_1=a_2,\quad a_2=a_3.$$

(Si l'on a à la fois $A=a_2=a_3$, Σ se réduit à une sphère pour le plan de symétrie $x=0$, et aucune des 3 vibrations n'est polarisée).

*Toutefois, en cas de solution double ($\alpha\beta\gamma=0$) le mode même d'établissement des équations (33) est en défaut et il convient de voir ce qui se passe dans le cas où la résolution du système (33) conduirait à une valeur nulle pour l'une des inconnues α^2, β^2 ou γ^2. Ces solutions ne peuvent convenir que si elles rentrent dans la 2^e série. On conçoit, comme cas limite qu'il doit bien en être ainsi. C'est ce que montre d'ailleure la vérification directe. La condition à réaliser est (avec $\gamma=0$):

$$\left[a_2-A_1+\frac{(B_2+a_2)(B_3+a_3)}{B_1+a_1}\right]\left[a_1-A_2+\frac{(B_1+a_1)(B_3+a_3)}{B_2+a_2}\right]=(a_3-a_1)(a_3-a_2).$$

On voit donc que, hors le cas des milieux spéciaux définis par l'une des conditions (35) (ou de celles qui s'en déduisent par permutation) aucun axe de symétrie n'est axe de polarisation. Dans le cas de ces milieux spéciaux, et alors seulement, les solutions dérivées de l'équation (34) peuvent donc être une direction d'axe de symétrie.

En définitive, *seize axes de polarisation pour les milieux généraux* à trois plans rectangulaires de symétrie de constitution.

Deuxième cas.—Milieux pour lesquels l'un des binomes B_i+a_i est nul.

1° Un seul des binomes est nul; soit $B_3+a_3=0$. Le terme en xy de l'équation (32) est nul; on doit écrire que l'un des autres rectangles est nul; si on laisse de côté le cas des axes étudié ci-dessus (les calculs faits sont, dans ce cas, indépendants de la nullité ou de la non-nullité de l'un des B_i+a_i) on est conduit à l'une des deux hypothèses suivantes:

$$(a)\ \gamma=0; \qquad (b)\ \alpha \text{ ou } \beta=0.$$

(a) Si $\gamma=0$, les trois termes rectangles sont nuls; on est donc conduit à écrire l'égalité de deux des coefficients des termes carrés de l'équation (32), c'est-à-dire, compte tenu de $\gamma=0$, de deux des binomes:

$$A_1\alpha^2+a_3\beta^2, \quad a_3\alpha^2+A_2\beta^2, \quad a_2\alpha^2+a_1\beta^2.$$

Chacune des équations ainsi obtenues donne deux directions du plan xoy symétriques deux à deux par rapport aux plans coordonnés. On trouve de la sorte *six solutions* réelles ou imaginaires, en général distinctes. Il est facile de former les conditions de réalité et celles pour lesquelles il y a des solutions confondues. Celles-ci se répartissent en deux groupes: solutions confondues avec ox ou oy, ou solution triple, en général formée de droites distinctes de ox et de oy.

(b) $\alpha=0$. Le terme en yz n'est pas nul, les autres rectangles le sont. On est conduit à écrire la dernière condition de révolution qui, compte tenu de $\alpha=0$, est:

$$(36) \qquad [(A_2-a_3)\beta^2+(a_1-a_2)\gamma^2][(a_1-a_3)\beta^2+(A_3-a_2)\gamma^2]-(B_1+a_1)^2\gamma^2\beta^2=0,$$

ce qui donne *quatre* solutions, en général différentes des précédentes (lorsqu'elles ne se réduisent pas à un axe, restriction que nous sous-entendrons désormais dans cette question).

L'hypothèse $\beta=0$ conduit à quatre solutions nouvelles.

En définitive, *lorsque l'un des B_i+a_i est nul sans qu'un second le soit, le milieu admet en général quatorze axes de polarisation.*

2° Deux binomes sont nuls; soit:

$$B_1+a_1=B_2+a_2=0,$$

le troisième

$$B_3+a_3\neq 0.$$

(a) Si on ne suppose pas nul le produit $\alpha\beta$, le terme rectangle en xy et lui seul est différent de zéro, d'où la condition

$$(37)\quad [(A_1-a_2)\alpha^2+(a_3-a_1)\beta^2+(a_2-A_3)\gamma^2]\ [(a_3-a_2)\alpha^2+(A_2-a_1)\beta^2+(a_1-A_3)\gamma^2] = (B_3+a_3)^2\alpha^2\beta^2.$$

On trouve ainsi un cône de directions d'axes du quatrième degré*.

(b) Si on suppose nul le produit $\alpha\beta$, les trois termes rectangles de l'équation (32) sont nuls. On serait, comme ci-dessus (1°, a), conduit à six nouvelles directions pour chacune des hypothèses $\beta=0$ ou $\alpha=0$, mais quatre d'entre elles seraient sur le cône (37) et par suite ne seraient pas à retenir. En définitive, outre le cône d'équation (37) on trouve quatre nouvelles directions définies par les relations

$$(38)\quad \begin{cases} \alpha=0 \text{ ou } \beta=0, \\ \text{et} \\ (A_1-a_3)\alpha^2-(A_2-a_3)\beta^2+(a_2-a_1)\gamma^2=0. \end{cases}$$

3° Les trois binomes B_i+a_i sont nuls; les trois termes rectangles sont alors nuls et il existe un lieu de directions d'axes formé de trois cônes du second degré dont on obtient les équations en égalant deux à deux les coefficients des termes en x^2, y^2 ou z^2 de l'équation (37).

Ces résultats et ceux qui seront établis ci-après (paragraphes XII et XIII) peuvent être résumés dans le théorème suivant:

*Il serait aisé d'en discuter les variétés en observant que la substitution: $u=\alpha^2$, $v=\beta^2$, $w=\gamma^2$, en ramène l'étude à celle d'un cône du second ordre.

Dans le cas des milieux doués d'isotropie transversale, ce cône doit se décomposer en deux cônes de révolution.

Pour ces milieux, la condition $B_1+a_1=B_2+a_2=0$ se réduit à $\mathfrak{L}+2\mathfrak{M}+a=0$. On vérifie que ce cône se décompose, en effet, en deux cônes de révolution:

$$(37^{\text{bis}})\qquad \frac{(a-\mathfrak{P})\gamma^2}{\alpha^2+\beta^2}=-(\mathfrak{L}+2\mathfrak{M}-a),$$

et

$$(37^{\text{ter}})\qquad \frac{(a-\mathfrak{P})\gamma^2}{\alpha^2+\beta^2}=a-\mathfrak{M}.$$

Ces cônes sont distincts, sauf lorsque $\mathfrak{L}+\mathfrak{M}=0$, condition qui, jointe à $\mathfrak{L}+2\mathfrak{M}+a=0$ définit des milieux très particuliers sur lesquels nous reviendrons [Note du 1[er] renvoi de la Remarque de la Section XII (2[e] cas § 6) et alinéa relatif à l'équation (48)].

Ils sont d'ailleurs, en général, distincts du cône d'équation (43). L'équation de ce dernier, pour les milieux dont il est question ici, se réduit à:

$$(43^{\text{bis}})\qquad (\mathfrak{M}-a)\ (\alpha^2+\beta^2)=2\gamma^2\ (\mathfrak{P}-a).$$

Si on laisse de côté le cas où $\mathfrak{P}-a=0$, pour lequel la seule direction réelle de ces trois cônes, réduits à des plans imaginaires (lorsque leur équation ne se réduit pas à une identité—voir Section XIII) est $\alpha=\beta=0$, on est conduit,

pour la coïncidence de 37^{bis} et de 43^{bis} à la condition: $2\mathfrak{L}+3\mathfrak{M}=a$,

" " 37^{ter} " 43^{ter} " : $\mathfrak{M}=a$.

(Dans ce dernier cas, d'ailleurs, le cône se réduit au plan d'isotropie $\gamma=0$.)

Théorème VI.—Les milieux de Green les plus généraux ont trente-six axes de polarisation.

Dans les milieux possédant trois plans rectangulaires de symétrie, ce nombre tombe à seize dans le cas le plus général. Il peut s'abaisser mais non s'élever à moins d'être remplacé par une infinité simple. Il peut même arriver, dans certains milieux très spéciaux, que toute direction de l'espace soit un axe de polarisation. Dans le cas des milieux généraux à trois plans de symétrie, les directions d'axe de polarisation se répartissent comme il suit: quatre directions non parallèles à ces plans mais deux à deux symétriques par rapport à eux; douze directions parallèles aux plans de symétrie.

XII.

Recherche des milieux de Green a trois plans de symétrie possédant une infinité d'axes de polarisation.

Nous avons vu ci-dessus, deux exemples de milieux possédant une infinité d'axes de polarisation. Nous nous proposons, maintenant, de rechercher tous les cas dans lesquels ce fait se produira.

Premier cas.—Aucun des binomes B_i+a_i n'est nul. On est alors conduit à écrire soit:

1° que les deux équations (33) se réduisent à une seule, ou tout au moins que, si l'on désigne par R, S, T, leurs trois membres, les différences $R-S$ et $T-S$ ont un facteur du premier degré commun. Cette dernière solution, eu égard à la forme des polynomes R, S et T, rentre d'ailleurs dans la précédente. Cela étant, on exprimera la proportionalité des coefficients des polynomes $R-S$ et $T-S$ et l'on trouvera des conditions analogues aux conditions (25)*.

2° que la première équation (34) soit une identité. Ceci donne les trois conditions:

$$(39) \qquad \begin{cases}(a_2-A_1)(a_2-a_3)=0, \quad (a_1-a_3)(a_1-A_2)=0,\\ 2a_1a_2-a_3(a_1+a_2)+A_1A_2-a_1A_1-a_2A_2=(B_3)^2+2a_3B_3.\end{cases}$$

Les premières équations (39) montrent qu'il y a lieu de considérer trois séries de solutions:

1e série $\qquad a_1=a_2=a_3,$

2e série $\qquad a_2=a_3, \quad a_1=A_2$ (ou $a_1=a_3$, $a_2=A_1$),

3e série $\qquad a_2=A_1, \quad a_1=A_2.$

Enfin, il faudrait, aux solutions trouvées, ajouter celles qui correspondent aux équations déduites de (34) par permutation circulaire.

*Ces conditions sont, en posant $C_i=\dfrac{(B_j+a_j)(B_k+a_k)}{B_i+a_i}$, $(i, j, k=1, 2, 3;\ i\neq j\neq k)$,

$$\frac{A_1-C_1-a_3}{A_1-C_1-a_2}=\frac{a_3-A_2+C_2}{a_3-a_1}=\frac{a_2-a_1}{a_2-A_3+C_3}.$$

1[e] *série.*—$a_1=a_2=a_3$. Désignons par μ leur valeur commune. Introduisant ces valeurs dans la troisième équation (39), celle-ci donne deux solutions définies par les conditions*:

(40) $$a_1=a_2=a_3=\mu,\ B_3=-\mu+\sqrt{\mu^2-\mu(A_1+A_2)+A_1A_2}\ ,$$

ou

(40^{bis}) $$a_1=a_2=a_3=\mu,\ B_3=-\mu-\sqrt{\mu^2-\mu(A_1+A_2)+A_1A_2}\ .$$

2[e] *série.*—Elle conduit à écrire $B_3+a_3=0$ et rentre dans un cas actuellement écarté.

3[e] *série.*—Deux types, comme pour la première série:

(41) $$a_2=A_1,\ a_1=A_2,\ B_3=-a_3+\sqrt{(a_3-A_1)(a_3-A_2)}\ ,$$

ou

(41^{bis}) $$a_2=A_1,\ a_1=A_2,\ B_3=-a_3-\sqrt{(a_3-A_1)(a_3-A_2)}\ .$$

Deuxième cas.—L'un des binomes B_i+a_i, au moins est nul.

(*a*) *Un seul binome est nul.*—Soit B_3+a_3.

On a vu (N° XI) qu'il y avait à examiner soit $\gamma=0$, soit α ou $\beta=0$.

1° $\gamma=0$; l'une des trois relations à joindre à cette condition doit être une identité. Les coefficients d'élasticité des milieux cherchés devront donc satisfaire à l'un des couples de relation:

$$a_3=A_1=A_2;\ a_2=A_1,\ a_3=a_1;\ a_1=A_2,\quad a_3=a_2$$

(ces deux derniers ne diffèrent pas au fond) en même temps qu'à la relation:

$$B_3+a_3=0.$$

2° $\alpha=0$ doit être jointe à la relation (36). Celle-ci doit donc se réduire à une identité; d'où:

(42) $$\begin{cases}(a_1-a_2)\ (A_3-a_2)=0,\\ (a_1-a_3)\ (A_2-a_3)=0,\\ (a_1-a_3)\ (a_1-a_2)+(A_2-a_3)\ (A_3-a_2)=(B_1+a_1)^2.\end{cases}$$

*Un type particulièrement intéressant est défini par:

$$A_1=A_2=A_3,$$
$$B_1=B_2=B_3.$$

Désignons par $-\lambda$ la valeur des A, celle des B sera:

$$B=-\mu\pm(\mu+\lambda),$$

c'est-à-dire

$$-(2\mu+\lambda)\ \text{ou}\ \lambda.$$

Le premier de ces milieux se confond avec les milieux isotropes; le second est défini, au point de vue élastique, par les formules:

$$N_i=\lambda\theta-2\lambda\epsilon_i,\quad T_i=\mu\gamma_i.$$

On pourrait évidemment trouver ainsi de nombreux milieux particuliers intéressants satisfaisant à la condition

$$A_1=A_2=A_3.$$

La considération des deux premières équations conduit aux quatre séries de solutions:

1° $a_1 = a_2 = a_3 (= a)$; $(B_1 + a_1)^2 = (A_2 - a)(A_3 - a)$,

et par suite, $B_3 = -a$.

Ces milieux dépendent donc des 5 paramètres A_1, A_2, A_3, B_2, a.

2° $a_1 = a_2$, $A_2 = a_3$. Cette solution conduit à $B_1 + a_1 = 0$. Elle rentre donc parmi les solutions correspondant à la nullité de deux binomes $B_i + a_i$, solution étudiée ci-après.

3° $a_1 = a_3$, $A_3 = a_2$; identique au fond à la précédente.

4° $A_3 = a_2$, $A_2 = a_3$, $(B_1 + a_1)^2 = (a_1 - a_3)(a_1 - a_2)$, $B_3 = -a_3$.

Ces milieux sont à *cinq paramètres*.

REMARQUE.—L'hypothèse $\beta = 0$ conduit à des résultats évidemment de même nature.

(*b*) *Deux ou trois binomes* $B_i + a_i$ *sont nuls*.

On a vu plus haut (N° XI) que, dans ce cas, il existait un cône d'axes de polarisation, ou même trois cônes algébriquement distincts dans le cas où les trois binomes sont nuls.

REMARQUE.—Il est évident a priori que les milieux à isotropie transversale doivent admettre un cône d'axes de polarisation dès qu'ils en admettent un différent de l'axe d'isotropie. Ce cône doit être formé d'un ou de plusieurs cônes de révolution autour de l'axe d'isotropie. Considérons d'abord ceux de ces milieux qui rentrent dans la première série. [Aucun $B_i + a_i$ n'est nul; c'est-à-dire ici: $\mathfrak{L} + \mathfrak{M} \neq 0$ et $\mathfrak{L} + 2\mathfrak{N} + a \neq 0$, voir pour le cas contraire ci-après note* ci-dessous, éq. 48 et renvoi de la Section XI (2^e^ cas).

Il est facile de voir que, dans ce cas, $R - S$ est identiquement nul. Le lieu est formé du cône d'équation $T - S = 0$.

Le calcul direct vérifie ces prévisions et donne l'équation du cône:

$$(\mathfrak{M} - a)\ (\alpha^2 + \beta^2) = \gamma^2 \left[\mathfrak{L} + 2\mathfrak{N} + 2\mathfrak{P} - a - \frac{(\mathfrak{L} + 2\mathfrak{N} + a)^2}{\mathfrak{L} + \mathfrak{M}} \right]. \tag{43}$$

Si l'on a:

$$\mathfrak{M} = a, \tag{44}$$

ce cône se réduit au plan d'isotropie, à moins que les coefficients vérifient en outre la relation:

$$\mathfrak{P} - \mathfrak{M} - \mathfrak{N} = \frac{2\mathfrak{N}^2}{\mathfrak{L} + \mathfrak{M}}, \tag{45}$$

auquel cas l'équation (43) devient une identité et *toute direction est un axe de polarisation*.

*Cette formule suppose $\mathfrak{L} + \mathfrak{M} \neq 0$; si $\mathfrak{L} + \mathfrak{M} = 0$, on se trouve dans le cas où $B_3 + a_3 = 0$. Si on laisse de côté le cas de $\alpha\beta\gamma = 0$, complètement étudié plus loin dans le texte, on est conduit à écrire que l'un des binomes $B_1 + a_1$ ou $B_2 + a_2$ est nul. Ils sont d'ailleurs égaux tous deux ici à $\mathfrak{L} + 2\mathfrak{N} + a$. On obtient donc la condition complémentaire $\mathfrak{L} + 2\mathfrak{N} + a = 0$. On trouve une solution qui sera examinée ci-après et qui répond d'ailleurs, comme on le verra, à la question.

Ce fait se présente pour les corps isotropes, caractérisés par $\mathfrak{P} = \mathfrak{M}$, $\mathfrak{N} = 0$; il fallait s'y attendre car on sait que, dans les corps isotropes, la vibration longitudinale seule est polarisée. Pour n'omettre aucune solution, il faut examiner ce que donnerait le cas $\alpha\beta\gamma = 0$. Il y a alors un nombre fini (cas général) ou un *plan de directions* d'axes de polarisation; ce dernier cas se présente quand la première équation (34) (ou celles qui s'en déduisent par permutation) est identiquement satisfaite.

S'il existe une infinité de telles directions dans le plan $\alpha = 0$, ou $\beta = 0$, comme tout l'ensemble est de révolution autour de *oz*, toutes les directions de l'espace doivent répondre à la question. L'ensemble des équations (44) et (45) définit un cas où il en est ainsi. Mais comme son établissement est fondé sur les équations (33) qui supposent $\alpha\beta\gamma \neq 0$ on peut avoir un doute sur la validité du résultat, au moins en ce qui concerne le plan d'isotropie. Il est évident que rien ne distinguant les plans $\alpha = 0$, $\beta = 0$ d'un plan quelconque passant par l'axe d'isotropie, le résultat général leur reste applicable, au moins autant que l'équation (43) n'est pas en défaut (Voir note du 3^e alinéa de la présente remarque). Si l'on veut examiner ce cas, on est conduit à étudier séparément l'hypothèse où le produit $\alpha\beta\gamma$ est nul. Par raison de symétrie, il suffira d'étudier les seules hypothèses $\alpha = 0$ ou $\gamma = 0$.

A la condition $\alpha = 0$ doit être associée la suivante, qui n'est autre chose que ce que devient l'équation (34) pour les milieux à isotropie transversale et lorsqu'on permute γ avec α:

$$(46) \qquad \beta^2\{(\mathfrak{L}+\mathfrak{M})(a-\mathfrak{M})\beta^2+[(\mathfrak{L}+\mathfrak{M})(\mathfrak{L}+2\mathfrak{N}+2\mathfrak{P}-a)-(\mathfrak{L}+2\mathfrak{N}+a)]^2\gamma^2\} = 0.$$

Cette équation doit se réduire à une identité, d'où les deux conditions:

$$(47) \qquad \begin{cases} (\mathfrak{L}+\mathfrak{M})\ (a-\mathfrak{M}) = 0, \\ (\mathfrak{L}+\mathfrak{M})\ (\mathfrak{L}+2\mathfrak{N}+2\mathfrak{P}-a)-(\mathfrak{L}+2\mathfrak{N}+a)^2 = 0. \end{cases}$$

On trouve d'abord la solution, déjà signalée:

$$(48) \qquad \mathfrak{L}+\mathfrak{M} = 0, \quad \mathfrak{L}+2\mathfrak{N}+a = 0.$$

Cette solution annule les trois binomes $B_i + a_i$ relatifs à ce milieu et l'équation de la quadrique de polarisation correspondante est:

$$(49) \qquad [-\mathfrak{L}(\alpha^2+\beta^2)+a\gamma^2]\ (x^2+y^2)+[a(\alpha^2+\beta^2)+(2\mathfrak{P}-a)\gamma^2]z^2 = 1,$$

équation d'une quadrique de révolution autour de *oz* quels que soient α, β et γ. Le milieu défini par les équations (48) admet donc toutes les directions de l'espace pour axes de polarisation*.

*Si l'on a:

$$-\mathfrak{L}(\alpha^2+\beta^2)+a\gamma^2 = a(\alpha^2+\beta^2)+(2\mathfrak{P}-a)\gamma^2$$

cette quadrique est une sphère et aucune des vibrations n'est polarisée. Ceci arrive, sans condition nouvelle pour les directions du cône de révolution représenté par l'équation ci-dessus ou pour toutes les directions de l'espace, si $a = -\mathfrak{L} = \mathfrak{P}$, mais ce dernier milieu n'est autre que le milieu isotrope très spécial pour lequel $\lambda + \mu = 0$.

Une seconde solution s'obtient en écrivant:

$$\mathfrak{M} = a,$$

et portant cette valeur dans la seconde équation (47). On retrouve les milieux définis par les conditions (44) et (45).

Reste à voir ce que donne $\gamma = 0$. Il est visible que, s'il existe dans ce plan une seule direction d'axe de polarisation ne passant pas en un point cyclique toutes ses directions répondent à la question.

Un calcul direct montre que dans les milieux à isotropie transversale, la première équation (34) devient:

$$(\mathfrak{L} + 2\mathfrak{M} - a)(\mathfrak{M} - a)(\alpha^2 + \beta^2)^2 = 0, \tag{50}$$

ce qui vérifie les prévisions. On voit, en outre, que les conditions pour que toutes les directions du plan d'isotropie soient des axes de polarisation sont:

$$\mathfrak{M} = a,$$

résultat déjà rencontré, ou

$$\mathfrak{L} + 2\mathfrak{M} - a = 0. \tag{51}$$

XIII.

Recherche des milieux de Green a trois plans rectangulaires de symétrie de contexture pour lesquels toute direction de l'espace est un axe de polarisation.

Nous avons rencontré ci-dessus deux milieux à isotropie transversale pour lesquels toute direction de l'espace était celle d'un axe de polarisation. Pour les avoir toutes il faut encore rechercher si l'équation de l'un des deux cônes (37^{bis}) ou (37^{ter}) ne peut se réduire à une identité. Pour le cône (37^{ter}) on trouve la condition:

$$\mathfrak{M} = \mathfrak{P} = a, \tag{52}$$

et pour le cône (37^{bis})

$$\mathfrak{L} + 2\mathfrak{M} = \mathfrak{P} = a, \tag{53}$$

conditions auxquelles il faut joindre la suivante:

$$\mathfrak{L} + 2\mathfrak{M} + a = 0. \tag{54}$$

Les milieux ainsi définis constituent deux solutions différentes des précédentes.

Nous nous proposons, dans ce paragraphe, d'étendre cette recherche aux milieux de Green généraux à trois plans rectangulaires de symétrie de constitution.

On peut éliminer de suite toute recherche supposant $\alpha\beta\gamma=0$, une telle spécification imposant aux directions le parallélisme à des plans bien déterminés*.

Cela posé, nous distinguerons toujours deux cas suivant qu'aucun des B_i+a_i n'est nul ou que l'un d'eux au moins est nul.

Premier cas. Aucun des B_i+a_i n'est nul.

Il faut alors écrire l'identité des trois membres des équations (33). Cela ne présente aucune difficulté et conduit à imposer aux a_i une même valeur que nous désignerons par μ et, remplacant, par analogie avec les notations de Lamé, B_i par λ_i, nous trouverons:

$$A_i=\frac{(\lambda_j+\mu)\ (\lambda_k+\mu)}{\lambda_i+\mu}+\mu, \qquad (i,\ j,\ k=1,\ 2,\ 3;\ i\neq j\neq k). \tag{55}$$

Le milieu correspondant, très intéressant, est défini, au point de vue élastique par les relations:

$$\left\{\begin{aligned} N_i&=\left[\frac{(\lambda_j+\mu)\ (\lambda_k+\mu)}{\lambda_i+\mu}+\mu\right]\epsilon_i+\lambda_j\epsilon_k+\lambda_k\epsilon_j,\\ T_i&=\mu\gamma_i.\end{aligned}\right. \tag{56}$$

Si $\lambda_1=\lambda_2=\lambda_3(=\lambda)$ on retrouve les milieux isotropes et les formules bien connues de Lamé.

Deuxième cas. L'un des B_i+a_i est nul.

Aucun des α, β, γ ne l'étant, il faut alors que l'un au moins des deux autres binomes $B+a$ le soit.

1° Deux binomes sont nuls; le troisième ne l'est pas.

Soit: $$B_1+a_1=B_2+a_2=0.$$

Il y a alors, en général, un cône du 4^e ordre, d'après l'équation (37), de directions d'axes de polarisation. Il faut donc et il suffit que son équation se

*Ceci, toutefois, peut prêter à objection, au moins dans le cas où les conditions de révolution utilisées supposent les termes rectangles différents de zéro. Si l'un, et par suite deux d'entre eux sont nuls, les conditions obtenues sont suffisantes, que le troisième soit nul ou non, et par suite, pas de difficulté. Nous pouvons donc nous demander si les conditions générales de révolution obtenues dans le premier cas étant en défaut pour $\alpha\beta\gamma=0$, les directions des droites des trois plans de symétrie ne feraient pas exception au résultat général, c'est-à-dire n'échapperaient pas à la propriété de toutes les directions de l'espace d'être des axes de polarisation. Le moyen le plus simple de lever la difficulté est de voir ce que devient, pour ces directions spéciales et pour les milieux en question, la quadrique de polarisation.

Il suffit, d'après ce qu'on vient de dire, d'examiner ce qui se passe pour les milieux définis par les formules (55) ou (56) et pour l'un, au choix, des plans de symétrie, par exemple $\gamma=0$, les trois jouant le même rôle dans les formules. L'équation de la quadrique de polarisation correspondante est:

$$\left\{\left[\frac{(\lambda_2+\mu)\ (\lambda_3+\mu)}{\lambda_1+\mu}+\mu\right]\alpha^2+\mu\beta^2\right\}x^2+\left\{\mu\alpha^2+\left[\frac{(\lambda_1+\mu)\ (\lambda_3+\mu)}{\lambda_2+\mu}+\mu\right]\beta^2\right\}y^2+\mu(\alpha^2+\beta^2)z^2$$
$$+2(\lambda_3+\mu)\alpha\beta xy=1,$$

et on vérifie sans peine que l'unique condition de révolution est une identité, ce qui lève toute difficulté.

réduise à une identité pour que toute direction de l'espace soit un axe de polarisation. On remarquera que si, dans cette équation, on fait $\gamma=0$, l'équation résiduelle doit être une identité. Les conditions pour lesquelles il en est ainsi ont été trouvées. Ce sont les conditions (39). C'est donc parmi les milieux correspondant que doivent être cherchés les solutions du problème que nous avons en vue.

1e *série:* $a_1=a_2=a_3$. Désignons par μ leur valeur commune:

$$B_3=-\mu\pm\sqrt{(\mu-A_1)(\mu-A_2)}\ .$$

L'équation (37) devient:

$$\gamma^2(\mu-A_3)[(A_1-\mu)\alpha^2+(A_2-\mu)\beta^2-(A_3-\mu)\gamma^2]=0. \tag{57}$$

La condition $A_3=\mu$ est nécessaire et suffisante pour que l'équation (57) se réduise à une identité. On obtient ainsi les deux types de solutions:

$$(58_1)\qquad A_3=a_1=a_2=a_3=\mu,\quad B_1=B_2=-\mu,\quad B_3=-\mu+\sqrt{(\mu-A_1)(\mu-A_2)}\ ,$$

et

$$(58_2)\qquad A_3=a_1=a_2=a_3=\mu,\quad B_1=B_2=-\mu,\quad B_3=-\mu-\sqrt{(\mu-A_1)(a_3-A_2)}\ .$$

2e *série:* $a_2=a_3=a$, $a_1=A_2$. On est conduit à écrire $B_3+a_3=0$. La solution rentre dans celles qui seront examinées ci-après (les trois B_i+a_i nuls).

3e *série:* $a_2=A_1$, $a_1=A_2$, $B_3=-a_3\pm\sqrt{(a_3-A_1)(a_3-A_2)}$.

En poursuivant les calculs, on trouve d'une part les milieux à isotropie transversale dont les coefficients sont liés par les relations: $\mathfrak{L}+2\mathfrak{M}=a$, $\mathfrak{N}+a=\mathfrak{M}$, $P=a$ et, d'autre part, des milieux définis, au point de vue élastique, par les formules à deux coefficients:

$$N_i=-a\theta+2a\epsilon_i (i=1,\ 2,\ 3),\quad T_1=a\gamma_1,\quad T_2=a\gamma_2,\quad T_3=\mu\gamma_3\ (a,\ \mu \text{ constantes}). \tag{59}$$

2° Les trois binomes B_i+a_i sont nuls. On trouve les milieux à isotropie transversale spéciaux caractérisés par les relations: $\mathfrak{L}+\mathfrak{M}=0$, $\mathfrak{L}+2\mathfrak{N}+a=0$, déjà rencontrées (Eq. 48).

XIV.

Recherche des milieux de Green a trois plans de symétrie rectangulaires pour lesquels il existe des ondes dont aucune vibration n'est polarisée.

Dans les recherches précédentes, on s'est posé le problème de trouver les conditions pour que les vibrations se comportent, au point de vue de la polarisation, comme dans les milieux isotropes généraux, c'est-à-dire, tels qu'une seule des trois vibrations soit polarisée. D'où la notion des axes de polarisation et les études qui s'y sont rattachées. Mais, de même qu'il existe un milieu isotrope (défini par $\lambda+\mu=0$) dans lequel aucune vibration n'est polarisée, nous allons chercher tous les milieux de Green à trois plans de symétrie tels qu'il existe des ondes dont aucune vibration ne soit polarisée. Il faut et il suffit pour cela que la quadrique de polarisation se réduise à une sphère. Nous dirons que les axes de polarisation correspondants sont *des axes spéciaux* de polarisation. Les résultats sont les suivants:

Si aucun des binomes B_i+a_i n'est nul, les seuls axes spéciaux possibles sont les axes de symétrie. Pour que la propriété soit réalisée, il faut et il suffit que les constantes du milieu satisfassent aux conditions:

(60) $$a_i=a_j=A_k \ (i\neq j\neq k),$$

l'indice k correspondant au plan de symétrie intéressé.

Si l'un des binomes B_i+a_i est nul, mais non deux, on trouve que, outre les solutions précédentes, il y aura, dans le cas le plus général, sous réserve d'une seule condition complémentaire entre les coefficients, quatre directions d'axes spéciaux. Pour $B_3+a_3=0$, par exemple, ces directions sont définies par les relations:

(61) $$(A_1-a_2)\alpha^2=(a_1-a_3)\beta^2, \quad (A_2-a_1)\beta^2=(a_2-a_3)\alpha^2, \quad \gamma=0,$$

avec la relation de compatibilité:

(62) $$(A_1-a_2)(A_2-a_1)=(a_3-a_1)(a_2-a_1), \quad (\text{et } B_3+a_3=0).$$

Mais *toutes* les directions du plan *xoy* sont des directions d'axes spéciaux pour les milieux plus particuliers dont les coefficients satisfont aux conditions:

(63) $$a_1=a_2=a_3=A_1=A_2, \quad (\text{avec } B_3+a_3=0).$$

Designons par a les valeurs communes des A et des a; ces milieux sont donc à quatre paramètres (B_1, B_2, A_3 et a).

A ces solutions il faut ajouter, bien entendu, celles qui s'en déduisent par permutation circulaire.

L'hypothèse où deux binomes B_i+a_i mais non les trois, sont nuls, ne conduit à aucune solution nouvelle.

Enfin, si les trois binomes B_i+a_i sont nuls, il y aura, en général, sans autre condition, quatre axes spéciaux de polarisation définis par les relations:

(64) $$A_1\alpha^2+a_3\beta^2+a_2\gamma^2=a_3\alpha^2+A_2\beta^2+a_1\gamma^2=a_2\alpha^2+a_1\beta^2+A_3\gamma^2.$$

Il y en aura une infinité, sur un cône du second degré si les coefficients satisfont aux conditions:

(65) $$\begin{cases} \dfrac{A_1-a_3}{a_2-a_3}=\dfrac{a_3-A_2}{a_1-A_2}=\dfrac{a_2-a_1}{A_3-a_1}, \\ B_i+a_i=0, \quad (i=1,\ 2,\ 3), \end{cases}$$

les premières relations (65) pouvant facilement être remplacées par les relations convenables lorsque l'un des dénominateurs est nul.

Enfin, il n'existe, en dehors du milieu isotrope spécial pour lequel $\lambda+\mu=0$, aucun milieu tel que, pour toute direction d'onde, aucune des trois vibrations ne soit polarisée.

XV.

On a recherché, ci-dessus, surtout pour les milieux de Green et, plus particulièrement, pour ceux d'entre eux qui possèdent trois plans rectangulaires de symétrie de contexture, les directions d'ondes telles que:

1° L'une des ondes soit longitudinale,

2° Une seule des trois ondes soit polarisée.

Ces deux propriétés sont possédées par toute onde se propageant en milieu isotrope. Nous donnons ci-dessous un tableau récapitulatif des conditions définissant un milieu pour lequel il existe au moins une infinité simple de directions jouissant de l'une de ces propriétés. Aux milieux définis par les relations données dans ce tableau, il faut ajouter ceux qui s'en déduisent par permutation circulaire.

Tableau des conditions définissant les milieux de Green spéciaux à trois plans rectangulaires de symétrie de constitution.

I. MILIEUX POSSÉDANT UNE INFINITÉ DE DIRECTIONS D'ONDES A UNE VIBRATION LONGITUDINALE.

1° *Milieux pour lesquels il existe une infinité simple de ces directions.*

Ces milieux sont definis par l'un des groupes de conditions suivants:

$$\text{(I)}\qquad \frac{B_2+2a_2-B_3-2a_3}{A_1-B_3-2a_3}=\frac{B_1+2a_1-A_2}{B_3+2a_3-A_2}=\frac{A_3-B_1-2a_1}{B_2-B_1+2a_2-2a_1}.$$

Équation du cône des normales aux plans d'ondes à vibration longitudinale:

$$[B_2+2a_2-B_3-2a_3]\alpha^2+[B_1+2a_1-A_2]\beta^2+[A_3-B_1-2a_1]\gamma^2=0.$$

REMARQUE*: Si l'un des dénominateurs des rapports (I) était nul, il conviendrait de changer la forme des conditions.

Si l'un des dénominateurs est nul les équations à prendre sont les suivantes que nous écrirons pour le cas où $B_2-B_1+2a_2-2a_1=0$:

$$(\text{I}^{\text{bis}})\qquad \begin{cases} B_2-B_1+2a_2-2a_1=0, \\ A_3-B_1-2a_1=0, \end{cases}$$

$$\frac{B_2+2a_2-B_3-2a_3}{A_1-B_3-2a_3}=\frac{B_1+2a_1-A_2}{B_3+2a_3-A_2}.$$

Le cône des normales se réduit aux deux plans:

$$(B_1+2a_1-B_3-2a_3)\alpha^2+(B_1+2a_1-A_2)\beta^2=0.$$

Si deux dénominateurs et non trois sont nuls, on écrira que les numérateurs correspondants le sont. Le cône se réduit à un plan double.

Si les trois dénominateurs sont nuls il n'y a pas d'autre condition; l'équation du cône a la même forme que dans le cas géneral et c'est, sauf exception, un cône véritable.

(II) $B_1+2a_1=A_2=A_3$. Équation du cône des normales: $\alpha=0$.

2° *Milieux pour lesquels toutes les directions de l'espace jouissent de cette propriété.*

$$(A)\qquad A_1=A_2=A_3=B_1+2a_1=B_2+2a_2=B_3+2a_3.$$

*Cette remarque s'applique au cas analogue rencontré ci-après; on ne la présentera plus.

II. Milieux possédant une infinité de directions d'axes de polarisation.

1° *Milieux possédant une infinité simple d'axes de polarisation.*

Posons:

$$C_i=\frac{(B_j+a_j)(B_k+a_k)}{B_i+a_i}, \quad (i,\ j,\ k=1,\ 2,\ 3;\ i\neq j\neq k).$$

Ces milieux sont définis par l'un des groupes de relations suivants:

(I′) $$\frac{A_1-C_1-a_3}{A_1-C_1-a_2}=\frac{a_3-A_2+C_2}{a_3-a_1}=\frac{a_2-a_1}{a_2-A_3+C_3};$$

équation du cône des axes de polarisation:

$$(A_1-C_1-a_3)\alpha^2+(a_3-A_2+C_2)\beta^2+(a_2-a_1)\gamma^2=0.$$

(II′) $$a_1=a_2=a_3(=\mu), \quad B_3=-\mu\pm\sqrt{(\mu-A_1)(\mu-A_2)}\ ;$$

équation du cône des axes de polarisation: $\gamma=0$.

(III′) $$a_1=A_2, \quad a_2=A_1, \quad B_3=-a_3\pm\sqrt{(a_3-A_1)(a_3-A_2)}\ ;$$

équation du cône des axes de polarisation: $\gamma=0$.

(IV′) $$B_3+a_3=0, \quad a_3=A_1=A_2;$$

équation du cône des axes de polarisation: $\gamma=0$.

(V′) $$B_3+a_3=0, \quad a_2=A_1, \quad a_3=a_1;$$

équation du cône des axes de polarisation: $\gamma=0$.

(VI′) $$a_1=a_2=a_3(=a), \quad B_3=-a, \quad B_1=-a\pm\sqrt{(a-A_1)(a-A_2)}\ ;$$

équation du cône des axes de polarisation: $\alpha=0$.

(VII′) $$A_3=a_2,\ A_2=a_3,\ B_1=-a_1\pm\sqrt{(a_1-a_2)(a_1-a_3)},\ B_3+a_3=0;$$

équation du cône des axes de polarisation: $\alpha=0$.

(VIII′) $$B_1+a_1=0, \quad B_2+a_2=0;$$

équation du cône des axes de polarisation:

$$[(A_1-a_2)\alpha^2+(a_3-a_1)\beta^2+(a_2-A_3)\gamma^2][(a_3-a_2)\alpha^2+(A_2-a_1)\beta^2+(a_1-A_3)\gamma^2]$$
$$=(B_3+a_3)^2\alpha^2\beta^2.$$

2° *Milieux pour lesquels toute direction de l'espace est un axe de polarisation.*

(A') $$A_i=\frac{(\lambda_j+\mu)(\lambda_k+\mu)}{\lambda_i+\mu}, \quad B_i=\lambda_i, \quad a_i=\mu,\ (i,\ j,\ k=1,\ 2,\ 3;\ i\neq j\neq k).$$

(B') $$A_3=a_1=a_2=a_3(=\mu), \quad B_3=-\mu\pm\sqrt{(\mu-A_1)(\mu-A_2)}, \quad B_1=B_2=-\mu.$$

(C') Milieux à isotropie transversale dont les coefficients satisfont aux relations

$$\mathfrak{L}+2\mathfrak{M}=a,\ \mathfrak{M}-\mathfrak{N}=a,\ \mathfrak{P}=a.$$

(D') $A_1=A_2=A_3=a_1=a_2=-B_1=-B_2=-B_3=a,\ a_3=\mu$ (valeur quelconque).

XVI.

Après avoir recherché les directions d'ondes possédant avec les ondes en milieu isotrope l'une des deux propriétés communes suivantes: avoir l'une des trois vibrations longitudinales ou avoir une seule vibration polarisée, nous allons indiquer comment on peut chercher les ondes possédant ces deux propriétés et auxquelles, pour abréger, nous donnerons le nom *d'ondes isotropoïdes.* Les milieux se classent, à ce point de vue, comme il suit:

1° D'une façon générale, *dans un milieu quelconque, il n'y a pas d'onde isotropoïde.*

En écrivant que l'une des directions, en nombre fini, d'axes de polarisation coïncide avec l'une des directions à vibration longitudinales, on aura la condition la plus générale d'existence d'une telle onde. Cela correspond en général à deux conditions entre les coefficients du milieu.

2° Mais, *si le milieu possède un cône d'axes de polarisation ou, au contraire, un cône de normales d'ondes à vibration longitudinale,* il suffira d'une seule condition pour écrire que ce cône contient au moins une direction de l'autre système et, par suite, que le milieu possède une direction isotropoïde. Cela fait, en tout, trois conditions, en général, deux assurant dans le cas général, l'existence du cône et la troisième étant la condition, pour ce cône, de contenir une direction de l'autre système.

Ces conditions sont longues et pénibles à expliciter et nous ne le ferons pas.

Il est à noter que le cas 1° est plus général que le 2°; il le contient comme cas particulier. Une observation analogue s'appliquerait à ce qui suit.

3° Le *milieu possède un cône de directions d'axes de polarisation et un cône de normales à des ondes longitudinales.* Cela correspond, en général, à quatre conditions.

Il y a alors *quatre directions isotropoïdes en général.*

Un cas intéressant est celui où les *cônes coïncident* en tout ou en partie. Il y a alors une infinité de directions isotropoïdes. Ce cas n'est pas le seul comme nous le verrons. La recherche systématique des conditions nécessaires et suffisantes pour la coïncidence de deux nappes des cônes définis ci-dessus est encore assez longue et nous nous bornerons à citer les exemples suivants:

(*a*) Milieux définis par les formules

$$(66) \qquad \begin{cases} N_1=\lambda\epsilon_1+\gamma_3\epsilon_2+\nu_2\epsilon_3, \\ N_2=\nu_3\epsilon_1+\lambda'\epsilon_2+(\lambda'+2\mu)\epsilon_3, \qquad T_i=\mu\gamma_i, \\ N_3=\nu_2\epsilon_1+(\lambda'+2\mu)\epsilon_2+\lambda'\epsilon_3. \end{cases}$$

Toutes les directions du plan $\alpha=0$ sont isotropoïdes.

(*b*) Milieux à isotropie transversale dont les coefficients satisfont à l'une des relations ou groupes de relations suivants:

(I) Les coefficients du milieu sont liés par la relation:

$$(\mathfrak{M}-a-\mathfrak{N})\left[\mathfrak{L}+2\mathfrak{N}+\mathfrak{P}-a-\frac{(\mathfrak{L}+2\mathfrak{N}+a)^2}{\mathfrak{L}+\mathfrak{M}}\right]=(\mathfrak{M}-a)(\mathfrak{P}-a). \tag{67}$$

Le cône des directions isotropoïdes est de révolution et a pour équation:

$$(a-\mathfrak{M}+\mathfrak{N})(\alpha^2+\beta^2)=(a-\mathfrak{P})\gamma^2. \tag{68}$$

(II) Les coefficients du milieu sont liés par le groupe des deux relations:

$$2\mathfrak{L}+3\mathfrak{M}=a,\quad \mathfrak{L}+2\mathfrak{N}+a=0. \tag{69}$$

On trouve, comme ci-dessus, un cône de révolution; son équation est:

$$\frac{2\gamma^2}{\alpha^2+\beta^2}=\frac{\mathfrak{M}-a}{\mathfrak{P}-a}. \tag{70}$$

(III) Les coefficients du milieu sont liés par les deux relations:

$$\mathfrak{M}=a,\quad \mathfrak{L}+2\mathfrak{N}+a=0; \tag{71}$$

le cône se réduit au plan d'isotropie $\gamma=0$.

XVII.

Enfin, *si un milieu de Green à trois plans rectangulaires de symétrie est tel que toutes les directions soient des directions d'axes de polarisation ou des directions des normales d'ondes à vibration longitudinale, il jouit en outre de la propriété d'avoir un cône* de directions de l'autre type.*

Ce cône est un cône de directions isotropoïdes.

Cette propriété se démontre en vérifiant que les coefficients des milieux qui satisfont aux relations (A'), (B'), (C') ou (D') satisfont, de ce fait, aux relations (I) (ou ses variantes) ou (II) et inversement, que les coefficients des milieux qui satisfont aux relations (A) vérifient aussi l'un des groupes (I′) à (VIII′).

Il est d'ailleurs bien évident que *tout milieu possédant un cône de directions isotropoïdes devra être obtenu comme il est dit ci-dessus* (N° XVI-3° *ou* N° XVII, les six premières lignes).

XVIII.

On est alors tout naturellement amené à se poser la question suivante:

Existe-t-il, en dehors des milieux isotropes, des milieux tels que toutes les directions de l'espace soient isotropoïdes?

*Le mot cône étant pris dans son sens général, il faut comprendre soit un cône indécomposable soit un cône dégénéré en plusieurs cônes distincts pouvant d'ailleurs eux-mêmes se réduire partiellement ou totalement à des plans.

En se bornant aux milieux à trois plans rectangulaires de symétrie, on est conduit à la seule solution constituée par le milieu à isotropie transversale dont les coefficients satisfont aux relations:

$$\mathfrak{L}+2\mathfrak{M}=a,\ \mathfrak{M}-\mathfrak{N}=\mathfrak{P}=a.$$

XIX.

On a surtout examiné, dans ce qui précède, les milieux de Green à trois plans de symétrie rectangulaires. Ce sont de beaucoup les plus intéressants.

Outre qu'ils comprennent plusieurs systèmes cristallins, il est assez vraisemblable que leur étude peut s'appliquer même à des systèmes qui ne rentrent pas dans cette catégorie par la seule nature de leur symétrie cristalline mais dont les propriétés élastiques possèdent cette symétrie.

C'est un fait analogue qui se produit, par exemple, pour les indices de réfraction qui, même dans le système anorthique, peuvent être représentés par l'ellipsoïde des indices.

D'ailleurs, dans le cas général, outre qu'on se heurte à une extrême complication des calculs, il ne faut pas se dissimuler ce que les résultats ont de physiquement illusoire, vu le nombre trop élevé des coefficients. Déjà les neuf coefficients des milieux à trois plans de symétrie sont un nombre bien considérable pour la technique actuelle.

Les recherches physiques qui pourraient être poursuivies dans la voie indiquée par la présente étude semblent devoir être aiguillées avec intérêt vers les milieux doués, non seulement de la symétrie trirectangulaire, mais encore de quelque autre propriété diminuant le nombre des coefficients élastiques offerts aux recherches du physicien.

QUANTUM THEORY OF PHOTOGRAPHIC EXPOSURE

By Dr. Ludwik Silberstein,
Research Laboratory of the Eastman Kodak Company, Rochester, New York, U.S.A.*

The photographic emulsion, though in appearance continuous and uniform, consists of a large number of minute grains of a silver halide, say Ag Br, crystals of the simple cubic system, shaped as flat thin plates, mostly triangular or round, ranging in size from a fraction of a micron to several microns. In practical photography the emulsion spread over the base, plate or film, contains many "layers" of grains, *i.e.*, a three-dimensional distribution of Ag Br crystals of a considerable depth. For experimental purposes, however, a single-layer coating is used, so as to simplify the conditions of exposure and to facilitate the counting and planimetrizing of the individual grains. A typical sample of such a single-layer plate, prepared and photographed under a magnification of 2500 by Mr. Trivelli, is shown in Fig. 1. The Ag Br grains are distributed irregularly over the area of the plate as flat targets of different sizes.

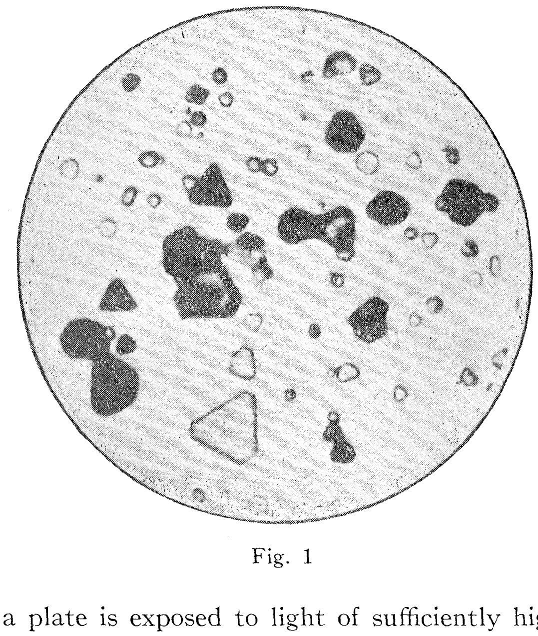


Fig. 1

When such a plate is exposed to light of sufficiently high frequency, and plunged into a developer, some of the grains are blackened, while others lying near around them, remain unchanged. A prolonged exposure would affect

*Communication No. 217.

these also, until one by one all the available grains will have succumbed to the incident light. Such a behaviour naturally reminded one of a rain or shower rather than a flood action, and suggested the treatment of the photographic exposure or formation of the latent image by means of Einstein's light quanta or, as I prefer to call them, light darts, as contrasted with the classical assumption of a continuous distribution of energy in the incident light beam.

The purpose of this paper is to give a review of the theoretical and experimental results obtained in the Eastman Research Laboratory during the last three years and described more fully in a number of papers on the Quantum theory of Photographic Exposure* and to indicate briefly some further yet unpublished results which seem to strengthen the evidence in favour of that theory.

It is true that Einstein's concept of light quanta, first proposed in 1905, has since found a stout support in several types of experiments, more especially those which gave, mainly in the hands of Millikan, a most excellent verification of Einstein's photo-electric law based on that concept. Yet these experiments prove in ultimate analysis only that for the liberation of each single electron a full quantum $h\nu$ of light energy is required and is actually taken up by the photo-electric plate, say a potassium cell, and that this process takes but an light immeasurably short time; but they do not actually show that the impinging is in itself a discrete assemblage of parcels of light (each containing a quantum) haphazardly distributed and rushing through space independently of each other. For, even if we knew beforehand such to be the case, the photo-electric plate or cell, being to all purposes a unique target, could not reveal the discrete nature of the onrushing shower of light darts. Again, the very nature of the photo-electric effect, as actually observed (viz., number of electrons ejected proportional to the impinging light energy) is such that the subdividing of the cell into many disconnected, independent little targets would in no way influence the total amount of the effect. *Caeteris paribus*, the total number of electrons liberated from all these minute targets would be the same as from the original unique large target.

The conditions in the case of a photographic emulsion (a plate exposed to light and "developed") are essentially different, and precisely such as to enable us to distinguish between, so to speak, a flood- and a rain-action. This difference between photography (which very likely is also a photo-electric phenomenon) and photo-electric experiments proper is mainly due to the well-established fact that a grain of an emulsion exposed to light is either not affected at all or is made developable entirely, as a whole, no matter how large it is. This holds even in the case of certain types of clumps, assemblages of grains in intimate contact with one another, each of these clumps, as ascertained by special experiments, playing the rôle of a single target for exposure purposes. In our earlier experiments (Paper I), aiming at a possibly wide range of target areas, an emulsion rich in clumps, containing from two up to thirty-three grains, was used. In most of the later experiments, however, the emulsions were practically composed of single grains, each acting as a single independent target.

*Phil. Mag., Vol. XLIV, p. 257-273; 956-968; Vol. XLV, p. 1062-1070. These will be shortly referred to as Papers I, II, III.

The sizes a (areas) of these targets which lie "flat", *i.e.*, almost parallel to the base, range from sub-microscopic ones up to 20 square microns, grains larger than this being rather sporadic. It will be well to mention here that all, even the most uniform emulsions practically obtainable and used in microphotographic experiments, consist of grains of perceptibly different sizes, the distribution of sizes a among them being in each case characterized by what is called the size-frequency curve of an emulsion. If $f(a)da$ be the number (per unit area of the photographic plate) of grains whose areas range from a to $a+da$, the emulsion may be briefly referred to as one of type $f(a)$, as *e.g.*, of exponential or of Gaussian (error-function) type. But the knowledge of the type is required only in order to derive the integral law, concerning the photographic "density" as a whole, from the elementary law of exposure concerning each size-class of grains separately. To begin with the latter, let there be, per unit area of the plate, N grains or targets each of size (area) a, which can be measured, and let n be the number of light-darts, again per unit area, impinging upon the plate. Let ϵ be the fraction* of the area of each target vulnerable to light, *i.e.*, such that on being hit by at least one light-dart of sufficiently small wave-length, it becomes itself and therefore also makes the whole target or grain developable. Then the total number k of grains affected or made developable by this shower of n darts will be, statistically (cf. Paper I),

$$k = N(1 - e^{-\epsilon a n}). \tag{1}$$

If we put for brevity

$$v = \log \frac{N}{N-k}, \tag{2}$$

this, the fundamental microphotographic formula can be written

$$v = \epsilon a n. \tag{3}$$

If the N grains are not rigorously equal in size but form a class of average size a and of a comparatively small breadth 2α, *i.e.*, ranging in size from $a-\alpha$ to $a+\alpha$, the last formula is to be replaced by the following one (Paper II)

$$v = \epsilon a n - \log \frac{\sinh(\epsilon \alpha n)}{\epsilon \alpha n},$$

which follows by integrating (1) over the whole breadth 2α under the assumption that the distribution of sizes within this narrow range is uniform, *i.e.*, $f(a)$ constant. Up to terms of the fourth order in $\epsilon \alpha n$ the last formula can be written, to all practical purposes,

$$v = \epsilon a n - \tfrac{1}{6}(\epsilon \alpha n)^2. \tag{3a}$$

If, in a somewhat broader treatment, the sensitive part ϵa of a grain, instead of being all in a lump, consists of a number κ of sensitive "spots", each of area ω, if $\bar{\kappa}$ be the average number of spots per grain (of a given size a), and if the

*Its numerical value to be determined presently from the experimental findings.

spots are assumed to be distributed haphazardly over the grains, then (Paper III) formula (3) is replaced by

$$v = \bar{\kappa}(1 - e^{-n\omega}) \doteq \bar{\kappa} n\omega(1 - \tfrac{1}{2}n\omega)$$

or, the average specific sensitivity ϵ being now given by $\epsilon = \bar{\kappa}\omega/a$,

$$v = \epsilon an(1 - \tfrac{1}{2}\epsilon an/\bar{\kappa}). \tag{4}$$

Thus far, however, our experiments were not accurate enough to distinguish conclusively between (3) and (4), though in a few cases we had indications of the refinement embodied in the second term of the latter formula, enabling us to obtain a rough estimate of the number $\bar{\kappa}$ as given by v_∞ or $\log \frac{N}{N-k}$ for $n = \infty$, and amounting to something like 4 to 6 spots per square micron in the case ot the Seed Lantern emulsion.

For the present the simpler formula (3) or, for a finite class breadth 2α, the formula (3*a*) is sufficient to cover adequately all our findings.

Passing to the experimental verification of the fundamental formula (1) or (3), with the class-breadth correction as in (3*a*), it may in the first place be mentioned that a considerable number of grain counts and measurements made by my colleagues during 1921-23 gave results in very satisfactory agreement with this formula with ϵ constant, *i.e.*, for a given emulsion independent of a and of n. Part of these results will be found in the papers quoted above. [Some of these were also shown on a slide in reading the paper.] Without re-discussing these older experiments I shall now pass to a brief description of some results not yet published, obtained more recently from microphotographic counts and absolute energy measurements made by my colleagues Mr. A. P. H. Trivelli and Mr. A. L. Schoen, and tending not only to confirm the constancy of the sensitivity fraction ϵ, but also to measure its absolute value under the explicit assumption that each impinging dart contains an actual quantum, $h\nu$, of energy.

The following table contains the values of the product ϵn_1 obtained by means of (3*a*) from the observed values of N, k,* and thence v, for ten successive "steps" or exposures, proceeding after the powers of $\sqrt{2}$, and for six size-classes of grains of a single-layer Seed Lantern emulsion exposed to light. The number of impinging light-quanta per square micron of the plate in the first step being denoted by n_1, that in the mth step was $2^{\frac{m-1}{2}} n_1$. The second (correction) term in (3*a*) amounted only to 0.060. Thus the quotients of the observed $v + 0.060$ into $2^{\frac{m-1}{2}}$ and into the average class-size, as indicated in the first column, should represent the product ϵn_1. There are in all thirty values†. It will be seen that apart from oscillations, which are on the whole irregular and well within the limits of experimental error, the product ϵn_1, and therefore also the specific sensitivity ϵ is fairly constant, *i.e.*, independent of the size of the grains

*For the technique of these measurements and counts see a paper by L. Righter and A. P. H. Trivelli in Phil. Mag., vol. XLIV, p. 252.

†The blanks are due to the technical difficulties of securing for the present reliable data for very small grains and weak exposures as for very large ones and strong exposures.

Step n/n_1	I	II	III	IV	V	VI	VII	VIII	IX	X	$n_1\bar{\epsilon}$
$\bar{a}_{\mu^2}$	1	$\sqrt{2}$	2	$2\sqrt{2}$	4	$4\sqrt{2}$	8	$8\sqrt{2}$	16	$16\sqrt{2}$	over steps
0.115						1.17	1.40	1.34	1.20	1.21	1.264
0.30				0.92	1.25	1.29	1.15	0.95	0.82	0.96	1.049
0.50		0.95	1.23	1.39	1.39	1.22	1.15	0.96	0.91		1.150
0.70	1.133	1.41	1.40	1.51	1.42	1.13					1.334
0.90	1.130	1.68	1.50								1.437
1.10			1.36								1.360
$n_1\bar{\epsilon}$ over sizes	1.132	1.345	1.373	1.273	1.353	1.203	1.233	1.083	0.977	1.085	1.217_9 / 1.217_7

and of the exposure. The arithmetical means over the sizes and the exposures are given in the last line and the last column respectively. The general mean of the 30 items is $\epsilon n_1 = 1.218$. In order to find the value of ϵ belonging to the grains of this emulsion, it remained to find the number of quanta of (blue) light to which the plate was exposed in the first step. Now, from Mr. Schoen's absolute energy measurements and with the wave-length $\lambda = 430\mu\mu$ as a compromise value for the standardized blue filter used in these experiments, it was found that $n_1 = 31{,}000$ per square micron. Whence, the specific sensitivity of the Seed Lantern emulsion, $\epsilon = 0.39.10^{-5}$, showing that only a very small fraction of the area of each grain is actually sensitive to light. This result was by no means unexpected. For our previous experiments (Paper III) with another emulsion, *W12C*, and $\lambda = 470$, gave $\epsilon = 1.2.10^{-5}$, a fraction of the same order. Again, from a recent set of experimental data for a Graflex emulsion obtained by my colleague Dr. E. P. Wightman ($\lambda = 460$), I find, for four different exposures, and in each case as an average over six size-classes of grains,

$$\epsilon = 3.26,\ 3.28,\ 3.38,\ 3.25.10^{-5},$$

with a general mean of $3.3.10^{-5}$.

In view of the probable photo-electric nature of the exposure or the production of the latent image, as already suggested by Joly, it is interesting to compare these photographic results with some properly photo-electric ones. Now, in Elster and Geitel's photo-electric experiments with what these eminent specialists considered a very sensitive potassium cell only one in 2300 blue light-quanta hitting the cell knocked out an electron*, and most of their cells were even several times less efficient. The simplest interpretation of this result is that only 1/2300 of the total area of the potassium cell was sensitive, *i.e.*, ready to part with an electron on being hit by a light-dart. Such behaviour

*This efficiency was only occasionally exceeded, about forty times, in certain experiments of Pohl and Pringsheim. Cf. A. L. Hughes, *Photo-Electricity*, 1914.

can reasonably be ascribed to some local irregularities* of the otherwise perfect surface of potassium and similarly, in our photographic case, some local crippling of the otherwise perfect cubic space-lattice structure of the silver bromide grains. Thus 1/2300 would simply be the value of ϵ to be ascribed to the said potassium cell, and gathering the results, we have the following little table of ϵ-values:

Seed Lantern	$0.4.10^{-5}$,
$W12C$	$1.2.10^{-5}$,
Graflex	$3.3.10^{-5}$,
Potassium cell	43.10^{-5}.

Thus one of the best potassium cells is only about 13 times as efficient as the grains of the most sensitive of these photographic emulsions, while this is over eight times as sensitive as our first emulsion. In fine, the sensitive fraction ϵ of the total area of these heterogeneous targets is very much of the same order, and its value, which at first might seem surprisingly small, receives thus a certain degree of plausibility. The actual smallness of ϵ, especially in the case of photography, may also have some practical interest. For it makes it reasonable to try to increase it considerably, pushing its value towards unity as its ideal limit. It may turn out that, as in the case of photo-electricity proper, the occlusion of foreign atoms within the lattice of the silver halide grains will help to solve this practical problem. But considerations of such a nature are entirely beyond the scope of the present paper.

Thus far the microscopic aspect of the theory relating, that is, to the individual grains of an emulsion. Passing to the macroscopic part of the subject, it will be enough to consider here only that integrated or density formula which holds for commercial (many-layered) emulsion coatings of the exponential type, in the sense explained at the beginning of the paper. In fact, it appears that the effective† frequency curve for most of the emulsions dealt with in experimental and in practical work is, to all purposes, adequately covered by the exponential function

$$f(a) = Ce^{-\mu a},$$

where C,μ are constants. Such a distribution, moreover, can safely be assumed to hold from $a=a_1=0$ to $a_2=\infty$. (The more general case of any finite a_1, a_2, given in Paper II, is of purely academic interest.) Under these circumstances, and integrating the fundamental formula (1), the total area affected by a shower of n light-darts is found to be

$$\sigma = A - \frac{A}{(1+\epsilon n/\mu)^2},$$

where $A = \int_0^\infty f(a)da = 1-\theta$, say, is the total available area of silver halide (always

*Due perhaps to the presence of some foreign atoms, such as the occlusion of minute amounts of argon in Pohl and Pringsheim's experiments.

†Thus to call the frequency curve when account has been taken of the mutual shielding of targets occurring in a many-layered coating. Cf. Paper I, p. 265.

per unit area of the plate or film). In all practical cases θ is but a small fraction, 10^{-3} and less, namely the fraction of the plate area uncovered by silver halide.

Now the so-called photographic density D is defined as the common logarithm of the opacity of the plate, which is the reciprocal of the transparency, and the latter can roughly be taken to be $1-\sigma$. Thus

$$D = -\text{Log }(1-\sigma), \tag{5}$$

where Log stands for log base 10. If E be the exposure in any conventional units, we can put

$$\frac{\epsilon n}{\mu} = \alpha E,$$

where α will be a plate-constant to be determined empirically. Thus the density as a function of exposure becomes

$$D = -\text{Log}\left[\theta + \frac{1-\theta}{(1+\alpha E)^2}\right]. \tag{6}$$

If D_m is the extreme density obtainable on the given plate, the directly ascertainable meaning of θ is given by $D_m = -\text{Log }\theta$, which follows from (5) for $\sigma = A = 1-\theta$ or also from the final formula for $E=\infty$. In most practical cases D_m certainly exceeds 2, so that $\theta < 0.01$.

This new type of formula turned out to represent very closely a large number of densitometric observations made on a variety of plates, exposed to either X-rays or ordinary light, by my colleagues Mr. R. B. Wilsey and Mr. Emery Huse.

To begin with the former, I tested first an important sub-case of (6), to wit, for negligibly small θ or rather for αE not large enough to reveal θ and thus to reach what is known as "the shoulder" of the characteristic curve* of a plate or film. This sub-case of the complete formula, with $\theta \doteq 0$, is

$$D = 2 \text{ Log }(1+\alpha E), \tag{6a}$$

a formula of remarkable simplicity, yet one that turned out to cover accurately the densities observed in a very wide range of exposures. In the following table the third line contains the densities observed by Mr. Wilsey on a film exposed to X-rays in the range of $E=1$ to $64\sqrt{2}$, the development conditions being the same for all fourteen exposures. The second line consists of the D-values calculated by (6a) with $\alpha = 0.151$. The agreement is very close indeed, the deviations δD being well within the limits of experimental error. An equally

E	1	$\sqrt{2}$	2	$2\sqrt{2}$	4	$4\sqrt{2}$	8	$8\sqrt{2}$	16	$16\sqrt{2}$	32	$32\sqrt{2}$	64	$64\sqrt{2}$
$D_{\text{calc.}}$	0.12	.17	.23	.30	.41	.54	.69	$.86_5$	1.07	1.29	1.53	1.79	2.06	2.33
$D_{\text{obs.}}$	0.08	.15	.20	.29	.40	.54	.72	.90	1.10	1.32	1.56	1.80	2.12	2.48
δD	+.04	+.02	+.03	+.01	+.01	.00	−.03	$-.03_5$	−.03	−.03	−.03	−.01	−.06	−.15

*D plotted against Log E.

good agreement with (6*a*) and in a yet wider range, was obtained for a set of X-ray experiments made more recently by Mr. Wilsey's temporary collaborator Mr. A. R. Riddle. An important point in connection with these, and all the following experimental verifications is the selection, with a given developer, of the proper development time. Several copies of each sensitometric strip should be prepared, or one cut up into several narrower strips, each extending over the whole range of exposures, and the different strips developed for various times so distributed as to give a sufficiently complete information about the progress of development. Of these strips one, corresponding to a certain development time, will be picked out and found to agree with the formula, with a constant coefficient α. For either a shorter or a longer development time the values of α determined from pairs of observed D and E will be found to vary systematically all along the range of exposures. Thus, for instance, 6 minutes was the development time for Mr. Wilsey's data, just quoted, giving an almost perfect agreement, whereas strips developed 5 minutes and 7 minutes led to a series of α-values decreasing and increasing, respectively, with increasing exposure. Manifestly, such a selection of development time is unavoidable, for, since our formula gives only the area of the silver halide "affected" by light or made developable, but not yet actually developed, certain development times will be too short to blacken entirely all the grains thus affected, while others will be long enough to blacken not only these but also a number of other grains not hit by light-darts at all (fog). A formula taking account explicitly of the development-time as well as of exposure has recently been constructed and will be given in a forthcoming paper. Meanwhile we shall be content with this representation of density as function of exposure alone. Thus far the sub-case (6*a*).

Next, to cover the experimental characteristic curves stretching over E-values high enough to show the "shoulder," we have to use the complete formula (6), with the two plate-constants θ and α, of which the former is readily determined from the highest steps by means of $D_m = -\mathrm{Log}\,\theta$, and the latter from a few of the lower steps, where θ plays no perceptible rôle. The close agreement

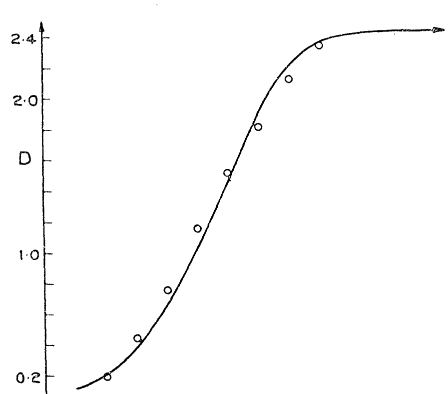


Fig. 2

of the formula with characteristic curves observed on several types of emulsions (plates and films) exposed to either light or X-rays will be seen from the Figs. 2, 3, 4, 5, in which the circlets represent data observed by Mr. Huse, and (in the case of Fig. 3) by Mr. Wilsey, and the full curves are drawn according to formula

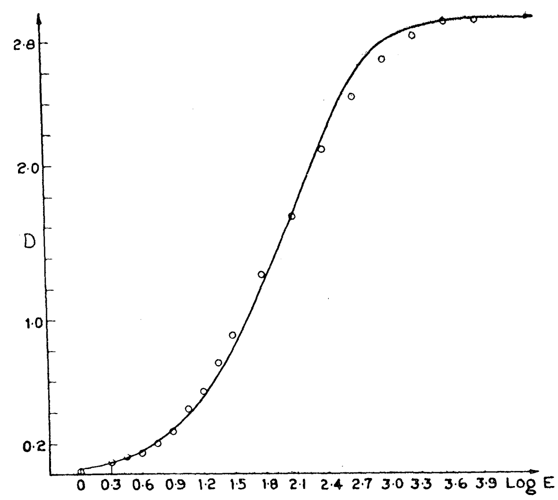


Fig. 3

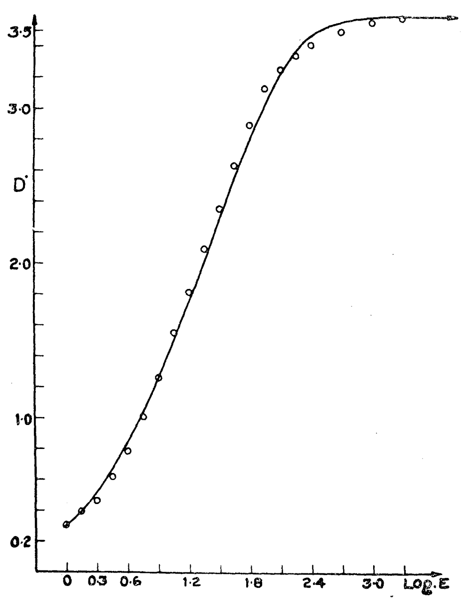


Fig. 4

(6) with the following values of the constants: For Fig. 2, relating to Seed 26, exposed to light, $\theta=0.0035$, $\alpha=0.150$; for Fig. 3, Eastman X-ray Film, single coating, exposed to X-rays, $\theta=0.00140$, $\alpha=0.0487$; for Fig. 4, relating to Eastman

Positive Film, exposed to light, $\theta = 0.00025$, $\alpha = 0.416$, and finally, for Fig. 5, concerning a Seed sample (reference No. 6429), exposed to light, $\theta = 0.00140$, $\alpha = 0.244$. The differences between the observed and the calculated densities scarcely exceed the limits of experimental error. A good number of other results of much the same kind, obtained in the course of the last two years,

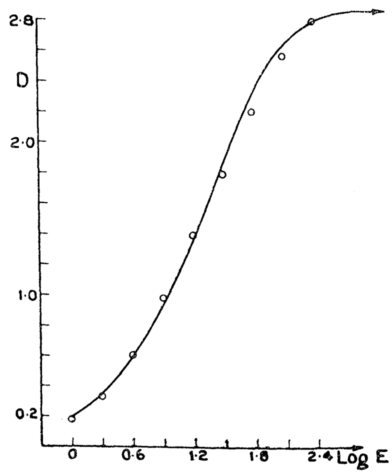


Fig. 5

might be shown, but these few, of which some cover a very wide range of exposures, will suffice to support the proposed densitometric formula and through it, also, to a certain extent, the fundamental formula (1), from which it followed by integration over all grain sizes under a reasonable assumption as to the type of the frequency curves of the emulsions.

Note added in reading the proof, October 31, 1927. In the course of the three years which have elapsed since this paper, as read at the Congress, was written, a further scrutiny of the microphotographic data has necessitated a complete modification of the proposed exposure theory, transferring the element of chance from the impinging light to some peculiarities (nuclei) of the silver halide grains themselves, haphazardly distributed among them. For details of criticism see the writer's recent paper, Criticism to the Theory of Photographic Exposure (now in the press of the Phil. Mag.), where also the alternative, nuclear, theory is fully developed.

A SHORT CONTRIBUTION TO THE KINETIC THEORY OF GASES

By Dr. Cristóbal de Losada y Puga,
Lima, Peru.

I intend to deal in this paper with the question of the probability of a free path taken at random reaching the length r, and with that of the probability that its length lies between r and $r+dr$. As soon as a molecule A has collided with another, which we will call B, molecule A is projected in a certain direction, along which it continues moving until it collides with another molecule C: the path described by molecule A between these two collisions is what we call the free path.

We may define the probability P of a free path reaching the length r as

$$P=f(r). \tag{1}$$

The probability Q that the length of the free path will lie between r and $r+dr$ is obtained on taking the first differential of P, that is to say

$$Q=-\frac{d.f(r)}{dr}\,dr. \tag{2}$$

According to its physical meaning, function (1) must be continuously decreasing; because in order to reach a given value, the length of the free path under consideration has to take in succession all the lower values.

In the Kinetic Theory of Gases, an expression for formula (1) is given, viz.,

$$P=e^{-\frac{r}{j}} \tag{3}$$

where j is the mean free path. (Cf., e.g., Jeans, *Dynamical Theory of Gases*, 3d ed., p. 257).

The plan of this paper is as follows:

I. Analysis of formula (3).

II. Deduction of a new form of equation (1), instead of (3).

I. Analysis of Formula (3)

Equation (3) defines P as a continuously decreasing function of r, and its first derivative is also monotonic. Now, in accordance with its physical meaning, function Q must not be monotonic, but first increasing and then decreasing.

In effect, let s be a length exceedingly small compared with the mean value of the distance from any molecule to the nearest one. Then, it seems that a free path of length $2s$ must be more probable than a path of length s, because for the path to have the length s it would be necessary that molecule C be at a distance s from the point where A and B would collide. Now, if the mean value of the distance from any molecule to the nearest one is very large compared with s, it will be more probable for molecule C to be at a distance $2.s$ than at a distance s: function Q, therefore, must be increasing for small values of r. On the other hand, the free paths of great length will have very few chances of occurring; and the greater the length, the less the probability. Consequently, for large enough values of r, Q must be a decreasing function, and it cannot be monotonic.

The reason why function (3) presents this inadequacy, is perhaps the method employed in deducing it, which involves the misstatement of assuming the decreasing probability even for small values of r.

But if the number of molecules tends to prevent the paths being too long, the distances between the molecules tend to prevent the paths being too short.

II. Deduction of a New Form for Equation (1) Instead of (3)

To find a new form of equation (1) which is free from the objections raised against (3), let us consider molecule A at the instant at which, owing to its collision with B, it is projected and begins to travel over the free path whose length is the variable r of our equations. Let us imagine the molecule A surrounded by its "protecting sphere" (sphère de protection, Abstandssphäre) whose radius σ is equal to the diameter of the molecule. As the molecule travels along the path r, the protecting sphere will generate a cylinder of volume $\pi\sigma^2 r$. The free path will end as soon as the centre of any molecule C lies within this cylinder. Let us imagine now a sphere of radius r, whose centre lies at the point where the centre of molecule A was at the instant of its collision with B: the volume of this sphere will be $\frac{4}{3}\pi r^3$. Now, if we suppose that within this sphere there is a single molecule, the probability that the centre of this molecule lies within the cylinder will be

$$\frac{\pi\sigma^2 r}{\frac{4}{3}\pi r^3} = \frac{3\sigma^2}{4r^2}$$

and the probability that it lies outside of the cylinder (*i.e.*, the probability that the free path of molecule A be uninterrupted within the sphere of radius r, or the probability that the free path reaches at least the length r) will be

$$1 - \frac{3\sigma^2}{4r^2}.$$

Now, if instead of there being a single molecule in the sphere of radius r, there should be N, the probability of the free path reaching the length r, would be

$$P = \left(1 - \frac{3\sigma^2}{4r^2}\right)^N. \tag{5}$$

The number N of molecules contained in the sphere of radius r is evidently $\frac{4}{3}\pi r^3 n$, n being the number of molecules contained in the unit of volume (this depends only on temperature and pressure). Replacing this value of N in (5), we have:

$$P=\left(1-\frac{3\sigma^2}{4r^2}\right)^{\frac{4}{3}\pi r^3 n}. \tag{6}$$

The probability Q that a free path lies within r and $r+dr$, is obtained by taking the first differential of (6), r being the independent variable; this gives the expression

$$Q=-\frac{4}{3}\pi r^2 n\left(1-\frac{3\sigma^2}{4r^2}\right)^{\frac{4}{3}\pi r^3 n-1}\left[3\left(1-\frac{3\sigma^2}{4r^2}\right)\log\left(1-\frac{3\sigma^2}{4r^2}\right)+\frac{3\sigma^2}{2r^2}\right]dr. \tag{7}$$

Finally, let us compare our fundamental equation (6) with the classical equation (3). To do this, equation (6) may be written:

$$P=\left[\left(\frac{1}{1-\dfrac{1}{\left(\dfrac{4r^2}{3\sigma^2}\right)}}\right)^{\frac{4r^2}{3\sigma^2}}\right]^{-\pi\sigma^2 nr}.$$

Calling

$$\frac{4r^2}{3\sigma^2}=m, \qquad \pi\sigma^2 n=\frac{1}{j} \text{ (Clausius' value)},$$

we have

$$P=\left\{\left[\frac{1}{1-1/m}\right]^m\right\}^{-r/j}=\left\{\left[\left(1+\frac{1}{m}\right)+\left(\frac{1}{m^2}+\frac{1}{m^3}+\frac{1}{m^4}+\ldots\right)\right]^m\right\}^{-r/j}.$$

and developing by the binomial theorem,

$$P=\left\{\left(1+\frac{1}{m}\right)^m+\frac{1}{m}\left(1+\frac{1}{m}\right)^{m-1}\left(1+\frac{1}{m}+\frac{1}{m^2}+\ldots\right)\right\}^{-r/j}.$$

As r/σ increases indefinitely, m becomes very great; and then the first term of the series becomes e in the limit and all the others zero: therefore, for great values of r/σ equations (3) and (6) tend to equality.

There are, as is known, several expressions for j: one deduced by Clausius on the hypothesis that the molecule under consideration moves among a cloud of molecules at rest; a second value deduced by the same author, on the supposition that all the molecules are moving in every direction but with the same velocity c; and yet another due to Maxwell, obtained by applying the principle of distribution of velocities.

Equation (6), proposed in this paper, coincides at the limit with the classical one, upon the condition that we take for j the first of the expressions just mentioned. This was to be expected, because in deriving the formula (6), I have supposed all the molecules except A, at rest. We might improve our equation (6) by taking account of the chaotic motion of the molecules.

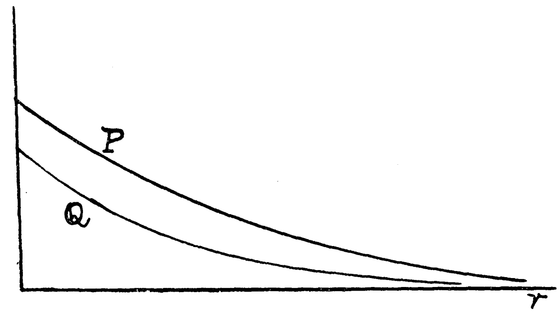


Fig. 1

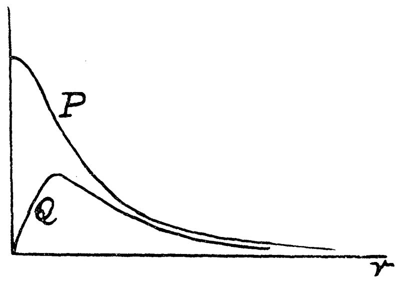


Fig. 2

Figs. 1 and 2 show approximately (they have not been computed) the values of P and Q according to equations (3) and (4), and to equations (6) and (7), respectively.

SUR UN GROUPE DE TRANSFORMATIONS QUI SE PRÉSENTENT EN ÉLECTRODYNAMIQUE

PAR M. S. ZAREMBA,
Professeur à l'Université de Cracovie, Cracovie, Pologne.

Les équations qui existent entre des éléments qui ont une signification physique jouissent ordinairement de certaines propriétés d'invariance faciles à apercevoir avec sûreté à priori. D'autre part, des propriétés de ce genre limitent à elles seules, d'une façon considérable, la variété des hypothèses admissibles, en ce qui concerne la forme des équations correspondantes.

Par conséquent, dans la recherche des équations de la Physique, il serait peut-être avantageux de commencer par tirer parti des propriétés d'invariance des équations demandées d'une façon plus systématique que cela ne se fait d'ordinaire. C'est ce que je me propose de montrer sur un exemple qui présente peut-être quelque intérêt.

Dans ce qui va suivre, j'adopterai la conception traditionnelle de l'espace-temps parce que, jusqu'à présent, on ne sait pas, comme je l'ai fait voir ailleurs,* établir, d'une façon satisfaisante, une correspondance entre les valeurs numériques des grandeurs considérées dans la théorie de la relativité et des opérations de mesure.

1. Pour définir un champ électromagnétique (C), dans le vide, il suffit de choisir un système de coordonnées déterminé et de faire connaître deux vecteurs comme fonctions du temps t et des coordonnées x_1, x_2, x_3 de l'origine de chacun d'eux; l'un de ces vecteurs, soit e, représentera alors la force électrique au point $(x_1\ x_2, x_3)$ à l'époque t, c'est-à-dire la force qui solliciterait à l'époque t un point matériel M chargé de l'unité d'électricité et situé au point (x_1, x_2, x_3); quant au second vecteur, soit m, il représentera la force magnétique au point et à l'époque considérés, c'est-à-dire la force qui solliciterait à l'époque t un pôle magnétique P, d'intensité égale à l'unité, situé au point (x_1, x_2, x_3). Les définitions précédentes impliquent que, par rapport au système de coordonnées (que nous allons nous représenter comme un système de coordonnées cartésiennes rectangulaires), les vitesses du point M et du pôle P sont nulles. Nous avons formulé les définitions considérées de sorte que la circonstance précédente se présente parce que la force dérivant du champ et sollicitant un point électrisé dépend en général non-seulement de sa position à l'époque considérée, mais aussi de sa vitesse par rapport au système de coordonnées (S) et qu'il en est peut-être de

*S. Zaremba. La théorie de la Relativité et les faits observés. Jour. de Math., 1922, p. 105.

même pour un pôle magnétique. Cela posé, il est tout indiqué de donner aux vecteurs e et m les noms de force électrique et de force magnétique du champ (C) *relatives* au système de coordonnées (S); c'est ce que nous allons faire dorénavant.

Voici maintenant un problème qui se présente de lui-même: *Connaissant les forces électrique e et magnétique m d'un champ électromagnétique (C) relatives á un système de coordonnées déterminé (S), déterminer les éléments analogues e' et m', relatifs à un système de coordonnées (S') qui se déplace d'une façon donnée par rapport au système (S).*

Je me propose d'étudier ce problème dans le cas particulier où le système (S') est animé d'un mouvement de translation rectiligne et uniforme par rapport au système (S).

2. Les notations précédentes étant conservées, désignons par u la vitesse constante du système (S') par rapport au système (S). Nous admettrons ce qu'admettent d'une façon plus ou moins explicite tous les physiciens, à savoir que la solution de notre problème peut être représentée par l'ensemble de deux équations vectorielles de la forme suivante:

$$(1) \qquad \begin{cases} e' = \phi(e, m, u), \\ m' = \psi(e, m, u), \end{cases}$$

où ϕ et ψ représentent des caractéristiques qu'il s'agit de déterminer.

Désignons d'une façon générale par

$$e_i', \ m_i', \ e_i, \ m_i, \ u_i, \qquad (i = 1, 2, 3),$$

les projections orthogonales respectives des vecteurs e', m', e, m et u sur l'axe de numéro i dans le système de coordonnées (S). Chacune des équations vectorielles (1) se décomposera en trois équations scalaires de sorte qu'en définitive nous aurons le système suivant de six équations scalaires:

$$(2) \qquad \begin{cases} e_i' = \phi_i(e_1, e_2, e_3, m_1, m_2, m_3, u_1, u_2, u_3), \\ m_i' = \psi_i(e_1, e_2, e_3, m_1, m_2, m_3, u_1, u_2, u_3). \end{cases} \quad (i = 1, 2, 3),$$

Voici d'abord une proposition qui résulte immédiatement de la nature vectorielle des égalités (1):

(A) *Si l'on désigne par ξ_1, ξ_2, ξ_3 les projections orthogonales d'un vecteur quelconque sur les axes du système (S), chacune des expressions:*

$$(3) \qquad \sum_{i=1}^{3} \xi_i \phi_i(e_1, \ldots, m_1, \ldots, u_1, \ldots),$$

$$(4) \qquad \sum_{i=1}^{3} \xi_i \psi_i(e_1, \ldots, m_1, \ldots, u_1, \ldots),$$

représentera un invariant du groupe des rotations.

Pour aller plus loin, considérons un troisième système de coordonnées (S'') animé, par rapport au système (S'), d'un mouvement de translation rectiligne et uniforme avec une vitesse u'. Désignons par

$$e_i'', \quad m_i'', \quad u_i',$$

les projections orthogonales respectives sur l'axe de numéro i dans le système (S) des forces électrique e'' et magnétique m'' du champ (C), relatives au système (S''), ainsi que celle du vecteur u'. Pour obtenir les expressions des e_i'' et m_i'' en fonction des quantités

$$e_k', \; m_k', \; u_k', \quad (k=1, 2, 3),$$

il suffit évidemment de substituer dans les équations (2) aux symboles:

$$e_i', \; m_i', \; e_i, \; m_i, \; u_i,$$

respectivement les symboles:

$$e_i'', \; m_i'', \; e_i', \; m_i', \; u_i'.$$

Nous aurons donc:

$$(5) \qquad \begin{cases} e_i'' = \phi_i(e_1', \ldots, m_1', \ldots, u_1', \ldots), \\ m_i'' = \psi_i(e_1', \ldots, m_1', \ldots, u_1', \ldots). \end{cases}$$

Cherchons maintenant les expressions des e_i'' et m_i'' au moyen des quantités

$$e_i, \; m_i, \quad (i=1, 2, 3).$$

Il y a deux manières d'obtenir les formules demandées:

1° On peut porter les valeurs (2) des e_i' et des m_i' dans les formules (5).

2° Le système (S'') se déplaçant par rapport au système (S) avec une vitesse qui, en projection sur les axes de ce système, a pour expressions

$$u_i + u_i' \quad (i=1, 2, 3),$$

on obtiendra encore les formules demandées en substituant, dans les formules (2) aux symboles

$$e_i', \; m_i', \; u_i, \quad (i=1, 2, 3),$$

les symboles

$$e_i'', \; m_i'', \; u_i+u_i', \quad (i=1, 2, 3),$$

ce qui donne:

$$\begin{aligned} e_i'' &= \phi_i(e_1, \ldots, m_1, \ldots, u_1+u_1', \ldots), \\ m_i'' &= \psi_i(e_1, \ldots, m_1, \ldots, u_1+u_1', \ldots). \end{aligned} \qquad (i=1, 2, 3).$$

Evidemment, les deux procédés doivent conduire aux mêmes expressions des e_i'' et m_i''. Cela posé, on reconnaît sans peine que l'on a la proposition suivante.

(B) *L'ensemble des formules* (2) *définit un groupe* G *de transformations ponctuelles à trois paramètres* u_1, u_2, u_3 *de l'espace arithmétique à* 6 *dimensions; dans ce groupe, les paramètres du produit de deux transformations sont toujours égaux aux sommes des paramètres homologues des facteurs.*

Le groupe désigné par G dans l'énoncé précédent jouit encore évidemment de la propriété que voici:

(C) *Le groupe* G *contient la transformation identique et cette transformation correspond aux valeurs nulles des paramètres.*

3. Pour appliquer la méthode de Sophus Lie à la détermination des transformations du groupe G, il suffirait de connaître les fonctions

(6) $$\xi_{s,i}\ ,\ \eta_{s,i}$$

définies par les formules:

(7)
$$\begin{cases} \xi_{s,i} = \left(\dfrac{\partial \phi_s}{\partial u_i}\right)_{u_1=u_2=u_3=0,} \\ \eta_{s,i} = \left(\dfrac{\partial \psi_s}{\partial u_i}\right)_{u_1=u_2=u_3=0.} \end{cases}$$

En effet, dans ce cas, les fonctions chercheés seraient définies sans ambiguité par les conditions suivantes:

1° Chacune de ces fonctions devra être une intégrale commune du système jacobien:

(8)
$$\frac{\partial f}{\partial u_i} = \sum_{s=1}^{3} \xi_{s,i} \frac{\partial f}{\partial e_s} + \sum_{s=1}^{3} \eta_{s,i} \frac{\partial f}{\partial m_s}, \quad (i=1,\ 2,\ 3).$$

2° Pour

$$u_1 = u_2 = u_3 = 0,$$

et en vertu de la proposition (C), les fonctions ϕ_s et ψ_s devront se réduire respectivement à e_s et m_s.

Cherchons donc les expressions des ξ_{si} et des η_{si}. A cet effet posons:

(9)
$$T = \sum_{s=1}^{3} \xi_s \phi_s(e_1,\ \dots,\ m_1,\ \dots,\ u_1,\ \dots).$$

Les formules (7) nous donneront alors:

(10)
$$\xi_{s,i} = \left(\frac{\partial^2 T}{\partial \xi_s \partial u_i}\right)_{u_1=u_2=u_3=0.} \quad (s,\ i=1,2,3),$$

Pour aller plus loin, adoptons la convention suivante: un symbole de la forme F_i étant défini seulement pour les valeurs 1, 2 et 3 de l'indice i, nous considérerons le symbole F_k, où k représente un entier quelconque, comme défini par la formule:

$$F_k = F_i$$

où i représente l'entier déterminé par l'ensemble des relations:

$$i \equiv k \ (\text{mod. } 3), \quad (1 \leqq i \leqq 3).$$

Posons maintenant

$$n_i = e_{i+1}\ m_{i+2} - e_{i+2}\ m_{i+1}.$$

La fonction T, définie par la formule (9), étant, en vertu de la proposition (A), un invariant du groupe des rotations, il résulte des formules (10) que les $\xi_{s,i}$ sont les composantes d'un tenseur du second degré. Cela posé, on trouve:

(12)
$$\xi_{s,i} = e_s f_1(i) + m_s f_2(i) + n_s f_3(i)$$

en posant

(13)
$$f_k(i) = p_{k,1}\, e_i + p_{k,2}\, m_i + p_{k,3}\, n_i, \quad (i,\ k=1,\ 2,\ 3),$$

où les $p_{k,t}$ sont des invariants du groupe des rotations, fonctions des seules variables e_i et m_i $(i=1, 2, 3)$. Cela posé, j'observe que les fonctions α_1, α_2, α_3, définies par les formules:

$$\alpha_1=\sum_{k=1}^{3} e_k^2,\ \alpha_2=\sum_{k=1}^{3} m_k^2,\ \alpha_3=\sum_{k=1}^{3} e_k\, m_k,$$

constituent un système complet d'invariants indépendants des variables e_i et m_i par rapport au groupe des rotations. Il résulte de là que les $p_{k,t}$ pourront être regardés comme fonctions des seules variables α_1, α_2, α_3.

D'une façon tout-à-fait analogue on trouve:

$$\eta_{s,i}=e_s\phi_1(i)+m_s\phi_2(i)+n_s\phi_3(i), \tag{14}$$

en posant

$$\phi_k(i)=q_{k,1}\, e_i+q_{k,2}\, m_i+q_{k,3}\, n_i, \tag{15}$$

où les $q_{k,t}$ représentent des fonctions des seules variables α_1, α_2, α_3.

Pour résoudre notre problème dans toute sa généralité, il faudrait déterminer d'abord, de la façon la plus générale, les p et les q par la condition que le système (8) soit un système jacobien, ce qui exigerait l'intégration d'un système compliqué d'equations aux dérivées partielles du premier ordre, et passer ensuite à l'intégration du système (8) lui-même. Toutefois, le problème se simplifie beaucoup quand on veut se borner au degré de généralité qui semble suffisant au point de vue de la physique.

En effet, d'après les hypothèses généralement admises, la différence

$$e'-e,$$

où e' est défini par la première des formules (1), représente un vecteur perpendiculaire à la vitesse u. D'autre part, il semble naturel d'admettre qu'il en est de même du vecteur:

$$m'-m.$$

Or, les hypothèses précédentes étant adoptées, on trouve que l'on a

$$\xi_{i,k}+\xi_{k,i}=0,\ \eta_{ik}+\eta_{ki}=0,\quad (i,\ k=1, 2, 3). \tag{16}$$

Posons

$$\lambda_i(f)=\sum_{s=1}^{3} \xi_{s,i}\,\frac{\partial f}{\partial e_s}+\sum_{s=1}^{3} \eta_{s,i}\,\frac{\partial f}{\partial m_s}.$$

En vertu des équations (16) les conditions nécessaires et suffisantes pour que le système (8) soit un système jacobien prendront la forme suivante:

$$\lambda_i(\xi_{s,k})=\lambda_i(\eta_{s,k})=0,\quad (i,\ k,\ s=1, 2, 3). \tag{17}$$

Il résulte immédiatement de là que les formules (2) s'écriront comme il suit:

$$\left\{\begin{aligned} e_i' &=e_i+\sum_{k=1}^{3} \xi_{ik}u_k,\\ m_i' &=m_i+\sum_{k=1}^{3} \eta_{ik}\, u_k. \end{aligned}\right. \tag{18}$$

En effet, à cause des équations (17) les valeurs (18) des e_i' et des m_i' seront des intégrales communes des équations (8). D'autre part, pour

$$u_1 = u_2 = u_3 = 0$$

les formules (18) donnent

$$e_i' = e_i,\ m_i' = m_i,$$

et cela achève de prouver que les formules (18) sont bien les formules cherchées.

4. D'après ce qui précède, le problème se ramène à la détermination des ξ_{ik} et des η_{ik}.

Les conditions (16) donnent:

$$\begin{aligned} &p_{ii} = q_{ii} = 0, \\ &p_{i,i+1} + p_{i+1,i} = 0, \quad (i = 1, 2, 3), \\ &q_{i,i+1} + q_{i+1,i} = 0, \end{aligned}$$

et, en définitive, on trouve que les formules (12) et (14) peuvent être ramenées à la forme suivante:

$$(19) \qquad \begin{cases} \xi_{i,i} = \eta_{i,i} = 0, \\ -\xi_{i+1,i} = \xi_{i,i+1} = p_1 e_{i+2} + p_2 m_{i+2} + p_3 n_{i+2}, \\ -\eta_{i+1,i} = \eta_{i,i+1} = q_1 e_{i+2} + q_2 m_{i+2} + q_3 n_{i+2}, \end{cases}$$

où les p_k et les $q_k (k = 1, 2, 3)$ représentent des fonctions des seules variables α_1, α_2, α_3. Les équations qui déterminent ces fonctions s'obtiennent en développant les conditions (17). Pour écrire les équations précédentes, posons

$$(20) \qquad \begin{cases} \Delta = \alpha_1 \alpha_2 - \alpha_3^2, \\ M = q_1 \alpha_1 - p_2 \alpha_2 + (q_2 - p_1)\alpha_3, \\ A = p_2(p_1 + q_2),\ B = p_1^2 + p_2 q_1,\ A' = q_1(p_1 + q_2),\ B' = q_2^2 + p_2 q_1, \end{cases}$$

et définissons trois opérateurs différentiels μ_1, μ_2 et μ_3 par les formules:

$$\mu_1 = -2p_3\alpha_3 \frac{\partial}{\partial \alpha_1} - 2q_3\alpha_2 \frac{\partial}{\partial \alpha_2} - (p_3\alpha_2 + q_3\alpha_3)\frac{\partial}{\partial \alpha_3},$$

$$\mu_2 = 2p_3\alpha_1 \frac{\partial}{\partial \alpha_1} + 2q_3\alpha_3 \frac{\partial}{\partial \alpha_2} + (p_3\alpha_3 + q_3\alpha_1)\frac{\partial}{\partial \alpha_3},$$

$$\mu_3 = -2p_2 \frac{\partial}{\partial \alpha_1} + 2q_1 \frac{\partial}{\partial \alpha_2} + (p_1 - q_2)\frac{\partial}{\partial \alpha_3};$$

les équations demandées pourront alors se mettre sous la forme suivante:

$$\begin{aligned} &\Delta\{\mu_1(p_1) - p_3 q_1\} + \alpha_2 p_3 M = 0, \\ &\Delta\{\mu_2(p_1) + 2p_3 p_1 + q_3 p_2\} - \alpha_3 p_3 M = 0, \\ &\Delta\mu_3(p_1) - \alpha_2 A - \alpha_3 B = 0, \\ &\Delta\{\mu_1(p_2) - p_3 p_1 - p_3 q_2 - q_3 p_2\} - \alpha_3 p_3 M = 0, \\ &\Delta\{\mu_2(p_2) + p_3 p_2\} + \alpha_1 p_3 M = 0, \end{aligned}$$

$$\Delta\mu_3(p_2)+\alpha_3A+\alpha_1B=0,$$
$$\Delta\{\mu_1(p_3)-p_3q_3\}+\alpha_2A+\alpha_3B=0,$$
$$\Delta\{\mu_2(p_3)+p_3^2\}-\alpha_3A-\alpha_1B=0,$$
$$\Delta\mu_3(p_3)+p_3M=0,$$
$$\Delta\{\mu_1(q_1)-q_3q_1\}+\alpha_2q_3M=0,$$
$$\Delta\{\mu_2(q_1)+p_3q_1+q_3p_1+q_3q_2\}-\alpha_3q_3M=0,$$
$$\Delta\mu_3(q_1)-\alpha_3A'-\alpha_2B'=0\ ,$$
$$\Delta\{\mu_1(q_2)-p_3q_1-2q_3q_2\}-\alpha_3q_3M=0,$$
$$\Delta\{\mu_2(q_2)+q_3p_2\}+\alpha_1q_3M=0,$$
$$\Delta\mu_3(q_2)+\alpha_1A'+\alpha_3B'=0,$$
$$\Delta\{\mu_1(q_3)-q_3^2\}+\alpha_3A'+\alpha_2B'=0,$$
$$\Delta\{\mu_2(q_3)+q_3p_3\}-\alpha_1A'-\alpha_3B'=0,$$
$$\Delta\mu_3(q_3)+q_3M=0.$$

Sans insister, pour le moment, sur la discussion un peu laborieuse du système précédent et en réservant pour un autre travail quelques applications des résultats obtenus, je me bornerai actuellement à faire remarquer que, si l'on adopte l'une des hypothèses sur lesquelles M. Lorenz* a fondé la théorie des électrons, hypothèse qui revient à admettre que

$$p_1=p_3=0,\ p_2=0,$$

on trouve

$$q_1=q_2=q_3=0,$$

ce qui exprime que, par rapport au groupe défini par les équations (18), la force magnétique est un invariant.

*H. A. Lorenz. *The Theory of Electrons.* Leipzig, B. G. Teubner, 1909, p. 14, formule 23.

SUR LA NATURE DES GRANDEURS ÉLECTRIQUES CONSIDÉRÉES EN ÉLECTROSTATIQUE

PAR M. J. B. POMEY,
Directeur de l'École Supérieure de Télégraphie, Paris, France.

1. On dit d'ordinaire que le calcul vectoriel est le langage naturel de la Physique, mais il a l'inconvénient de supposer que l'on connait les propriétés métriques de l'espace. Or, on peut définir le champ électrique et le déplacement électrique sans faire appel à la métrique; il parait donc préférable d'employer les notations du calcul tensoriel.

2. Un vecteur est défini si on se donne son origine (a_1, a_2, a_3) et son extrémité (b_1, b_2, b_3); si l'on considère comme données seulement les différences $(b_1-a_1, b_2-a_2, b_3-a_3)$ on arrive à la notion de vecteur libre. Le vecteur libre déterminé par ses composantes suivant un système fondamental de vecteurs unités est un vecteur contrevariant.

De même, un plan peut être défini par l'équation

$$a_1x_1+a_2x_2+a_3x_3=d. \tag{1}$$

Mais si nous associons à ce plan un plan parallèle

$$a_1y_1+a_2y_2+a_3y_3=d+h, \tag{2}$$

nous formons, avec leur ensemble, un doublet. Par différence, nous obtenons:

$$a_1(y_1-x_1)+a_2(y_2-x_2)+a_3(y_3-x_3)=h.$$

Les différences $(y_1-x_1, y_2-x_2, y_3-x_3)$ sont les composantes d'un vecteur libre qui a son origine dans le plan (1) et son extrémité dans le plan (2), et l'on peut faire abstraction de la valeur particulière donnée à d. Le doublet peut alors être transporté à volonté, parallèlement à lui-même: il est caractérisé par le nombre h. L'équation

$$a_1z^1+a_2z^2+a_3z^3=h,$$

où l'on se donne h et où z^1, z^2, z^3 sont les composantes d'un vecteur libre contrevariant quelconque, allant d'un point quelconque du premier plan à un point quelconque du second, définit un tenseur du premier degré dont les composantes covariantes sont (a_1, a_2, a_3). La quantité h est supposée indépendante du choix du système de référence.

3. La notion de vecteur contrevariant se rattache à la notion de translations équipollentes, celle de tenseur covariant à une répartition de surfaces de niveau planes.

4. On peut définir le champ électrique par le travail nécessaire au déplacement élémentaire d'une charge donnée.

Trois expériences le long des chemins dx^1, dx^2, dx^3 suivant trois vecteurs unités fondamentaux donnent les travaux dT_1, dT_2, dT_3 et l'on a:

$$dT_1 = E_1 dx^1, \quad dT_2 = E_2 dx^2, \quad dT_3 = E_3 dx^3.$$

Si l'on déplace la charge suivant le chemin élémentaire ds qui a pour composantes contrevariantes dx^1, dx^2, dx^3, on a un travail dT et l'experience montre que l'on a:

$$dT = dT_1 + dT_2 + dT_3,$$

ou

$$dT = \Sigma E_i dx^i.$$

L'ensemble des expériences montre qu'en opérant avec une autre charge d'épreuve, les travaux seraient proportionnels. Ce rapport caractérise donc la charge d'épreuve, et, si on la suppose égale à l'unité, les coefficients E_1, E_2, E_3 caractériseront le champ en chaque point; dT joue le rôle de la quantité h et dx^1, dx^2, dx^3, celui des composantes z^1, z^2, z^3.

Donc, E_1, E_2, E_3 sont les composantes d'un tenseur covariant du premier degré, puisque dT est évidemment indépendant du système fondamental d'unités choisi au point considéré.

On voit que tant que la métrique n'est pas précisée, le champ électrique se présente comme un tenseur covariant.

5. On peut, de même, définir le déplacement de Maxwell par le phénomène de l'électrisation par influence. Exposons au champ un petit disque conducteur; il prendra, sur l'une de ses faces, une quantité d'électricité $+q$ et sur la face opposée la quantité d'électricité $-q$. Cette quantité est evidemment un invariant; elle ne dépend pas du choix du système fondamental de vecteurs unités. Pour la déterminer, il suffit de faire trois expériences en plaçant chaque fois le petit disque dans une des faces du trièdre de référence au point considéré. En opérant avec d'autres disques, on met en évidence par l'expérience ce fait que la charge induite par influence électrostatique est proportionnelle à la surface (Cf. N° 6, notion de volume).

Si le petit plan d'épreuve est un parallélogramme (S) construit sur les translations élémentaires dx^1, dx^2, dx^3; δx^1, δx^2, δx^3 on peut considérer successivement les parallélogrammes élémentaires situés dans les faces du trièdre de référence:

$$dx^2\delta x^3 - dx^3\delta x^2, \; dx^3\delta x^1 - dx^1\delta x^3, \; dx^1\delta x^2 - dx^2\delta x^1,$$

et l'on aura les quantités induites:

$$q_1 = D_{23}(dx^2\delta x^3 - dx^3\delta x^2),$$
$$q_2 = D_{31}(dx^3\delta x^1 - dx^1\delta x^3),$$
$$q_3 = D_{12}(dx^1\delta x^2 - dx^2\delta x^1).$$

Dès lors, l'expérience montre que pour le parallélogramme (S) on aura une charge q égale à $q_1 + q_2 + q_3$ soit

$$q = q_1 + q_2 + q_3$$

et, par suite,

$$q=\begin{vmatrix} D_{23} & D_{31} & D_{12} \\ dx^1 & dx^2 & dx^3 \\ \delta x^1 & \delta x^2 & \delta x^3 \end{vmatrix}.$$

De là résulte que le déplacement électrique, qui caractérise cette électricité induite, est entièrement défini par les trois coefficients D_{23}, D_{31}, D_{12}, et comme l'expression générale d'un tenseur du second degré est

$$q=\sum_i \sum_j D_{ij} dx^i \delta x^j,$$

expression qui comporte six coefficients dans le cas général, on en conclut qu'on a affaire à un tenseur dégénéré du second ordre.

6. Remarquons maintenant que si l'on considère le volume du parallélipipède construit sur trois vecteurs comme côtés, on peut, indépendamment de toute métrique, définir ce volume comme égal, à un coefficient près, à la fonction trilinéaire alternée

$$\begin{vmatrix} \xi^1 & \xi^2 & \xi^3 \\ \eta^1 & \eta^2 & \eta^3 \\ \zeta^1 & \zeta^2 & \zeta^3 \end{vmatrix},$$

où (ξ^1, ξ^2, ξ^3), (η^1, η^2, η^3), $(\zeta^1, \zeta^2, \zeta^3)$, sont respectivement les composantes contrevariantes des trois arêtes.

Le coefficient par lequel il convient de multiplier ce déterminant est d'ailleurs le volume construit sur les vecteurs unités du système fondamental choisi. Soit v ce coefficient. Le volume sera donc

$$v\begin{vmatrix} \xi^1 & \xi^2 & \xi^3 \\ \eta^1 & \eta^2 & \eta^3 \\ \zeta^1 & \zeta^2 & \zeta^3 \end{vmatrix}.$$

Il deviendra égal à q si je remplace

$$\begin{array}{rl} \eta^1,\ \eta^2,\ \eta^3 & \text{par } dx^1,\ dx^2,\ dx^3, \\ \zeta^1,\ \zeta^2,\ \zeta^3 & \text{par } \delta x^1,\ \delta x^2,\ \delta x^3, \\ v\xi^1,\ v\xi^2,\ v\xi^3 & \text{par } D_{23},\ D_{31},\ D_{12}. \end{array}$$

Comme le volume est un invariant et que q est aussi un invariant, les quantités

$$v\xi^1,\ v\xi^2,\ v\xi^3,$$

d'une part, et

$$D_{23},\ D_{31},\ D_{12},$$

d'autre part, seront soumises aux mêmes transformations.

On peut donc poser $D_{23}=vD^1$, $D_{31}=vD^2$, $D_{12}=vD^3$, (D^1, D^2, D^3) étant mis à la place de (ξ^1, ξ^2, ξ^3) pour représenter les composantes contrevariantes d'un vecteur.

La présence du facteur v indique que l'on doit considérer (D_{23}, D_{31}, D_{12}) comme équivalent à une densité tensorielle.

7. Il résulte de là que l'intensité de champ est un tenseur covariant et que le déplacement est une densité vectorielle. Ce sont donc des grandeurs d'espèces différentes. Pour cette raison, il y a lieu de donner des noms différents aux unités avec lesquelles on les mesure.

8. La relation vectorielle

$$D = E$$

qui est vraie en dehors de la matière ou des électrons, ne peut être exprimée qu'autant que l'on a fixé la métrique, car on peut alors passer des composantes covariantes aux composantes contrevariantes, et, par suite, mettre cette relation sous une forme invariante.

SUR QUELQUES APPLICATIONS DES ÉQUATIONS INTÉGRALES AU PROBLÈME DE L'HYSTÉRÉSIS MAGNÉTIQUE

PAR M. V. A. KOSTITZIN,
Professeur à l'Université de Moscou, Moscou, Russie.

1. Pour expliquer le phénomène de l'induction magnétique, nous avons des théories moléculaires en vertu desquelles chaque particule de matière est considérée comme un aimant élémentaire assujetti aux forces de liaison et sur la nature desquelles on peut former des hypothèses très variées. En désignant par $x(t)$ l'intensité du champ magnétisant, par $y(t)$ l'intensité d'aimantation induite, et par t le temps, on trouve une relation de la forme

$$y = MF(x) \tag{1}$$

où $F(x)$ est une fonction impaire avec les propriétés suivantes:

$$\begin{cases} F'(x) > 0, \; F''(x) \leq 0, \\ F(0) = 0, \; F(\infty) = 1. \end{cases} \tag{2}$$

Si l'on applique la théorie bien connue de Weber, on trouve

$$\begin{cases} F(x) = \dfrac{2}{3}\dfrac{x}{D}, & (x \leq D), \\ F(x) = 1 - \dfrac{1}{3}\dfrac{D^2}{x^2}, & (x \geq D), \end{cases} \tag{3}$$

tandis que, d'après M. Langevin, on a:

$$F(x) = \frac{e^{\alpha x} + e^{-\alpha x}}{e^{\alpha x} - e^{-\alpha x}} - \frac{1}{\alpha x}. \tag{4}$$

Les points de départ des théories de Weber et de M. Langevin sont tout à fait différents, mais l'allure générale des courbes définies par les équations (3) et (4) est la même, et les conditions (2) sont vérifiées dans les deux cas. Donc, il nous est permis de supposer que la fonction $F(x)$, dans tous les cas, vérifie les conditions (2). Or tous les faits observés ne sont expliqués, ni par la formule (3) ni par la formule (4). Si l'on suppose, par exemple, que $x(t)$ croisse indéfiniment avec le temps, la courbe expérimentale représentant y en fonction de x, a un point d'inflexion, et, tandis que $y''(x)$ est positive pour les valeurs très petites de x, au contraire, elle est négative pour les valeurs assez grandes de x, ce que les théories de Weber et de M. Langevin ne permettent pas de prévoir; et le phénomène d'hystérésis reste inexpliqué.

2. On peut obtenir une formule plus souple et plus proche de la réalité en introduisant dans la formule (1) un terme complémentaire. En effet, il est très probable que le magnétisme induit ne disparait pas en même temps que le champ magnétisant. On peut alors remplacer l'équation (1) par l'équation suivante:

$$y(t) = MF[x(t)] + \lambda \int_{t_0}^{t} k(t-u) y(u) du, \tag{5}$$

en désignant par λk le coefficient d'hérédité. La fonction k est une fonction positive et décroissante; $k(0)=1$, et la constante λ, étant positive, ne dépasse pas l'unité. En résolvant l'équation (5) on a:

$$y(t) = MF[x(t)] + \lambda M \int_{t_0}^{t} K(\lambda;\ t-u) F[x(u)] du. \tag{6}$$

Le noyau résolvant $K(\lambda; u)$ est aussi une fonction positive et décroissante de u, égale à *un* pour $u=0$.

3. L'équation (6) nous permet de faire l'étude d'un cas important. Supposons que:

$$x(t) = \beta(t-t_0).$$

On a:

$$\left\{ \begin{aligned} y(t) &= MF[\beta(t-t_0)] + \lambda M \int_0^{t-t_0} K(\lambda;\ u) F[\beta(t-t_0-u)] du, \\ y'(t) &= M\beta F'[\beta(t-t_0)] + \lambda M\beta \int_0^{t-t_0} K(\lambda;\ u) F'[\beta(t-t_0-u)] du, \\ y''(t) &= M\beta^2 F''[\beta(t-t_0)] \\ &\quad + \lambda M\beta^2 \int_0^{t-t_0} K(\lambda;\ u) F''[\beta(t-t_0-u)] du + \lambda M\beta K(\lambda;\ t-t_0) F'(0). \end{aligned} \right. \tag{7}$$

Si l'on suppose que l'intégrale

$$\int_0^{t-t_0} K(\lambda;\ u) du$$

a une limite finie quand $t \to \infty$, il est facile de faire voir que $y(t)$ a aussi une limite finie et que $y''(t)$ est négative pour les valeurs de t suffisamment grandes. Mais si l'on pose $t=t_0$, on a:

$$y''(t_0) = M\beta^2 F''(0) + \lambda M\beta F'(0).$$

Or, dans les deux hypothèses que nous avons envisagées, l'hypothèse de Weber et celle de M. Langevin, on a toujours:

$$F''(0) = 0,\ F'(0) > 0.$$

Nous voyons donc que:

$$y''(t_0) > 0$$

et que par conséquent, l'existence du point d'inflexion découle bien de notre hypothèse.

4. Il y a encore un fait expérimental dont notre hypothèse nous permet de rendre compte. Si le champ magnétisant disparait d'une manière brusque, l'aimantation induite résiduelle est beaucoup plus petite que dans le cas où le champ magnétique tend vers zéro d'une manière continue.

En effet, posons, dans le premier cas:

$$\begin{cases} x(t)=\beta(t-t_0), & \text{pour } t_0<t<T, \\ x(t)=0 & \text{pour } T<t, \end{cases} \tag{8}$$

et dans le deuxième cas:

$$\begin{cases} x(t)=\beta(t-t_0) & \text{pour } t_0<t<T, \\ x(t)=\dfrac{\beta(T-t_0)(T_1-t)}{T_1-T} & \text{pour } T<t<T_1, \\ x(t)=0 & \text{pour } T_1<t. \end{cases} \tag{9}$$

Dans le premier cas, on a pour $t>T_1$:

$$y_1(t)=\lambda M\int_{t_0}^{T} K(\lambda;\, t-s)F[\beta(s-t_0)]ds,$$

et dans le second cas:

$$y_2(t)=\lambda M\int_{t_0}^{T} K(\lambda;\, t-s)F[\beta(s-t_0)]ds+\lambda M\int_{T}^{T_1} K(\lambda;\, t-s)F\left[\frac{\beta(T-t_0)(T_1-s)}{T_1-T}\right]ds.$$

La comparaison est facile, et il est évident que la différence

$$y_2(t)-y_1(t)=\lambda M\int_{T}^{T_1} K(\lambda;\, t-s)F\left[\frac{\beta(T-t_0)(T_1-s)}{T_1-T}\right]ds$$

est positive.

5. Il nous est possible maintenant de passer à l'étude du cas beaucoup plus important où $x(t)$ est une fonction périodique du temps. Supposons que $t_0=0$, et que

$$4hn\leq t\leq 4h(n+1),$$

$4h$ étant la valeur de la période. En désignant par u la différence $t-4nh$ et par $Z_n(u)$ la valeur de la fonction $y(4nh+u)$, on trouve:

$$Z_n(u)=MF[x(u)]+\lambda M\int_0^{4h} S_n(\lambda;\, u-s)F[x(s)]ds+\lambda M\int_0^{u} K(\lambda;\, u-s)F[x(s)]ds,$$

où

$$S_n(\lambda;\, u-s)=\sum_{k=1}^{n} K(\lambda;\, 4hn+u-s).$$

Quand $n\to\infty$ la limite de $S_n(\lambda;\, u-s)$ existe bien, et en la désignant par

$$S(\lambda;\, u-s)=\sum_{k=1}^{\infty} K(\lambda;\, 4hn+u-s),$$

nous obtenons pour le cycle limite:

$$(10)\qquad Z(u)=\lim_{n\to\infty} Z_n(u)=MF[x(u)]+\lambda M\int_0^{4h} S(\lambda;\, u-s)F[x(s)]ds$$

$$+\lambda M\int_0^{u} K(\lambda;\, u-s)F[x(s)]ds.$$

La fonction $S(\lambda;\, u-s)$ vérifie la relation suivante:

$$(11)\qquad S(\lambda;\, u-s+4h)=S(\lambda;\, u-s)-K(\lambda;\, u-s+4h).$$

Si l'on suppose pour faciliter les calculs que le noyau K est une fonction exponentielle

$$K(u-s)=e^{-\mu(u-s)},$$

on a

$$K(\lambda;\, u-s)=e^{-(\mu-\lambda)(u-s)},$$

$$S(\lambda;\, u-s)=\frac{e^{-(\mu-\lambda)(u-s)-4\lambda(\mu-\lambda)}}{1-e^{-4\lambda(\mu-\lambda)}}.$$

Il faut ensuite faire des hypothèses complémentaires sur la fonction $x(u)$ qui dépend de la technique expérimentale. Si l'on suppose, par exemple, que

$$x(u)=A\,\frac{u-h}{h},\quad 0\leq u\leq 2h,$$

$$x(u)=A\,\frac{3h-u}{h},\quad 2h\leq u\leq 4h,$$

l'équation (10) prend une forme particulièrement simple, et il est facile de faire voir que la courbe représentant y en fonction de x reproduit très bien la courbe expérimentale d'hystérésis. Pour expliquer par cette théorie toutes les particularités des courbes d'hystérésis, il y a toujours à notre disposition le coup de pouce classique. Malheureusement, la place me manque ici pour développer tous les calculs et pour plus de détails, je renvoie à mon mémoire qui va paraître au fascicule 3 - 4 du t. XXXI du Recueil Mathématique de la Société Mathématique de Moscou.

ALTERNATING CURRENT DISTRIBUTION IN CYLINDRICAL CONDUCTORS

By Dr. Chester Snow,
Physicist, Bureau of Standards, Washington, D.C., U.S.A.

The fundamental equations of Maxwell for the scalar and vector potentials ϕ and A are simplified by an approximation which is valid from the lowest to the highest (radio) frequencies. This approximation is due to the difference in the order of magnitude of the electrical conductivity of dielectrics and conductors. The case of N conducting cylindrical groups surrounded by any number of dielectrics is discussed and forms of solution obtained for ϕ and A which are proportional to $e^{ipi-\gamma z}$ representing a simple possible type of propagation along the cylinders. These lead to general formulae for the coefficients of leakage, capacity, resistance, and inductance (all of which are functions of the frequency in the general case) as well as expressions for the attenuation and phase velocity. The mean energy relations are also expressed in terms of these coefficients. Application is made to the case of two circular cylinders of different conductivity, permeability, and radius, surrounded by a homogeneous, slightly conductive, dielectric. Asymptotic formulae for the alternating current resistance R and inductance L of the line at high frequencies are obtained, which together with the exact expressions for the coefficients of leakage and capacity lead to high frequency expressions for the attentuation and phase velocity. These are all functions of the frequency due to the fact that the current distribution is not uniform in the conductors.

I. Introduction—Statement of Problem

A knowledge of the mode of propagation and distribution of alternating current in a cylindrical system is important, for such currents are used in practice, and, moreover, the propagation of an arbitrary type of wave, or "transient", may theoretically be found in terms of the periodic solution by the use of Fourier's integral or by complex integration*. In working out such distinct problems as that of propagation along parallel wires where the dielectric extends to infinity, and along a cable system where a conductor extends to infinity, a number of formal similarities become evident. It may be difficult to see just how general are these similarities if the initial formulation of the problem is not sufficiently general. Some of the approximations to be made are common to all problems of this type. Moreover, the connections between the field vectors and such physical concepts as resistance, inductance, and capacity may be made in a

*Thornton C. Fry., Phys. Rev. Aug. 1919., p. 115.

manner which is quite inclusive. In this paper a discussion of the general problem is undertaken in the hope that it may prove useful in outlining the procedure or forecasting the results in particular problems of this type.

II. Fundamental Equations

1. *Type of Waves. The Propagation Constant* γ.

If all the conductors and dielectrics have their generating lines parallel to the z axis, the electric vector may be represented by the real part of $Ee^{ipt-\gamma z}$ and the magnetic induction by the real part of $Be^{ipt-\gamma z}$ where p is 2π times the frequency, and γ is the complex propagation constant which may be written

$$\gamma = b + \frac{ip}{V}. \tag{1}$$

The real quantities b and V represent the attenuation constant and phase velocity, respectively, of this type of wave. There will be a finite number of types of waves or values of γ possible for a given system, and we shall first consider the case where only one type is present. It will be more convenient to deal with the complex scalar and vector potentials ϕ and A respectively, and to derive E and B from them by the relations

$$\begin{cases} E = -\nabla\phi - ipA, \\ B = \text{curl } A = \mu H. \end{cases} \tag{2}$$

The electromagnetic c.g.s. system of units will be used throughout with the exception of the dielectric constant k which will be taken in electrostatic c.g.s. units. The electrical conductivity being λ, and the ratio of the units being $c = 3\times 10^{10}$, the Maxwell equations

$$\begin{cases} \text{curl } B = 4\pi\mu\left(\lambda + \dfrac{ipk}{4\pi c^2}\right)E, & \text{Div } B = 0, \\ \text{curl } E = -ipB, & \text{Div } E = 0, \end{cases} \tag{3}$$

require that

$$\begin{cases} \nabla^2 A_x + h^2 A_x = \dfrac{\partial \chi}{\partial x}, \\ \nabla^2 A_y + h^2 A_y = \dfrac{\partial \chi}{\partial y}, \\ \nabla^2 A_z + h^2 A_z = \dfrac{\partial \chi}{\partial z} = -\gamma\chi, \\ \nabla^2 \phi + h^2 \phi = -\dfrac{\partial \chi}{\partial t} = -ip\chi, \end{cases} \quad \begin{cases} \text{where } \nabla^2 \equiv \dfrac{\partial^2}{\partial x^2} + \dfrac{\partial^2}{\partial y^2}, \\ h^2 = \gamma^2 - 4\pi i p\mu\left(\lambda + \dfrac{ipk}{4\pi c^2}\right), \\ \chi = \dfrac{\partial A_x}{\partial x} + \dfrac{\partial A_y}{\partial y} - \gamma A_z + \dfrac{\gamma^2 - h^2}{ip}\phi. \end{cases} \tag{4}$$

Conversely any pair ϕ and A which satisfy these equations (4) and give continuity to the tangential components of E and H will give the correct field vectors.

2. *General Approximations.*

It will be assumed that γ is a small quantity of the order of $\frac{1}{c}$ (first order). In all materials k is of the order of unity, and hence the complex conductivity a defined by

$$a \equiv \lambda + \frac{ipk}{4\pi c^2}, \tag{5}$$

will be very approximately equal to λ in conductors, that is, a finite magnitude (for copper $\lambda = .0006$).

In dielectrics λ is of the same order, or even smaller than the term $\frac{ipk}{4\pi c^2}$. Thus for gutta-percha $\lambda = 10^{-24}$ and $k = 4$ so that $\frac{pk}{4\pi c^2} = \frac{2f}{10^{21}}$. It is therefore evident that for a range of frequency from one cycle up to the highest, λ will not be greater than the second order term $\frac{pk}{4\pi c^2}$ even if the dielectric is ten thousand times a better conductor than gutta percha, and consequently the complex conductivity will be a second order infinitesimal in all dielectrics. For the sake of the most general results we shall assume that λ may be of the same order as $\frac{pk}{4\pi c^2}$ or γ^2. The magnetic permeability μ will have the value 1 in all dielectrics and in non-magnetic metals.

The cause of the wave is assumed to be in certain electromotive forces applied to the terminals of the conductor at say $z = 0$ and $z = l$ where the cylindrical conductors are connected by networks of known impedances. We shall assume that there is nothing analogous to these applied electromotive forces in a magnetic sense; no applied "magnetomotive forces" tending to magnetize the cylinders in the z direction. Hence the component of magnetic field in this direction H_z will be everywhere infinitesimal since it could only be produced by the x,y components of current which will be infinitesimal in the dielectrics and also in the conductors since the normal component of current in conductors must be continuous with its value outside. Hence A_x and A_y will be negligible. We may therefore understand A to mean the z component A_z in all that follows since the others are negligible.

The first approximation to the solution will therefore lead to

$$\left\{\begin{aligned} E_x &= -\frac{\partial\phi}{\partial x}, & B_x &= \frac{\partial A}{\partial y}, \\ E_y &= -\frac{\partial\phi}{\partial y}, & B_y &= -\frac{\partial A}{\partial x}, \\ E_z = E &= -\frac{\partial\phi}{\partial z} - ipA_z = \gamma\phi - ipA, & B_z &= 0, \end{aligned}\right. \tag{6}$$

which hold at all points in the xy plane.

3. *Differential Equations and Boundary Conditions for ϕ and A.*

The electric component E_z (which will be designated by E from now on) is of finite order of magnitude everywhere and is continuous at all boundaries. Since $E=\gamma\phi-ipA$, it follows that A and $\gamma\phi$ are also finite. If ϕ and A are continuous at all boundaries, the continuity of E will be assured. Since E_x and E_y are negligible in conductors, ϕ must have a constant value at all points in each conducting section. Let $\phi=c_n$ at all points in the plane section $S_n(z=\text{constant})$ of the nth conductor or group of conductors in contact. Suppose there are N such groups. The tangential component of the electric field in the xy plane will be zero (and hence continuous) at all boundaries between a dielectric and conductor or between two conductors. It will also be continuous at boundaries between two different dielectrics if ϕ is continuous there. At such surfaces the conservation of electricity, or solenoidal property of the total electric current, requires that $a\frac{\partial\phi}{\partial n}$ be continuous. Since a is a second order quantity, the last of the equations (4) reduces to $\nabla^2\phi=0$. Similarly at the boundary between two materials of different magnetic permeability, the continuity of H_s demands that

$$\frac{1}{\mu}\frac{\partial H}{\partial n}=-H_s$$

be continuous. The continuity of B_n is assured by the continuity of A since $B_n=\frac{\partial A}{\partial s}$.

Therefore the differential equations which ϕ and A must satisfy are:

$$(7)\quad\begin{cases}(a)\ \nabla^2\phi=0 \text{ everywhere,}\\(b)\ \phi=c_n \text{ on the section } S_n \text{ of the } n\text{th conducting group,}\\(c)\ \phi \text{ is continuous everywhere,}\\(d)\ a\dfrac{\partial\phi}{\partial n} \text{ is continuous at the boundaries between two different dielectrics,}\\(e)\ \lim\limits_{r=\infty}\left(r^2\dfrac{\partial\phi}{\partial r}\right) \text{ exists for all directions of } r;\end{cases}$$

$$(8)\quad\begin{cases}(a)\ \nabla^2A=0 \text{ in dielectrics,}\\ \qquad\quad =4\pi ip\mu\lambda\left(A-\dfrac{\gamma\phi}{ip}\right) \text{ in conductors,}\\(b)\ A \text{ is continuous at all boundaries,}\\(c)\ \dfrac{1}{\mu}\dfrac{\partial A}{\partial n} \text{ is continuous at all boundaries,}\\(d)\ \lim\limits_{r=\infty}\left(r^2\dfrac{\partial A}{\partial r}\right) \text{ exists for all directions of } r.\end{cases}$$

The relation (8*a*) follows from (4). The conditions at infinity (7*e*) and (8*d*) are derived as follows:

Let I_n denote the total z component of current through the section S_n of the nth conductor (or nth group of conductors in contact).

The set of closed contours or artificial boundaries, shown by dotted lines in Fig. 1 (see p. 202), are drawn so that each artificial boundary encloses a homogeneous material. In the case of materials extending to infinity in the xy plane,, the artificial boundary is closed by arcs of a circle with centre at some finite point, and indefinitely large radius. The natural boundaries are shown by full lines. The normal to the boundary is taken in each medium pointing toward the boundary, and the positive direction of an element ds of an artificial boundary curve is such that in going around this contour in the positive direction the homogeneous medium encircled is on the left hand. Thus the directions of n and ds are related like those of the x and y axes respectively. With this understanding, the statement of the conservation of electricity applied to the nth conductor is

$$\gamma I_n = \int a \frac{\partial \phi}{\partial n} ds_n, \quad (n=1, 2, 3, \ldots\ldots, N), \tag{9}$$

where the integral is taken around any closed curve in the dielectric which encircles this conductor only, the normal n pointing toward the conductor.

Since ϕ and $a \frac{\partial \phi}{\partial n}$ are continuous and $\nabla^2 \phi = 0$ in every dielectric, it follows that $\int a \frac{\partial \phi}{\partial n} ds = 0$ taken over any closed curve which does not encircle a conductor. Consequently, by (9),

$$\gamma \sum_1^N I_n = \sum_{n=1}^N \int a \frac{\partial \phi}{\partial n} ds_n = -\operatorname*{limit}_{r=\infty} r \int_0^{2\pi} a \frac{\partial \phi}{\partial r} d\theta,$$

where r is the radius of a circle with some finite point taken as centre.

Since the terminal apparatus located in finite portions of the planes $z=0$, and $z=l$ are assumed to have no mutual capacity or coupling with any of these cylinders, it follows that the sum of the currents entering either end must be zero at every instant, so that

$$\sum_{n=1}^N I_n = 0. \tag{10}$$

Hence the $\operatorname*{limit}_{r=\infty} r \frac{\partial \phi}{\partial r}$ must be zero. But if we assume that the density of surface change on the natural boundaries is finite and falls off so rapidly as the point on the boundary moves off to infinity that the integral which defines its logarithmic potential is convergent, then it follows that $\frac{\partial \phi}{\partial r}$ must vanish in such a manner that $\operatorname*{limit}_{r=\infty} r^2 \frac{\partial \phi}{\partial r}$ exists which is the condition (7e).

Similarly, it is assumed that the density of current (surface densities of molecular currents on boundaries, or magnetic surface changes) are so distributed that their logarithmic potential exists, and if $r\frac{\partial A}{\partial r}$ vanishes it must vanish canonically. The definition of current takes the form

$$I_n = \frac{1}{4\pi}\int \frac{\partial A}{\partial n}\, ds_n, \tag{11}$$

the integral being taken in the dielectric as in (9). A similar reasoning leads to the condition (8*d*).

In the next two sections it will appear that the conditions (7) and (8) are necessary and sufficient to uniquely determine ϕ and A in terms of the potentials $c_1, c_2, \ldots, c_n$. These solutions of (7) and (8) will make $\sum_1^N I_n = 0$, which is reciprocal with the requirement that $r\frac{\partial \phi}{\partial r}$ and $r\frac{\partial A}{\partial r}$ shall vanish at infinity and in such a manner that $r^2\frac{\partial \phi}{\partial r}$ and $r^2\frac{\partial A}{\partial r}$ shall exist. If no conductor extends to infinity then ϕ and A are harmonic and the values of $\phi(\infty)$ and $A(\infty)$ will be determinate since one cannot arbitrarily assign either the value of the harmonic function or of its normal derivative at infinity (if the c's are all arbitrary). In the case where a conductor, say the Nth, extends to infinity then A and E will not be harmonic in this conductor but will satisfy

$$\nabla^2 A - 4\pi i p a A = 0 = \nabla^2 E - 4\pi i p a E,$$

since

$$\phi(x_N, y_N) - \phi(\infty) = 1 - 1 = 0$$

if x_N, y_N is on this conductor. Consequently the vanishing of $r\frac{\partial A}{\partial r}$ $\left(\text{and hence } r\frac{\partial E}{\partial r}\right)$ brings with it the fact that both A and E must vanish at infinity.

Whether either of the finite constants $\phi(\infty)$ or $A(\infty)$ have a value zero is a question without physical significance since the x and y components of electric and magnetic field will vanish properly at infinity. But the linear combination $\gamma\phi(\infty) - ipA(\infty)$ has the physical significance $E(\infty)$ $[=E_z(\infty)]$. Consequently we must conclude that for the type of wave assumed, the first approximations here attempted will give a definite value to $E(\infty)$ which cannot be arbitrarily assigned the value zero. As a matter of fact, it will be found that $E(\infty)$ will not have the value zero in the case where one conductor of finite section surrounds all the others. Even in this case the formal solution we shall obtain will give the first approximation (*i.e.*, the finite terms) to the current distribution in conductors. A further approximation would not appreciably affect the value of E in conductors, but would give a function E_z which is not strictly harmonic in the dielectric and hence which vanishes at infinity since $\frac{\partial E}{\partial r}$ must do so.

In case more than one conductor extends to infinity we shall assume that they are at the same potential and may be treated as one group of conductors in contact. This avoids the case where two conductors have infinite coefficients of capacity or leakage per unit length.

III. The Complex Scalar Potential ϕ

1. *Uniqueness of a Solution.*

If the complex values $c_1, c_2, \ldots, c_N$ which ϕ must assume on the conductors are arbitrarily assigned, there cannot be more than one function ϕ satisfying the conditions (7) of Section III. If there were another, their difference $\phi'+i\phi''$ would satisfy all these conditions and vanish on each conductor, where ϕ' and ϕ'' are real. The boundary conditions (7d) imply the two real relations

$$\lambda_1 \frac{\partial \phi'}{\partial n_1} + \lambda_2 \frac{\partial \phi'}{\partial n_2} = \frac{p}{4\pi c^2}\left(k_1 \frac{\partial \phi''}{\partial n_1} + k_2 \frac{\partial \phi''}{\partial n_2}\right),$$

$$\lambda_1 \frac{\partial \phi''}{\partial n_1} + \lambda_2 \frac{\partial \phi''}{\partial n_2} = -\frac{p}{4\pi c^2}\left(k_1 \frac{\partial \phi'}{\partial n_1} + k_2 \frac{\partial \phi'}{\partial n_2}\right).$$

Also ϕ' and ϕ'' are harmonic within each dielectric section S_j and

$$\underset{r=\infty}{\text{limit}} \int_0^{2\pi} \frac{\partial \phi}{\partial r} r d\theta = \underset{r=\infty}{\text{limit}} \int_0^{2\pi} \frac{\partial \phi}{\partial r} r d\theta = 0$$

and

$$\int \left(\phi' \frac{\partial \phi''}{\partial n} - \phi'' \frac{\partial \phi'}{\partial n}\right) ds_j = 0$$

taken around the artificial boundary of S_j or only around that part of it which is adjacent to another dielectric, since ϕ' and ϕ'' vanish at conducting boundaries and the integral vanishes for arcs of the circle at infinity.

By the use of Green's theorem it is easy to show that

$$\iint \lambda \left[\left(\frac{\partial \phi'}{\partial x}\right)^2 + \left(\frac{\partial \phi'}{\partial y}\right)^2 + \left(\frac{\partial \phi''}{\partial x}\right)^2 + \left(\frac{\partial \phi''}{\partial y}\right)^2\right] dS_j = \int \lambda \left[\phi' \frac{\partial \phi'}{\partial n} + \phi'' \frac{\partial \phi''}{\partial n}\right] ds_j.$$

Summing all such equations for the entire dielectric region of the xy plane gives

$$\iint \lambda \left[\left(\frac{\partial \phi'}{\partial x}\right)^2 + \left(\frac{\partial \phi'}{\partial y}\right)^2 + \left(\frac{\partial \phi''}{\partial x}\right)^2 + \left(\frac{\partial \phi''}{\partial y}\right)^2\right] ds = \sum_j \int \lambda \left[\phi' \frac{\partial \phi'}{\partial n} + \phi'' \frac{\partial \phi''}{\partial n}\right] ds_j,$$

where the double integral on the left extends over the entire dielectric region, the line integrals on the right being taken around all the closed artificial boundaries of all the dielectrics. The terms corresponding to the conducting part of these boundaries vanish because ϕ' and ϕ'' are zero there. The terms corresponding to the boundaries between two dielectrics may be written by virtue of the above boundary conditions in the form

$$\frac{p}{4\pi c^2} \int k \left(\phi' \frac{\partial \phi''}{\partial n} - \phi'' \frac{\partial \phi'}{\partial n}\right) ds$$

over both sides of all the artificial boundaries between two dielectrics, and this has been shown to be zero. Hence ϕ' and ϕ'' must be constant over the entire dielectric region since they are continuous there. They vanish on conductors and hence $\phi' = \phi'' = 0$, which shows that there cannot be more than one funcion ϕ which satisfies the conditions (7) of Section II.

2. *Existence and Uniqueness of a Generalized Green's Function G.*

To find whether it is possible for any function ϕ to exist which satisfies the condition (7), and if so to obtain an integral representation of ϕ, we may make use of a generalization of Green's function $G(xy, \xi\eta)$ which may be regarded as the potential at any point x, y in the dielectric region, when there is a line source of unit strength (per unit length along z) which is parallel to z, at some fixed point ξ, η in the dielectric region, and when all the conductors are kept at zero potential and all together receive the unit current from the line source.

The properties of this function are

$$(a) \qquad \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right) G(xy, \xi\eta) = 0$$

where x, y is a point within any dielectric section.

$$(b) \qquad G(xy, \xi\eta) \text{ and } a(xy) \frac{\partial}{\partial n} G(xy, \xi\eta)$$

change continuously when the point x, y moves across the boundary line between two dielectrics.

(*c*) $G(xy, \xi\eta)$ vanishes when the point x, y moves up to a conducting boundary of the dielectric region.

(*d*) When the point x, y approaches the fixed point ξ, η, $G(xy, \xi\eta)$ becomes infinite in such a manner that it differs from $-\dfrac{2 \log r(xy, \xi\eta)}{4\pi a(\xi\eta)}$ by a finite quantity. The strength of the source is $\displaystyle\int a \frac{\partial G(xy, \xi\eta)}{\partial n} ds$ taken around an infinitesimal circle with centre at ξ, η, and is thus equal to 1, if n is drawn toward the circle, on the outside. This represents the efflux of electricity from the line source per second per unit length along z (leakage and displacement current together).

(*e*) When the dielectric region extends to infinity, if x,y moves off to infinity in any direction in the x,y plane (ξ, η being a fixed finite point), then $\dfrac{\partial G}{\partial r}$ vanishes (so that the $\displaystyle\lim_{r=\infty} r^2 \frac{\partial G(xy, \xi\eta)}{\partial r}$ exists). Hence G takes on the asymptotic form $G(xy, \xi\eta) = G(\infty, \xi\eta) + \text{terms in } \dfrac{1}{r}$ and higher. The constant $G(\infty, \xi\eta)$ is calculable, not assignable. From these properties it is easily shown by applying Green's theorem to the two functions $G(xy, x'y')$ and $G(\xi\eta, x'y')$, where x,y and ξ, η are any two distinct points whatever in the dielectric region, that this function is a symmetrical one of the two points, *i.e.*,

$$G(xy,\xi\eta)=G(\xi\eta,xy). \tag{12}$$

This means that the potential at x,y due to a line charge at ξ,η will have the same value as the potential at ξ,η due to a line charge at x,y. From this interpretation it is evident that $\left(\frac{\partial^2}{\partial\xi^2}+\frac{\partial^2}{\partial\eta^2}\right)G=0$ and that if x,y is a fixed point and ξ,η the variable, then G and $a\frac{\partial G}{\partial n}$ will be continuous.

If we apply the integral transformation

$$\iint a\nabla^2 G dS_j=\int a\frac{\partial G}{\partial n}ds_j$$

to each homogeneous dielectric section S_j and sum for all of them, excluding the point ξ,η by an infinitesimal circle, we find that

$$\sum_{k=1}^{N}\int a\frac{\partial G(x_k y_k,\xi\eta)}{\partial n_k}ds_k=-1, \tag{13}$$

where the kth integral is taken around any closed curve in the dielectric which encircles the kth conducting group only. This merely states the fact that the total flow from the line source goes into the conductors and there is no flow to infinity.

To prove the existence and uniqueness of such a function G having the desired properties, we may begin by assuming that a Green's function $g(xy,\xi\eta)$ has been constructed for the entire dielectric region as if it were homogeneous. In the case of a closed finite region, this is usually effected by writing

$$g(xy,\xi\eta)=-\frac{1}{2\pi}\log r(xy,\xi\eta)+v(xy,\xi\eta)$$

and choosing $v(xy,\xi\eta)$ as a function which satisfies $\frac{\partial^2 v}{\partial x^2}+\frac{\partial^2 v}{\partial y^2}=0$ in the region, and assumes the value $\frac{1}{2\pi}\log r(x'y',\xi\eta)$ when x,y approaches the point x',y' on the boundary, thus making $g(xy,\xi\eta)$ vanish when x,y approaches the boundary point x',y'. The existence and uniqueness of v is therefore a special case of the more general problem of Dirichlet of finding a harmonic function which assumes any assigned boundary values. Since the development of the theory of Fredholm's integral equation it may be taken as satisfactorily proven that a harmonic function exists and is unique if either

(*a*) its value is assigned on all points of the boundary,

or

(*b*) its value is assigned on some parts of the boundary, and the values of its normal derivative on all remaining parts*.

In the case of an unbounded region, the value of $g(xy,\xi\eta)$ when x,y goes to infinity is not prescribed, but the condition is imposed that $\lim\limits_{r=\infty} r^2 \dfrac{\partial\phi}{\partial r}$ shall exist. This together with the facts that g shall be harmonic in x,y and vanish on all the internal boundaries of the region and become logarithmically infinite when x,y approaches ξ,η serves to uniquely determine g and hence determines the value which it will assume at infinity.

If we apply Green's theorem to the two functions $a(xy)G(xy,\xi_1\eta_1)$ and $a(xy)g(xy,\xi\eta)$, the point x,y being the variable of integration, we find, if ξ_1,η_1 is within the homogeneous dielectric section S_j and ξ,η in S_k:

$$-g(\xi_1\eta_1,\xi\eta) = -\int a\left[G(xy,\xi_1\eta_1)\frac{\partial g(xy,\xi\eta)}{\partial n} - g(xy,\xi\eta)\frac{\partial G(xy,\xi_1\eta_1)}{\partial n}\right]ds_j,$$

$$a(\xi\eta)G(\xi\eta,\xi_1\eta_1) = -\int a\left[G(xy,\xi_1\eta_1)\frac{\partial g(xy,\xi\eta)}{\partial n} - g(xy,\xi\eta)\frac{\partial G(xy,\xi_1\eta_1)}{\partial n}\right]ds_k,$$

the integrals being taken around the artificial contours of the corresponding dielectric section. The contour integral on the right has the value zero if taken around any other artificial boundary of a dielectric. Hence, adding all such terms to cover the entire dielectric region we get

$$a(\xi\eta)G(\xi\eta,\xi_1\eta_1) + \int [a_1(s)-a_2(s)]\,\frac{\partial g(s,\xi\eta)}{\partial n_1}G(s,\xi_1\eta_1)ds = g(\xi_1\eta_1,\xi\eta) \tag{14}$$

where the line integral is taken once over all the boundaries between different dielectrics, the point $s(xy)$ being the point of integration on this curve. The values $a_1(s)$ and $a_2(s)$ correspond to the medium on the left and right of the curve respectively at the point s. The positive direction of ds may be arbitrary but the direction of the normal n must be such that n_1 and ds_1 are related to each other like the directions of the x and y axes respectively.

The continuity of G, $a\dfrac{\partial G}{\partial n}$, g and $\dfrac{\partial g}{\partial n}$ have been used in obtaining this equation. (The point $s(xy)$ on the boundary s is an ordinary point for Green's function g).

If the point ξ,η now approaches a point $s'(x'y')$ on the path of the line integral from the left side, the equation becomes

$$a_1(s')\,G(s',\xi_1\eta_1) - \left[\frac{a_1(s')-a_2(s')}{2}\right]G(s',\xi_1\eta_1)$$
$$+\int [a_1(s)-a_2(s)]\,\frac{\partial g}{\partial n_1}(s,s')\,G(s,\xi_1\eta_1)ds = g(\xi_1\eta_1,x'y') = g(\xi_1\eta_1,s'),$$

*Cf Volterra, *Leçons sur les Équations Intégrales*, page 126. Max. Mason, *Boundary Value Problems*, New Haven Math. Colloquium.

or

$$(15) \qquad G(s',\xi_1\eta_1)+2\int\frac{a_1(s)-a_2(s)}{a_1(s')+a_2(s')}\frac{\partial g(s,s')}{\partial n_1}G(s,\xi_1\eta_1)ds=\frac{2g(\xi_1\eta_1,s')}{a_1(s')+a_2(s')}.$$

The same result is obtained when ξ,η approaches $s'(x'y')$ from the right, since the contribution of the infinitesimal element of the line integral at s' is now

$$\left[\frac{a_1(s')-a_2(s')}{2}\right]G(s',\xi_1\eta_1).$$

This is found from the fact that the principal part of $\dfrac{\partial g(s,\xi\eta)}{\partial m_1}$ becomes $\dfrac{\partial\theta(s,\xi\eta)}{\partial s_1}$.

If the characteristic determinant of the integral equation (15) does not vanish, it suffices to uniquely determine $G(s,\xi_1\eta_1)$ at all points s on the boundary curve, and hence by (3) at all points in the dielectric region. To prove that this determinant cannot vanish we make use of the fact that if it does there must be at least one solution $G_0(s,\xi_1\eta_1)$ not identically zero of the homogeneous integral equation obtained by replacing the second member of (15) by zero; *i.e.*,

$$(16) \qquad G_0(s',\xi_1\eta_1)+2\int\frac{a_1(s)-a_2(s)}{a_1(s')+a_2(s')}\frac{\partial g(s,s')}{\partial n_1}G_0(s,\xi_1\eta_1)ds=0.$$

But if this were the case we could construct a scalar potential function $\phi_0(\xi\eta)$ for any point ξ,η within the dielectric region by the formula

$$(17) \qquad \phi_0(\xi\eta)=-\frac{1}{a(\xi\eta)}\int[a_1(s)-a_2(s)]\frac{\partial g(s,\xi\eta)}{\partial n_1}G_0(s,\xi_1\eta_1)ds$$

which could not be identically zero in all parts of the plane. From the property of Green's function g, it is evident that $\phi_0(\xi\eta)$ will vanish when ξ,η moves up to any conducting boundary, will be harmonic in ξ,η at all ordinary points. Also it is evident that $a\dfrac{\partial\phi_0}{\partial n}$ will be continuous at the boundaries between dielectrics, and $\lim\limits_{r=\infty}\left(r^2\dfrac{\partial\phi}{\partial r}\right)$ will exist. Finally if ξ,η approaches a point on the boundary between two dielectrics, then the homogeneous integral equation (6), which G_0 is assumed to satisfy, will show that ϕ_0 also is continuous there. Thus ϕ_0 possesses all the properties which have been shown sufficient to assure its non-existence. Therefore it may be concluded that a function $G(xy,\xi\eta)$ and one only may be found.

3. *The Existence of ϕ and its Integral Representation.*

Assuming that the function G has been constructed, we may apply Green's theorem to the two functions $\phi(xy)$ and $G(xy,\xi\eta)$, and obtain

$$-\int a_j\left[\phi(xy)\frac{\partial G(xy,\xi\eta)}{\partial n}-G(xy,\xi\eta)\frac{\partial\phi(xy)}{\partial n}\right]ds_j=\phi(\xi\eta)$$

if ξ,η is a point within the homogeneous dielectric section S_j, the integration being taken around the artificial boundary of S_j. Applied to any other dielectric section S, the theorem gives

$$-\int a\left[\phi(xy)\frac{\partial G(xy,\xi\eta)}{\partial n}-G(xy,\xi\eta)\frac{\partial\phi(xy)}{\partial n}\right]ds=0.$$

Adding the first expression to all the others of the second type so as to include the entire dielectric region, it will be found that all the integrals over the boundaries between two dielectrics cancel, because of the continuity of ϕ and $a\frac{\partial\phi}{\xi n}$ on the one hand and of G and $a\frac{\partial G}{\partial n}$ on the other. Also the integrals $G(xy,\xi\eta)\frac{\partial\phi}{\partial n}$ over the conducting boundaries of the region vanish because G vanishes there. The integrals over the infinite circle vanish because $r\frac{\partial\phi}{\partial r}$ and $r\frac{\partial G}{\partial r}$ vanish at infinity.

Hence we obtain the integral representation

$$\phi(\xi\eta)=-\sum_{k=1}^{N}c_k\int a(xy)\frac{\partial G(xy,\xi\eta)}{\partial n_k}\,ds_k=\sum_{k=1}^{N}c_k\phi_k(\xi\eta) \tag{18}$$

where

$$\phi_k(\xi\eta)\equiv-\int a(xy)\frac{\partial G(xy,\xi\eta)}{\partial n_k}\,ds_k, \tag{19}$$

this integral being taken around any closed contour in the dielectric region which surrounds the kth group of conductors only.

The N equations of type (9), Section II, for the conservation of electricity give N equations of the type

$$\gamma I_n=\int a\frac{\partial\phi}{\partial n_n}\,ds_n=\sum_{k=1}^{N}c_k\int a\frac{\partial\phi_k}{\partial n_n}\,ds_n=\sum_{k=1}^{N}\alpha_{nk}c_k \tag{20}$$

where the complex coefficient α_{nk} is defined by

$$\alpha_{nk}\equiv\int a\frac{\partial\phi_k}{\partial n_n}ds_n \tag{21}$$

where this integral is taken over any closed curve in the dielectric region which encircles the nth group of conductors only. From this definition and the definition (8) of ϕ_k it follows that the coefficients α_{nk} constitute a symmetrical array, that is,

$$\alpha_{nk}=\alpha_{kn}. \tag{22}$$

This is seen by inserting the definition (8) of ϕ_k in (21), letting x,y denote a point on or near the boundary of the conducting section S_n and ξ,η a point on or near the boundary of the conducting section S_k

$$a_{nk} = -\int\int ds_n ds_k a(xy) a(\xi\eta)\, \frac{\partial^2 G(xy,\xi\eta)}{\partial n_n \partial n_k}.$$

Since $G(xy,\xi\eta) = G(\xi\eta,xy)$ this is easily seen to be identical with the definition of a_{kn}.

The unique solution or integral representation (19) for $\phi(\xi\eta)$ has been obtained by assuming a function ϕ to exist which satisfies all the conditions in (7) of Section II. It remains to be shown that the function $\phi(\xi\eta)$ which is given by the right side of (19) does indeed satisfy all these conditions, and that if each of the N currents $I_1 I_2, \ldots, I_N$ satisfies the equation of conservation of electricity (20), the sum of these currents will be identically zero, whatever the value of the N constant potentials $c_1 c_2, \ldots, c_N$ or of γ.

To do this and to obtain an idea of their physical meaning we may examine the N partial potentials of type $\phi_k(\xi\eta)$ defined by (19).

4. *Properties and Physical Interpretation of the Partial Potentials ϕ_k.*

From the properties of $G(xy,\xi\eta)$ it follows that $\phi_k(\xi\eta)$ has the following properties

(*a*)
$$\left(\frac{\partial^2}{\partial \xi^2} + \frac{\partial^2}{\partial \eta^2}\right)\phi_k(\xi\eta) = 0$$

if ξ,η is any point not on a boundary. This is evident by differentiating under the integral sign twice with respect to ξ and to η, which is allowable since ξ,η is not a point on the line of integration. Since

$$\left(\frac{\partial^2}{\partial \xi^2} + \frac{\partial^2}{\partial \eta^2}\right) G(xy,\xi\eta) = 0$$

it follows that

$$\frac{\partial^2 \phi_k}{\partial \xi^2} + \frac{\partial^2 \phi_k}{\partial \eta^2} = 0.$$

(*b*) When ξ,η moves across any boundary between two dielectrics $\phi_k(\xi\eta)$ changes continuously because $G(xy,\xi\eta)$ is also a continuous function of ξ,η for any fixed values of x,y.

(*c*) At such boundaries $a\,\dfrac{\partial \phi_k}{\partial n}$ is also continuous because of the similar continuity of $a\dfrac{\partial G(xy,\xi\eta)}{\partial n}$ when x,y is any fixed point.

(*d*) The $\lim\limits_{r=\infty}\left(r^2 \dfrac{\partial \phi_k}{\partial r}\right)$ exists when ξ,η goes to infinity in any direction in the x,y plane, because of the similar property of $G(xy,\xi\eta)$ when x,y is a finite point.

(*e*) When ξ,η moves off to infinity in any direction in the x,y plane $\phi_k(\xi\eta)$ will become constant, say $\phi_k(\infty)$, which will not necessarily be zero. From the definition (8) it is evident that $\phi_k(\xi\eta)$ is equal to the total conduction and displacement current (per unit length along z) which flows into the k^{th} conductor, from the unit line source at ξ,η, when all the conductors are at zero potential *and all together receive the total current unity*, from the line source. It is thus evident that if these conditions are satisfied, that $\phi_k(\xi\eta)$ will not in general vanish when ξ,η goes to infinity. If there is one conductor extending to infinity, the N^{th} say, then $\phi_N(\infty)$ will be unity and $\phi_n(\infty)=0$, if $n\neq N$, since the infinite conductor will in this case receive all of the current from the line source at ξ,η when the latter moves to infinity. The same statements are true if ϕ_N surrounds all the others.

(*f*) When the point ξ,η approaches any conducting section S_n, then $\phi_k(\xi\eta)$ approaches zero if $n\neq k$ and $\phi_k(\xi\eta)$ approaches unity at the boundary of conducting section S_k. This is evident from the above interpretation of $\phi_k(\xi\eta)$ as the current into S_k when there is a unit line source at ξ,η. When ξ,η comes infinitely close to the boundary of any conductor, the unit current will all flow from the line source into the conductor in its immediate neighbourhood and hence the flow into any other conductor will be zero.

Thus ϕ_k possesses all the properties which we have shown are necessary to make it unique and it may be determined by the methods of harmonic analysis.

From these properties of each ϕ_k it is evident that the function $\phi(\xi\eta)$ given by (18) does satisfy all the conditions (7) Section II. It is the only one which exists.

The N functions $\phi_1, \phi_2, \ldots, \phi_N$ are not all independentfor there exists a linear relation between them:

$$\sum_{k=1}^{N} \phi_k(\xi\eta)=1. \tag{23}$$

This also follows from the interpretation of $\phi_k(\xi\eta)$ as current into the section S_k, since altogether the N conducting sections receive the current unity from the line source at ξ, η, equations (19) and (13). From (23) it follows that there are N homogeneous linear relations existing between the coefficients a_{nk}. Thus operating on (23) by $\int a \dfrac{\partial}{\partial n_n} ds_n$ gives

$$\sum_{n=1}^{N} a_{nk}=0, \quad (k=1, 2, 3, \ldots, N). \tag{24}$$

From this it is evident that $I_1+I_2+\ldots+I_N=0$ identically whatever the values of $c_1, c_2, \ldots, c_N$ and γ. The last of the N equations (20) is not independent of the first $N-1$ equations but may be derived from them by simply adding them together. The determinant of this array of coefficients vanishes and it is not possible to solve this set of equations for all the c's in terms of the I's.

Hence when each ϕ_k has been found, then if each of the N currents satisfies the equation of conservation of electricity (20) the necessary relation (10), that the sum of all the currents shall vanish, will be automatically satisfied.

The physical meaning of $\phi_k(\xi\eta)$ follows from its properties. It represents the potential at any point ξ,η when the k^{th} conductor is at unit potential and all the others at zero potential in the presence of the various conducting dielectrics, subject to the additional condition that the whole flow out from the k^{th} conductor goes into the remaining ones, as there is no flow to infinity in the x,y plane. This is the meaning of the relations (23) and (24).

It is evident that ϕ_k is in general complex on account of the surface charges at the boundaries between two dielectrics. These charges are eliminated from appearance by the use of $G(xy,\xi\eta)$ through which function their influence is exerted, so to speak. If the dielectric were entirely homogeneous both as to its conductivity λ and its dielectric constant k, then the function ϕ_k would be a *real* electrostatic function, which would not involve λ, k or p, but would depend only upon the geometric configuration of the conducting sections which bound the dielectric region.

More generally, if the ratio

$$\frac{\lambda_1}{\lambda_2}=\frac{k_1}{k_2}=\frac{a_1}{a_2}=\frac{\lambda_1+\dfrac{ipk_1}{4\pi c^2}}{\lambda_2+\dfrac{ipk_2}{4\pi c^2}}$$

is real, for every pair of dielectrics in contact, then each ϕ_k is real, and involves the values of these real ratios but not the frequency. In particular if all the dielectrics are non-conducting, this ratio is real and ϕ_k will not involve the frequency.

In general the surface charges and their phases, at the boundaries between different dielectrics, although they are eliminated from consideration, are the cause of ϕ being complex and involving the frequency.

It is evident that if a real formula for ϕ_k can be obtained for the case of *steady* flow through the given dielectrics, with given conductivities $\lambda_1,\lambda_2,\ldots$, etc. then the complex solution ϕ_k may be obtained by substituting in this expression the corresponding complex conductivities $a_1,a_2,\ldots$, in place of their real values.

It is to be noted that ϕ_k does not involve γ.

5. *The Complex Coefficients of Leakage and Capacity.*

The complex charge of free electricity Q_n upon, and the leakage current G_n from, the n^{th} conducting group per unit length along z are given by:

$$(25)\qquad \begin{aligned} G_n&=\int\lambda\frac{\partial\phi}{\partial n_n}ds_n=\sum_{k=1}^{N}c_k\int\lambda\frac{\partial\phi_k}{\partial n_n}ds_n,\\ Q_n&=\frac{1}{4\pi c^2}\int k\frac{\partial\phi}{\partial n_n}ds_n=\sum_{k=1}^{N}c_k\frac{1}{4\pi c^2}\int k\frac{\partial\phi_k}{\partial n_n}ds_n.\end{aligned}$$

Consequently if we resolve the complex coefficient a_{nk} into its real and imaginary components

$$-a_{nk}=G_{nk}+ipC_{nk} \text{ where } \begin{cases} G_{nk}=G_{kn} \text{ and } \sum_{k=1}^{N} G_{nk}=0, \\ C_{nk}=C_{kn} \text{ and } \sum_{k=1}^{N} C_{nk}=0, \end{cases} \tag{26}$$

the equation of conservation of electricity (20) takes the form

$$\gamma I_n=G_n+ipQ_n=-\sum_{k=1}^{N}(G_{nk}+ipC_{nk})c_k. \tag{27}$$

On adding to this equation the identically zero quantity

$$-c_n\sum_{k=1}^{N}a_{nk}=-\sum_{k=1}^{N}a_{nk}c_n=\sum_{k=1}^{N}(G_{nk}+ipC_{nk})c_n$$

it takes the form

$$\gamma I_n=G_n+ipQ_n=\sum_{k=1}^{N-1}(G_{nk}+ipC_{nk})\ (c_n-c_k), \tag{28}$$

which shows that the real coefficients G_{nk} and C_{nk} where $n\neq k$ are coefficients of leakage and capacitance respectively. The coefficients G_{nn} and C_{nn} do not occur in the form (28) and may be regarded simply as the negative of the sum of all the other coefficients G_{nk} and C_{nk} respectively.

If the frequency is very low, or if the entire dielectric region is homogeneous, or more generally whenever the functions ϕ_n are real, G_{nk} and C_{nk} will not be functions of the frequency and will have their ordinary electrostatic significance since they are then defined by

$$G_{nk}=\int\lambda\frac{\partial\phi_k}{\partial n}ds_n,\quad C_{nk}=\frac{1}{4\pi c^2}\int k\frac{\partial\phi_k}{dn}\partial s_n. \tag{29}$$

It is evident that in general the formula (29) may be put in a form similar to (28), namely,

$$\gamma I_n=\sum_{k=1}^{N-1}a_{nk}(c_k-c_N),\ (n=1,\ 2,\ 3,\ \ldots,\ N). \tag{29'}$$

IV. The Complex Vector Potential A

1. *Existence and Properties of a Magnetic Flux Function M.*

Let $s(xy)$ be any point on the boundary between materials of different permeability. Let ξ,η and ξ_1,η_1 be two other points not on this curve. Consider the vector potential function whose x and y components are zero, but whose z component has at the point ξ,η the value given by the logarithmic potential

$$M(\xi\eta,\xi_1\eta_1)=\mu(\xi\eta)\left[-2\log r(\xi\eta,\xi_1\eta_1)-2\int\sigma(s)\frac{\partial\log r(xy,\xi\eta)}{\partial n}ds\right] \tag{31}$$

where the direction of ds along the curve may be taken at pleasure, but that of dn is that of the normal drawn toward the curve in the medium on the left side. The integral represents the value at ξ,η of the logarithmic potential of a double distribution of strength $\sigma(s)$ upon the boundary curves where μ is discontinuous. This is harmonic and has continuous normal derivatives [ξ,η being the variable point] at all points. It also vanishes when ξ,η moves off to infinity. It may also be written in the form

$$-2\int \sigma(s)\frac{\partial \log r(xy,\xi\eta)}{\partial n}\,ds = -2\int \sigma(s)\frac{\partial\theta(xy,\xi\eta)}{\partial s}\,ds$$

where $\theta(xy,\xi\eta)$ is the angle between the positive direction of the x axis and the line drawn *from* $s(xy)$ *to* ξ,η.

Regarding ξ_1,η_1 as the fixed point and ξ,η variable, it is evident that (whatever the function $\sigma(s)$ so long as the integral is convergent), the function $M(\xi\eta,\xi_1\eta_1)$ satisfies $\left(\frac{\partial^2}{\partial\xi^2}+\frac{\partial^2}{\partial\eta^2}\right)M(\xi\eta,\xi_1,\eta_1)=0$ and takes the form $-2\mu(\xi\eta)\log r(\xi\eta,\xi_1\eta_1)$ when ξ,η goes to infinity, and that $\frac{1}{\mu(\xi\eta)}\frac{\partial}{\partial n}M(\xi\eta,\xi_1\eta_1)$ is continuous everywhere. The integral, however, is discontinuous at the boundary curves, but if $\sigma(s)$ can be so chosen as to keep $M(\xi\eta,\xi_1\eta_1)$ continuous there, then it is evident that $M(\xi\eta,\xi_1\eta_1)$ will represent a (z-component of) vector potential function at ξ,η due to a steady unit current filament at ξ_1,η_1 parallel to z, in the presence of all magnetic materials.

This is evident since the normal components of the magnetic induction derived from it is just $\frac{\partial M}{\partial s}$ and this is continuous if M is so. The tangential component of B divided by μ, that is of H, is $-\frac{1}{\mu}\frac{\partial M}{\partial n}$ and this will also be continuous. If we compute the line integral

$$\int H_s\,ds = \frac{1}{\mu}\int -\frac{\partial M(\xi\eta,\xi_1\eta_1)}{\partial n}\,ds$$

around an infinitesimal circle with centre at ξ_1,η_1 the result is 4π.

To determine the density $\sigma(s)$ we must express the fact that $M(\xi\eta,\xi_1\eta_1)$ approaches the same value when the point ξ,η approaches a point $s'(x'y')$ on the boundary curve first from the left and then from the right. In the first case, the equation gives, if $\mu_1(s')$ is the value on the left, $\mu_2(s')$ on the right of the curve at s',

$$-\frac{M(s',\xi_1\eta_1)}{2\mu_1(s')} = \log r(s',\xi_1\eta_1)+\pi\sigma(s')+\int \sigma(s)\frac{\partial\theta(s,s')}{\partial s}ds.$$

In the second case

$$-\frac{M(s',\xi_1\eta_1)}{2\mu_2(s')} = \log r(s',\xi_1\eta_1)-\pi\sigma(s')+\int \sigma(s)\frac{\partial\theta(s,s')}{\partial s}ds.$$

Eliminating $M(s',\xi_1\eta_1)$ gives the integral equation

$$(32)\qquad \sigma(s')+\frac{1}{\pi}\int\left(\frac{\mu_1-\mu_2}{\mu_1+\mu_2}\right)_{s'}\frac{\partial\theta(s,s')}{\partial s}\,\sigma(s)ds = -\frac{1}{\pi}\left(\frac{\mu_1-\mu_2}{\mu_1+\mu_2}\right)_{s'}\log r(s',\xi_1\eta_1)$$

to determine the density of the double distribution $\sigma(s)$. The nucleus of this equation becomes infinite when $s=s'$ but is integrable, and the equation may be reduced to one with a finite and continuous nucleus by multiplying by this nucleus and integrating over the range. This equation uniquely determines the function σ provided the Fredholm determinant is not zero.

That this cannot be the case may be proven from the known fact that when this determinant vanishes there is at least one solution $\sigma_0(s)$, not identically zero, of the homogeneous equation obtained by placing the right side of the above equation equal to zero. With this function $\sigma_0(s)$ we could then form the potential $v_0(\xi\eta)$ for every point ξ,η by the formula

$$(33) \qquad v_0(\xi\eta) = -2\mu(\xi\eta)\int \sigma_0(s)\,\frac{\partial \log r(xy,\xi\eta)}{\partial n}\,ds = -2\mu(\xi\eta)\int \sigma_0(s)\,\frac{\partial \theta(xy,\xi\eta)}{\partial s}\,ds.$$

This function v_0 would satisfy $\left(\frac{\partial^2}{\partial \xi^2}+\frac{\partial^2}{\xi\eta^2}\right)v_0=0$ at all points, and $\frac{1}{\mu}\frac{\partial v}{\partial n}$ would be continuous everywhere and would vanish at infinity. In addition, v_0 itself would be continuous because of the assumed property of the function $\sigma_0(s)$. Moreover v_0 could not be identically zero in all sections of the plane. It is easy to show that in this case by transformations similar to those previously used that

$$\iint \frac{1}{\mu}\left[\left(\frac{\partial v_0}{\partial \xi}\right)^2+\left(\frac{\partial v_0}{\partial \eta}\right)^2\right] d\xi d\eta = 0$$

the integral being taken over the entire plane, and since μ is a real positive value everywhere, and v_0 is real, it is evident that v_0 must be a constant everywhere, and hence zero since it vanishes at infinity. Therefore the Fredholm determinant does not vanish and there is a unique solution $\sigma(s)$and the function $M(\xi\eta,\,\xi_1\eta_1)$ may be uniquely determined.

By applying Green's theorem to the two functions $M(xy,\xi\eta)$ and $M(xy,\xi_1\eta_1)$ we find that M is a symmetrical function of the two points x,y and ξ,η in whatever magnetic media the two points may be. That is

$$(34) \qquad M(xy,\xi\eta) = M(\xi\eta,xy).$$

In the special case where all the materials in space have the same magnetic permeability $M(xy,\xi\eta)$ reduces to $-2\mu\log r(xy,\xi\eta)$.

2. *Integral Equations for A and E.*

By means of the function M we may formulate the conditions (8), Section II, for A as an integral equation involving surface integrals over the conducting sections but free from boundary terms.

To do this apply Green's theorem to the two functions $A(xy)$ and $M(xy,\xi\eta)$ where ξ,η is some point within the section S_j which may be either a dielectric or conducting section. This point ξ,η being excluded by an infinitesimal circle, the theorem gives

$$-\frac{1}{4\pi}\iint\frac{M(xy,\xi\eta)\nabla^2 A}{\mu}dS+\frac{1}{4\pi}\int\frac{1}{\mu}\left[M(xy,\xi\eta)\frac{\partial A}{\partial n}-A\frac{\partial M(xy,\xi\eta)}{\partial n}\right]ds$$

$$=A(\xi\eta),\ \text{if}\ \xi,\eta\ \text{is within}\ S_j,$$
$$=0\ \text{for all other sections.}$$

Adding together all such expressions corresponding to the entire x,y plane, the line integrals over both sides of all boundaries between different magnetic media cancel, because of the continuity of A, N, $\frac{1}{\mu}\frac{\partial A}{\partial n}$ and $\frac{1}{\mu}\frac{\partial M}{\partial n}$. The integral over the infinite circle gives the term $A(\infty)$. (However, it is important to notice at this point that the same result would have been obtained if A were assumed to take the asymptotic form $A=A(\infty)+A'\log r$ at infinity). Replacing $-\frac{\nabla^2 A}{4\pi\mu}$ by its value $-\lambda(ipA-\gamma\phi)$ from (8a) of Section II, gives, after interchanging the notation of the points x,y and ξ,η:

$$A(xy)+ip\iint M(xy,\xi\eta)\lambda(\xi\eta)\left[A(\xi\eta)-\frac{\gamma\phi(\xi\eta)}{ip}\right]d\xi d\eta=A(\infty). \tag{35}$$

Or since $E=\gamma\phi-ipA$, by (6) Section II,

$$E(xy)+ip\iint M(xy,\xi\eta)\lambda(\xi\eta)E(\xi\eta)d\xi d\eta=\gamma\phi(xy)-ipA(\infty) \tag{36}$$

$$\equiv F(xy)\equiv\gamma[\phi(xy)-\phi(\infty)]+E(\infty),\ [\text{since}\ E(\infty)=\gamma\phi(\infty)-ipA(\infty)],$$

where the integration extends over all conducting sections. The equation must hold whether x,y is a point in a conductor or in a dielectric. Hence after the value of E has been determined for all points in conducting sections, this equation becomes an explicit formula giving the value of E at any other point. As a matter of fact the value of E at outside points is seldom of practical interest, and the value of E in conductors may be found without reference to its value in dielectrics. All of the conditions (8) of Section II, except the last, have been assumed in the derivation of (35) and (36), and it is easy to show from the properties of M that if this equation is satisfied all of these conditions will be fulfilled with the exception of the last one, which requires that $r\frac{\partial E}{\partial r}\left(\text{or}\ r\frac{\partial A}{\partial r}\right)$ shall vanish canonically at infinity, a condition which is reciprocal with the relation $\sum_1^N I_n=0$. This condition will therefore determine the value of $A(\infty)$ or $E(\infty)$.

The integral equation (36) has a unique solution given in the form of a definite integral when the second member $F(xy)$ is any function given in all conducting sections which makes this integral convergent, so that the undetermined constant $E(\infty)$ which made its appearance in (36) must be so chosen as a linear function of the N constants $c_1, c_2, \ldots, c_N$ that $\sum_{n=1}^N I_n$ (where I_n is defined by (11) Section II) shall be zero whatever the values of $c_1, c_2, \ldots, c_N$. These constants may then all be considered as arbitrarily assignable. It may be noted that in case the N^{th} conductor extends to infinity the value of $E(\infty)$ must be zero since I_N is assumed to be finite. Also, in this case $\phi_N(\infty)=1$ and $\phi_n(\infty)=0$ when $n\neq N$. Consequently, the second member of (36), namely $F(xy)$, which has different constant values on each conducting group, will have the value zero when x,y is on the N^{th} conductor, since

$$\gamma[\phi(x_N y_N)-\phi_N(\infty)]+E(\infty)=\gamma[1-1]+0=0.$$

On any other conducting section S_n it has the value γc_n if $n\neq N$.

3. *Existence and Uniqueness of a Solution of* (36).

Multiplying the expression (36)

$$[E(x'y')+ip\iint M(x'y',\xi\eta)\lambda(\xi\eta)E(\xi,\eta)d\xi d\eta-F(x'y')]=0$$

by $ipM(xy,x'y')\lambda(x'y')dx'dy'$ and integrating with respect to x',y' over the entire range shows that if E satisfies (36) it must also satisfy the equation

$$\text{(37)}\qquad ip\iint\lambda(x'y')M(xy,x'y')dx'dy' \times[E(x'y')+ip\iint M(x'y',\xi\eta)\lambda(\xi\eta)E(\xi\eta)d\xi d\eta-F(x'y')]=0.$$

Conversely, if this equation is satisfied, we may show that the bracket must be identically zero and therefore (36) will also be satisfied. To prove this let the bracket be represented by $h'(xy)+ih''(xy)$. Then $h'(xy)$ must be a real function such that

$$\iint M(xy,x'y')h'(x'y')dx'dy'=0$$

identically for all values of x,y, and similarly for h''. Denoting this integral by $V(xy)$, it is evident that if $V(xy)=0$ identically then $\nabla^2 V(xy)=0$ everywhere in the range. But $\nabla^2 V(xy)=4\pi\mu(xy)h(xy)$ as is seen by differentiation of the integral, and from the properties of M. Consequently $h(xy)$ must be zero everywhere if h is such that $\iint h^2 dx\,dy$ exists. Therefore the equations (36) and (37) are reciprocal.

By making use of (36) the equation (37) may be written

$$\text{(38)}\qquad E(xy)+p^2\iint dx'dy'M(xy,x'y')\lambda(x'y')\iint M(x'y',\xi\eta)\lambda(\xi\eta)E(\xi\eta)d\xi d\eta = F(xy)-ip\iint M(xy,\xi\eta)\lambda(\xi\eta)F(\xi\eta)d\xi d\eta.$$

The second member of this equation will be finite even if the N^{th} conducting section extends to infinity since $F(xy)$ will vanish on this section.

We may for the present limit the discussion to the case where all conducting sections are finite. The form thus obtained for the solution then suggests methods of dealing with the equation with open sections. With finite range, the order of integration in preceding integral may be interchanged and the equation takes the form

$$E(xy)+p^2\iint N(xy,\xi\eta)\lambda(\xi\eta)E(\xi\eta)d\xi d\eta=f(xy) \tag{39}$$

$$=F(xy)-ip\iint M(xy,x'y')\lambda(x'y')F(x'y')dx'dy'$$

where the new nucleus

$$N(xy,\xi\eta)=N(\xi\eta,xy)=\iint M(xy,x'y')M(\xi\eta,x'y')\lambda(x'y')dx'dy' \tag{40}$$

is not only symmetrical in the two points but is everywhere finite—a property not possessed by M which becomes logarithmically infinite when the two points approach each other. The theorems of Fredholm are applicable to this equation and show that if p is not a root of the characteristic determinant, there is one and only one solution.

$$E(xy)=F(xy)-ip\iint N(xy,\xi\eta,ip)F(\xi\eta)d\xi d\eta \tag{41}$$

where

$$N(xy,\xi\eta,ip)=\frac{D_N(xy,\xi\eta,ip)}{D_N(ip)}$$

is the resolving nucleus of the primitive equation and is given by Fredholm's formula. It satisfies the two integral equations

$$M(xy,x'y')\lambda(x'y')-N(xy,x'y',ip)$$

$$=ip\iint M(xy,\xi\eta)\lambda(\xi\eta)N(\xi\eta,x'y',ip)d\xi d\eta \tag{42}$$

$$=ip\iint M(\xi\eta,x'y')\lambda(x'y')N(xy,\xi\eta,ip)d\xi d\eta.$$

It is evident that $N(xy, xy)$ is finite since

$$N(xy,\ xy)=\iint [M(x'y',\xi\eta)]^2\lambda(x'y')dx'dy'$$

and the surface element when x',y' is near the point ξ,η may be written in the form

$$\int_0^{2\pi}d\theta\int_0^{\epsilon}rdr\Big[-2\mu\log r\Big]^2=8\pi\mu^2(\xi\eta)\int_0^{\epsilon}r\ (\log r)^2dr=4\pi\mu^2(\xi\eta)\epsilon^2\Big[\log^2\epsilon-\log\epsilon+\frac{1}{2}\Big]$$

and this vanishes with ϵ showing that the element of the surface integral near the point ξ,η contributes an infinitesimal amount to the integral.

If p were a root of $D_N(ip)=0$ then N would become infinite and the solution impossible in general. But it is known that in this case there must be at least one solution $E_0(xy)$, which is not identically zero, of the homogeneous equation obtained by placing the second member of (9) equal to zero. This solution $E_0(xy)$ would also satisfy

$$E_0(xy)+ip\iint M(xy,\xi\eta)\lambda(\xi\eta)E_0(\xi\eta)d\xi d\eta=0. \tag{43}$$

If we write $E_0\ (xy)=E'_0(xy)+iE_0''(xy)$ where E_0' and E_0'' are real functions, and multiply this equation by

$$[E_0'(xy)-iE_0''(xy)]\lambda(xy)dxdy$$

and integrate over the entire range, we obtain

$$\iint\lambda(xy)[E_0'^2(xy)+E_0''^2(xy)]dxdy$$

$$+p\iint dxdy\iint d\xi d\eta M(xy,\xi\eta)\lambda(xy)\lambda(\xi\eta)[E_0'(\xi\eta)E_0''(xy)-E_0''(\xi\eta)E_0'(xy)] \tag{44}$$

$$=-ip\iint dxdy\iint d\xi d\eta M(xy,\xi\eta)\lambda(xy)\lambda(\xi\eta)[E_0'(\xi\eta)E_0'(xy)+E_0''(\xi\eta)E_0''(xy)].$$

The second member of this equation is a pure imaginary, the first member real, hence each side must vanish. The second integral on the left vanishes because of the symmetrical property $M(xy,\xi\eta)=M(\xi\eta,xy)$. Hence we must have

$$\iint\lambda(E_0'^2+E_0''^2)ds=0, \quad \text{or } E_0'\equiv E_0''\equiv E_0(x)\equiv 0,$$

and therefore no real value of p can be a characteristic constant for the primitive integral equation. A unique solution exists for all real values of p.

4. *Other Methods of Solving the Integral Equation* (36).

(*a*) *By Normal Functions.*

It is worth while to examine the equation from the point of view of normal functions. The nucleus $N(xy,x'y')$ is not only symmetrical and finite but it is *definite*, by which it is meant that there is no real function $h(xy)$ whose square is integrable (*i.e.*, such that $\iint h^2(xy)dxdy$ exists and is not zero), which will make the integral $\iint dxdy\iint dx'dy'N(xy,x'y')h(xy)h(x'y')$ vanish. This may be proven by noting that

$$\iint dxdy\iint dx'dy'N(xy,x'y')h(xy)h(x'y')$$

$$=\iint\lambda(\xi\eta)d\xi d\eta)\iint M(xy,\xi\eta)h(xy)dxdy\iint M(x'y',\xi\eta)h(x'y')dx'dy'$$

$$=\iint\lambda(\xi\eta)d\xi d\eta)\iint M(xy,\xi\eta)h(xy)dxdy]^2$$

which is always positive unless there is a function h for which

$$\iint M(xy, \xi\eta)h(xy)dxdy = 0,$$

identically for all points ξ, η in the range. We have just shown that no such function exists.

From this "definiteness" of N it is known that its characteristic constants are infinite in number, real, and positive, and form a denumerable ensemble of isolated points. To each such constant τ_n^2, there corresponds a normal function $u_n(xy)$ which is a fundamental solution of the equation

$$u_n(xy) - \tau_n^2 \iint N(xy, \xi\eta)u_n(\xi\eta)d\xi d\eta = 0. \tag{45}$$

The functions $u_1(xy)$, $u_2(xy)$, . . . are infinite in number and constitute a closed set of normal functions such that

$$\iint u_n(xy)u_n(xy)dxdy = 0 \qquad \text{if } n \neq m,$$
$$= 1 \qquad \text{if } n = m,$$

and the nucleus $N(xy, \xi\eta)$ is equal to the uniformly convergent series

$$N(xy, \xi\eta) = \sum_{n=1}^{\infty} \frac{u_n(xy)u_n(\xi\eta)}{\tau_n^2}. \tag{46}$$

An arbitrary function $f(xy)$ may be developed in a uniformly convergent series of these functions in the form

$$f(xy) = \sum_{n=1}^{\infty} f_n u_n(xy) \tag{47}$$

where the Fourier coefficient f_n is given by

$$f_n = \iint F(xy)u_n(xy)dxdy \tag{48}$$

provided the series $\sum_1^\infty f_n^2$ is convergent. Since the set of normal functions is "closed" there is no function which is normal to all of them, *i.e.*, orthogonal to $N(xy, \xi\eta)$. The equation (39) may then be written

$$E(xy) + p^2 \sum_{n=1}^{\infty} \frac{u_n(xy)}{\tau_n^2} \iint u_n(\xi\eta)\lambda(\xi\eta)E(\xi\eta)d\xi d\eta = f(xy) = \sum_{n=1}^{\infty} f_n u_n(xy). \tag{49}$$

In the simplest case where λ has the same value in all materials, assume $E(xy) = \sum_1^\infty E_n u_n(xy)$ and substitute in (49). This gives

$$\sum_{n=1}^{\infty} u_n(xy) \left[E_n + \frac{p^2\lambda}{\tau_n^2} E_n - f_n \right] = 0,$$

giving

$$E_n = \frac{f_n}{1+\lambda \frac{p^2}{\tau_n^2}},$$

and the solution is

$$(50) \qquad E(xy) = \sum_{n=1}^{\infty} \frac{F_n u_n(xy)}{1+\lambda \frac{p^2}{\tau_n^2}} = \iint F(\xi\eta) d\xi d\eta \sum_{n=1}^{\infty} \frac{u_n(xy) u_n(\xi\eta)}{1+\lambda \frac{p^2}{\tau_n^2}}$$

$$= \iint d\xi d\eta F(\xi\eta) \sum_{n=1}^{\infty} \frac{u_n(xy) u_n(\xi\eta)}{1+\lambda \frac{p^2}{\tau_n^2}}.$$

In the general case where λ has different constant values in different conductors, if we write

$$\lambda(xy) E(xy) = \sum_{n=1}^{\infty} b_n u_n(xy)$$

and

$$E(xy) = \sum_{n=1}^{\infty} E_n u_n(xy)$$

and substitute in the equation, we find

$$\sum_{n=1}^{\infty} u_n(xy) \left[E_n + \frac{p^2}{\tau_n^2} b_n - f_n \right] = 0$$

which will be satisfied if we may make

$$\frac{\tau_n^2}{p^2} E_n + b_n = \frac{\tau_n^2 f_n}{p^2}, \qquad (n = 1, 2, 3, \ldots \infty).$$

Now

$$b_n = \iint u_n(xy) \lambda(xy) E(xy) dx dy = \iint u_n(xy) \lambda(xy) \sum_{k=1}^{\infty} E_k u_k(xy) dx dy$$

or

$$b_n = \sum_{k=1}^{\infty} g_{nk} E_k$$

where

$$g_{nk} = g_{kn} = \iint \lambda(xy) u_n(xy) u_k(xy) dx dy.$$

Hence the constants E_n will be the solution of an infinite set of linear equations

$$\frac{\tau_n^2}{p^2} E_n + \sum_{k=1}^{\infty} g_{nk} E_k = \frac{\tau_n^2 f_n}{p^2}, \qquad (n = 1, 2, 3, \ldots, N).$$

or

$$\left(\frac{\tau_1^2}{p^2}+g_{11}\right)E_1+g_{12}E_2+g_{13}E_3+\ldots=\frac{\tau_1^2 f_1}{p^2},$$

$$g_{21}E_1=\left(\frac{\tau_2^2}{p^2}+g_{22}\right)E_2+g_{23}E_3+\ldots=\frac{\tau_2^2 f_2}{p^2},$$

$$g_{31}E_1+g_{32}E_2+\left(\frac{\tau_3^2}{p^2}+g_{33}\right)E_3+\ldots=\frac{\tau_3^2 f_3}{p^2}.$$

This set may be solved because its determinant cannot vanish for any real value of p.

In case λ has the same constant value on all conductors,

$$g_{nn}=\lambda,\ g_{nk}=0, \quad \text{if } n\neq k;$$

and we obtain the previous formula.

In the case where a conductor extends to infinity, the range of integration is no longer finite and the characteristic values of a finite symmetrical nucleus would no longer constitute a set of isolated points but would become uniformly distributed along the real axis everywhere equally dense. The representation of an arbitrary function in an infinite series of normal functions, over a definite range, would then give place to its representation over the infinite range by a definite integral of which Fourier's integral is an example. Instead of a solution in an infinite series of normal functions, one may expect a solution in the form of a definite integral.

(*b*) *Method of Successive Integration at Low Frequencies.*

It may be noted that the method of iterated integrations which is always applicable to Volterra's type of equation, may be successfully applied to the present problem if the frequency is small. This gives the solution in ascending powers of ip and is only applicable for values of p less than the modulus of the first characteristic constant of the primitive equation. The solution will be identical with that obtained from the Maclaurin development of

$$N(xy,\xi\eta,ip)=\frac{D_N(xy,\xi\eta,ip)}{D_N(ip)}$$

in ascending powers of p. Although no real value of p can be a root of $D_N(ip)=0$, nevertheless the Maclaurin development is limited to the circle in the complex p-plan whose radius is less than the modulus of the smallest complex root.

(*c*) *Method of Harmonic Analysis.*

It seems probable that in the majority of cases, there will be less labour involved in solving the differential equation with its boundary conditions by some sort of series of harmonic functions, such as Fourier-Bessel expansions, than in solving the integral equation for A (or E).

The differential equation for A suggests a certain type of expansions for A as a series of appropriate functions in each conductor or group of conductors which will be normal functions for that section. If there are groups of conductors in contact for which λ or μ have different values, the continuity of A and of $\frac{1}{\mu}\frac{\partial A}{\partial n}$ at such internal boundaries will lead to certain relations between the coefficients. The mutual influence which the conducting groups exert upon each other across the intervening dielectrics may then be found by assuming certain forms of expansion for A in the dielectric and by then making A and $\frac{1}{\mu}\frac{\partial A}{\partial n}$ continuous at every boundary between conductors and dielectric. This step may, however, be replaced by the following process which avoids any reference to the dielectric.

(*d*) *Mixed Method.*

Let x, y be a point within any conducting section, and with ξ,η as the variable of integration apply Green's theorem to the two functions $A(\xi\eta)$ and $\log r(xy,\xi\eta)$ for the entire dielectric region. This gives

$$A(\infty)+\frac{1}{2\pi}\int\left[A(\xi\eta)\frac{\partial \log r(xy,\xi\eta)}{\partial n}-\log r(xy,\xi\eta)\frac{\partial A(\xi\eta)}{\partial n}\right]ds=0,$$

where the integration is taken in the dielectric just outside the conductors, and around the complete boundary of all conductors. Since A and $\frac{1}{\mu}\frac{\partial A}{\partial n}$ are continuous at such boundaries, this necessary condition which A must satisfy becomes

$$A(\infty)=\frac{1}{2\pi}\int\left[A(\xi\eta)\frac{\partial \log r(xy,\xi\eta)}{\partial n_i}-\log r(xy,\xi\eta)\frac{1}{\mu}\frac{\partial A}{\partial n_i}\right]ds, \tag{51}$$

where the integration is taken around the complete boundary of all conducting groups as before but just inside the conductors. The internal normal n_i points toward the boundary. If the series for the internal values of A at each conducting group be introduced into this integral, the result must be identically true in whatever conducting section the point x,y may be. By expressing this fact when x,y is in each group in succession, the required number of equations between the coefficients are obtained. This method is illustrated in the case of two circular cylinders at the end of this paper.

5. *Form and Properties of the Solution.*

The N functions $\omega_1(xy)$, $\omega_2(xy)$, ..., $\omega_N(xy)$, which are defined as the solutions of integral equations of type

$$\omega_k(xy)+ip\iint M(xy,\xi\eta)\lambda(\xi\eta)\omega_k(\xi\eta)d\xi d\eta=\phi_k(xy) \tag{52}$$

are all linearly independent, for if there were a relation of the form $\sum_{k=1}^{N} h_k\omega_k(xy)=0$ the above equation, on being multiplied by h_k, and the sum taken for all values of k leads to

$$\sum_{k=1}^{N} h_k\omega_k(xy)+ip\iint M(xy,\xi\eta)\lambda(\xi\eta)\sum_{k=1}^{N} h_k\omega_k(\xi\eta)d\xi d\eta=\sum_{k=1}^{N} h\phi_k(xy).$$

The hypothesis leads to the conclusion that $\sum_{k=1}^{N} h_k\phi_k(xy)=0$ for all values of x,y in the range. But since $\phi_k(xy)=1$ when x,y is on conducting section S_k and equals 0 when x,y is on S_n where $n\neq k$, it follows that $h_1=h_2=\ldots=h_n=0$.

It is also evident that the only solution of

$$\omega(xy)+ip\iint M(xy,\xi\eta)\lambda(\xi\eta)\omega(\xi\eta)d\xi d\eta=1$$

is

$$\omega(xy)=\sum_{k=1}^{N}\omega_k(xy).$$

The solution for E is therefore

(53) $$E(xy)=\sum_{k=1}^{N}[\gamma c_k-ipA(\infty)]\omega_k(xy)$$

and

(54) $$I_n=\sum_{k=1}^{N}[\gamma c_k-ipA(\infty)b_{nk}]$$

where

(55) $$b_{nk}=\iint\lambda\omega_k dS_n$$

the integration being taken over the section S_n of the nth conducting group. The array of complex constants b_{nk} is symmetrical, that is,

(56) $$b_{nk}=b_{kn}.$$

To show this, multiply the equation for ω_k by $\lambda(xy)\omega_n(xy)dxdy$ and integrate over the entire range. This gives

$$\iint\lambda(xy)\omega_n(xy)\omega_k(xy)dxdy+ip\iint dxdy\iint d\xi d\eta M(xy,\xi\eta)\lambda(xy)\omega_n(xy)\lambda(\xi\eta)\omega_k(\xi\eta)$$
$$=\iint\lambda(xy)\omega_n(xy)\phi_k(xy)dxdy=\iint\lambda\omega_n dS_k=b_{kn}.$$

If we multiply the equation for ω_n by $\lambda\omega_k\, dxdy$ and integrate, the right hand side is b_{nk} and the left side is easily seen to be identical with the above on account of the symmetrical property $M(xy,\xi\eta)=M(\xi\eta,xy)$. The constant $A(\infty)$ must be so chosen that the sum of all the currents I_n will vanish. If the expression for

E be multiplied by $\lambda(xy)dxdy$ and the integration taken over all conductors, this gives

$$\iint \lambda E dxdy = \sum_{n=1}^{N} I_n = 0$$

$$\gamma \sum_{k=1}^{N} c_k \iint \lambda(xy)\omega_k(xy)dxdy - ipA(\infty) \iint \lambda \sum_{k=1}^{N} \omega_k(xy)dxdy = 0.$$

Or

$$ipA(\infty) = -E(\infty) + \gamma\phi(\infty) = \frac{\gamma \sum_{n=1}^{N} \sum_{k=1}^{N} b_{nk}c_k}{\sum_{n=1}^{N} \sum_{k=1}^{N} b_{nk}} = \gamma \sum_{k=1}^{N} \Psi_k(\infty)c_k, \tag{57}$$

where

$$\Psi_k(\infty) \equiv \frac{\sum_{n=1}^{N} b_{nk}}{\sum_{n=1}^{N} \sum_{s=1}^{N} b_{ns}}, \tag{58}$$

which shows that

$$\sum_{k=1}^{N} \Psi_k(\infty) = 1.$$

In order that it shall always be possible to thus choose the constant $A(\infty)$ as in (57) so as to make $\sum I_n = 0$, the denominator in the above formula for $\Psi_k(\infty)$, namely,

$$\sum_{n=1}^{N} \sum_{k=1}^{N} b_{nk} \text{ or } \iint \lambda \sum \omega_k(xy)dxdy$$

must never be zero. To prove that this can never vanish let

$$\omega(xy) = \sum_{k=1}^{N} \omega_k(xy) = u(xy) + iv(xy)$$

where u and v are real functions satisfying

$$u(xy) - p \iint M(xy, \xi\eta)\lambda(\xi\eta)v(\xi\eta)d\xi d\eta = 1$$

$$v(xy) + p \iint M(xy, \xi\eta)\lambda(\xi\eta)u(\xi\eta)d\xi d\eta = 0.$$

Multiply the first of these by $\lambda(xy)u(xy)dxdy$, the second by $\lambda(xy)v(xy)dxdy$, add the results, and integrate over the range. This gives, on account of the symmetry of M,

$$\iint \lambda(xy)[u^2(xy) + v^2(xy)]dxdy = \iint \lambda(xy)u(xy)dxdy.$$

Now if the sum in question were zero, then

$$\iint \lambda u dxdy = \iint \lambda v dxdy = 0,$$

and in this case the above relation would require that u and v and hence

$$\omega \equiv \sum_{k=1}^{N} \omega_k(xy)$$

be identically zero at all points of the range, which is never the case since the functions $\omega_k(xy)$ are not linearly dependent. This shows not only that the double sum in the denominator of (58) never vanishes, but that its real part is always positive. Consequently it will always be possible to choose $A(\infty)$ by (57) and (58) so that the sum of the currents will vanish. A similar treatment of $\omega_k(xy)$ shows that the real part of b_{kk} is positive and less than $\dfrac{1}{R_k(0)}$.

Inserting the value of $ipA(\infty)$ in the expression (53) for $E(xy)$ gives

$$E(xy) = \gamma \sum_{k=1}^{N} c_k \omega_k(xy) - \gamma \sum_{s=1}^{N} \omega_s(xy) \sum_{k=1}^{N} \Psi_k(\infty) c_k$$

$$= \gamma \sum_{k=1}^{N} c_k [\omega_k(xy) - \Psi_k(\infty) \sum_{s=1}^{N} \omega_s(xy)].$$

If we define the new set of functions $\Omega_1, \Omega_2, \ldots, \Omega_n$ by

(59) $$\Omega_k(xy) \equiv \omega_k(xy) - \Psi_k(\infty) \sum_{s=1}^{N} \omega_s(xy), \quad (k = 1, 2, \ldots, N),$$

then

(60) $$E(xy) = \gamma \sum_{k=1}^{N} c_k \Omega_k(xy).$$

From this definition it follows that

(61) $$\iint \lambda(xy) \Omega_k(xy) dx dy = 0, \quad (k = 1, 2, 3, \ldots, N).$$

Substituting

$$\omega_k(xy) = \Omega_k(xy) + \Psi_k(\infty) \sum_{s=1}^{N} \omega_s(xy)$$

in the integral equation (52) for ω_k gives

$$\Omega_k(xy) + ip \iint M(xy, \xi\eta) \lambda(\xi\eta) \Omega_k(\xi\eta) d\xi d\eta$$

$$+ \Psi_k(\infty) \sum_{s=1}^{N} [\omega_s(xy) + ip \iint M(xy, \xi\eta) \lambda(\xi\eta) \omega_s(\xi\eta) d\xi d\eta] = \phi_k(xy).$$

The bracket in this equation is just $\phi_s(xy)$ and $\sum_{s=1}^{N} \phi_s(xy) = 1$. Hence $\Omega_k(xy)$ the solution of the integral equation

(62) $$\Omega_k(xy) + ip \iint M(xy, \xi\eta) \lambda(\xi\eta) \Omega_k(\xi\eta) d\xi d\eta = \phi_k(xy) - \Psi_k(\infty).$$

This equation suffices to uniquely determine each function $\Omega_k(xy)$ at all points conductors. However, the N functions thus determined will not all be independent, for since

$$\sum_{k=1}^{N} \Psi_k(\infty) = 1$$

it follows from the definition of Ω_k that

$$\sum_{k=1}^{N} \Omega_k(xy) = 0. \tag{63}$$

If we define the coefficient β_{nk} by

$$\beta_{nk} \equiv \iint \lambda \Omega_k dS_n, \tag{64}$$

then the array of coefficients β_{nk} is symmetrical:

$$\beta_{nk} = \beta_{kn}. \tag{65}$$

This is proven in a manner precisely similar to that used in proving $b_{nk} = b_{kn}$. The constants β_{nk} also possess the property corresponding to equation (31):

$$\sum_{n=1}^{N} \beta_{nk} = 0, \quad (k = 1, 2, 3, \ldots, N). \tag{66}$$

The equation of definition of current for each conducting group takes the form

$$I_n = \gamma \sum_{k=1}^{N} \beta_{nk} c_k = \gamma \sum_{k=1}^{N} \beta_{nk}(c_k - c_N), \quad (n = 1, 2, 3, \ldots, N). \tag{67}$$

From some points of view it is more simple to introduce in place of the function Ω_k the functions $\Psi_1(xy), \ldots, \Psi_N(xy)$ defined by

$$\Psi_k(xy) = \phi_k(xy) - \Omega_k(xy). \tag{68}$$

Each must be the solution of an integral equation of the type

$$\Psi_k(xy) + ip \iint M(xy, \xi\eta) \lambda(\xi\eta) [\Psi_k(\xi\eta) - \phi_k(\xi\eta)] d\xi d\eta = \Psi_k(\infty). \tag{69}$$

This equation suffices to uniquely determine $\Psi_k(xy)$ at all points in conductors, and serves as a definition of $\Psi_k(xy)$ when the point x,y is in a dielectric. The vector potential $A(xy)$ is given at any point x,y by

$$A(xy) = \frac{\gamma}{ip} \sum_{k=1}^{N} c_k \Psi_k(xy), \tag{70}$$

which presents a certain formal analogy to the expression for the scalar potential:

$$\phi(xy) = \sum_{k=1}^{N} c_k \phi_k(xy).$$

The set of functions $\Psi_k(xy)$ and $\phi_k(xy)$ present a close analogy, for the symmetrical set of complex coefficients β_{nk} are analogous to the set a_{nk}. In fact the previous definition of β_{nk} may be reduced to a form precisely similar to that of α_{nk}, namely,

$$\beta_{nk}=\frac{1}{4\pi ip}\int\frac{\partial\Psi_k}{\partial n}\,ds_n, \tag{71}$$

where the integration is taken around any closed curve in the dielectric region which encircles the nth conducting group only, and where the normal points towards this conduction. This is evident since Ψ_k is harmonic in the dielectric, from its definition. Also since

$$\iint\lambda\Omega_k dxdy=\sum_{n=1}^{N}\iint\lambda\Omega_k dS_n=\iint\lambda\Psi_k dxdy=\sum_{n=1}^{N}\beta_{nk}=0,$$

it follows that $\Psi_k(xy)$ approaches the finite constant value $\Psi_k(\infty)$, when the point x,y moves off to infinity. The $\lim\limits_{r=\infty} r\dfrac{\partial\Psi_k}{\partial r}=0$. By differentiating the integral, it is seen that

$$\begin{cases}\nabla^2\Psi_k(xy)=0 & \text{if } x,y \text{ is in a dielectric,}\\ \qquad =4\pi ip\mu\lambda\Psi_k(xy), & \text{if } x,y \text{ is in conductor section } S_n \text{ and } n\neq k,\\ \qquad =4\pi ip\mu\lambda[\Psi_k(xy)-1], \text{ if } n=k. & \end{cases} \tag{72}$$

The analogy to $\sum\limits_{k=1}^{N}\phi_k(xy)=1$ is the relation

$$\sum_{k=1}^{N}\Psi_k(xy)=1. \tag{73}$$

It is also evident from the properties of M that $\Psi_k(xy)$ and $\dfrac{1}{\mu}\dfrac{\partial\Psi_k}{\partial n}$ are everywhere continuous. The integral equation for Ψ_k will determine its value in conductors without reference to the dielectric. However, it may be more convenient in practice to determine these functions by the methods of harmonic analysis. The foregoing differential equations and boundary conditions suffice to uniquely determine them at all points. Only $N-1$ of them need be computed, the remaining one being given by (73), which holds at all points. The condition that $\lim\limits_{r=\infty}\left(r\dfrac{\partial\Psi_k}{\partial r}\right)=0$ will (together with the other conditions) determine the value of $\Psi_k(\infty)$.

It is evident from this point of view that the case where a conductor extends to infinity presents no exception to the general form of solution, or the properties of these functions. The foregoing method of determining the N functions of Ψ_k, only $N-1$ of which are independent, amounts to a determination of the constant of integration $A(\infty)$ at the outset so as to satisfy the condition $\sum\limits_{n=1}^{N}I_n=0$ for all values of $\gamma, c_1, c_2, \ldots, c_N$.

The following alternative procedure amounts to an elimination of this constant $A(\infty)$, and leads immediately to a value of γ and the reactance of the system, in fact to all information of practical value. While not as symmetrical as the first method, it is essentially the same and possesses the advantage of leading naturally to a notation which reduces to a familiar one when the conductors are linear or the frequency low. It consists in finding the N independent functions $\omega_1(xy), \omega_2(xy), \ldots, \omega_N(xy)$, or what is the same thing, the N independent functions which will be denoted by the small letters $\psi_1(xy), \psi_2(xy), \ldots, \psi_N(xy)$ and defined by

$$\psi_k(xy) = \phi_k(xy) - \omega_k(xy). \tag{74}$$

The integral equation for $\omega_k(xy)$ shows that $\psi_k(xy)$ is the solution of the integral equation

$$\begin{aligned}\psi_k(xy) + ip\iint M(xy,\xi\eta)\lambda(\xi\eta)\psi_k(\xi\eta)d\xi d\eta \\ = ip\iint M(xy,\xi\eta)\lambda(\xi\eta)\phi_k(\xi\eta)d\xi d\eta.\end{aligned} \tag{75}$$

This serves to uniquely determine $\psi_k(xy)$ at all points in conductors and thence to define its value at all dielectric points. For the practical evaluation of $\psi_k(xy)$ by harmonic analysis, as the solution of a differential equation with certain boundary conditions, we may derive the following equations and conditions which $\psi_k(xy)$ must satisfy from a consideration of this integral equation.

By differentiation under the integral sign one finds from the properties of M that

(76)
- (a) $\nabla^2\psi_k(xy) = 0$ if x,y is in a dielectric,
  $= 4\pi ip\mu\lambda[\psi_k(xy) - \phi_k(xy)]$ if x,y is in a conductor,
- (b) ψ_k and $\dfrac{1}{\mu}\dfrac{\partial\psi_k}{\partial n}$ are continuous,
- (c) $\psi_k(xy) = D_k \log r(00,xy)$ at infinity, where D_k is a determinate constant not assignable.

These conditions are easily seen to be reciprocal with the integral equation, which may in fact be derived from them just as the equation for A and E was derived. Since

$$E(xy) = \gamma\phi(xy) - ipA(xy) = \gamma\sum_{k=1}^{N} c_k\phi_k(xy) - ipA(xy)$$

the vector potential $A(xy)$ at any point in the plane will be given by

$$ipA(xy) = ipA(\infty) + \sum_{k=1}^{N}[\gamma c_k - ipA(\infty)]\psi_k(xy) \tag{77}$$

and

$$I_n = \sum_{k=1}^{N}[\gamma c_k - ipA(\infty)]b_{nk}. \tag{78}$$

The definition of b_{nk} already given is easily seen to be equivalent to

$$b_{nk}=\frac{1}{4\pi ip}\int\frac{\partial\psi_k}{\partial n}ds_n \tag{79}$$

taken around any closed curve in the dielectric which encircles the nth group of conductors only.

6. *Coefficients of Resistance and Inductance.*

If it can be shown that the determinant $|b_{nk}|$ formed with the array of constants b_{nk} is never zero, then the equation (78) may be solved, giving the forms

$$\gamma c_s-ipA(\infty)=\sum_{k=1}^{N}z_{sk}I_k,\ (s=1,2,3,\ldots,N), \tag{80}$$

where

$$z_{sk}=R_{sk}+ipL_{sk} \tag{81}$$

and the real constants R_{sk} and L_{sk} are coefficients of resistance and inductance respectively. Since they are derived from a system of equations with a symmetrical determinant, the coefficients z_{nk} are symmetrical:

$$z_{nk}=z_{kn}\quad\text{or}\quad \begin{aligned}R_{nk}&=R_{kn},\\ L_{nk}&=L_{kn}.\end{aligned} \tag{82}$$

The determinant $|b_{nk}|$ cannot vanish for any finite value of p with finite conductivities. For we have shown that it is always possible to so choose the constant $A(\infty)$, without making each current I_n vanish, that the sum of the currents $\sum_{n=1}^{N}I_n$ shall be zero for every possible assignment of values to the N constants $c_1, c_2, \ldots, c_N$, and a direct contradiction to this fact may be obtained by assuming the determinant $|b_{nk}|$ to vanish. For in that event, it would be possible to choose the constants $c_1, c_2, \ldots, c_N$ not all zero, such that

$$\sum_{k=1}^{N}b_{nk}c_k=0,\quad (n=1,2,3,\ldots,N),$$

in which case each current I_n would be given by

$$I_n=-ipA(\infty)\sum_{k=1}^{N}b_{nk},$$

and if

$$\sum_{n=1}^{N}I_n=0=-ipA(\infty)\sum_{n=1}^{N}\sum_{k=1}^{N}b_{nk}$$

this will require that $A(\infty)=0$ since it has been shown that double the sum can never vanish. Consequently this would require that each current $I_1=I_2=\ldots=I_N=0$.

Therefore the equations (78) may always be solved and the coefficients R_{nk} and L_{nk} found. This being done, it will appear that the value of γ and the reactance of the line may be found without the necessity of computing the value of $A(\infty)$.

In case the frequency is very small, the function $\omega_k(xy)$ is almost identical with $\phi_k(xy)$ so that b_{nk} approaches zero with vanishing frequency if $n\neq k$, and b_{kk} approaches $\frac{1}{R_k(0)}$ where $R_k(0)$ is the direct current resistance per unit length of the kth conducting group. The determinant $|b_{nk}|$ approaches

$$(83)\qquad \begin{vmatrix} \frac{1}{R_1(0)} & 0 & 0 & \dots & 0 \\ 0 & \frac{1}{R_2(0)} & 0 & \dots & 0 \\ 0 & 0 & \frac{1}{R_3(0)} & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \frac{1}{R_N(0)} \end{vmatrix} .$$

The current distribution is practically uniform over the section of each conductor and R_{nn} approaches $R_n(0)$ while R_{nk} becomes negligible if $n\neq k$ with vanishing frequency.

This approximation neglects ψ_k because of the smallness of p, but the next approximation gives

$$\psi_k(xy)=ip\iint M(xy,\xi_k\eta_k)\lambda(\xi_k\eta_k)dS_k,$$

which gives

$$b_{nk}=-ip\iint dS_n\iint dS_k\lambda(x_ny_n)M(x_ny_n,\xi_k\eta_k)\lambda(\xi_k\eta_k)$$

$$=-ip\frac{L_{nk}}{R_n(0)R_k(0)} \qquad \text{if } n\neq k.$$

And when $n=k$

$$b_{nn}=\frac{1}{R_n(0)}-ip\iint dS_n\iint dS_k\lambda(x_ny_n)M(x_ny_n,\xi_n\eta_n)\lambda(\xi_n\eta_n)=\frac{1}{R_n(0)}-ip\frac{L_{nn}}{R_n(0)^2}.$$

The equations of definition of current (78) become

$$(84)\qquad I_n=\frac{\lambda c_n-ipA(\infty)}{R_n(0)}-ip\sum_{k=1}^{N}\frac{\lambda c_k-ipA(\infty)L_{nk}}{R_k(0)R_n(0)},\ (n=1,2,3,\dots,N).$$

The first approximation to a solution is

$$\frac{\gamma c_k-ipA(\infty)}{R_k(0)}=I_k,$$

and placing this value in the above summation gives

$$\gamma c_n - ipA(\infty) = R_n(0)I_n + ip\sum_{k=1}^{N} L_{nk}I_k, \tag{85}$$

which shows that for low frequencies $R_{nn}=R_n(0)$ and $R_{nk}=0$ $(n\neq k)$ while

$$L_{nk}=L_{kn}=R_n(0)R_k(0)\iint\lambda(x_ny_n)dS_n\iint\lambda(\xi_k\eta_k)dS_k M(x_ny_y,\xi_k\eta_k)$$

$$=\frac{\iint\lambda(x_ny_n)dS_n\iint\lambda(\xi_k\eta_k)dS_k M(x_ny_n,\xi_k\eta_k)}{\iint\lambda(x_ny_n)dS_n\iint\lambda(\xi_k\eta_k)\,dS_k},$$

If λ is constant over S_n and S_k then

$$L_{nk}=\frac{\iint M(x_ny_n,\xi_n\eta_n)dx_ndy_nd\xi_k\gamma d\eta_k}{S_nS_k}.$$

If further there are no ferromagnetic materials in the system, L_{nk} reduces to twice the negative of the geometric mean distance of the sections S_n and S_k. The approximations have been carried far enough to show therefore that the functions of frequency R_{nk} and L_{nk} reduce (for low frequencies) to their ordinary values which are familiar in the discussion of linear circuits.

V. The Propagation Constant γ. Attenuation and Phase Velocity

1. *Determinantal Equation for γ.*

The attenuation b and phase velocity V are determined by the equation (1) when the complex propagation constant γ is known. This must be so chosen as to render compatible the two systems of equations (29′) and (67).

$$\gamma I_n=\sum_{k=1}^{N-1}\alpha_n(c_k-c_N)\,,\; I_n=\gamma\sum_{k=1}^{N-1}\beta_{nk}(c_k-c_N),\quad (n=1,2,3,\ldots,N),$$

where

$$\alpha_{nk}=\alpha_{kn},\ \sum_{k=1}^{N}\alpha_{nk}=0\,,\quad \beta_{nk}=\beta_{kn},\ \sum_{k=1}^{N}\beta_{kn}=0.$$

Eliminating I_n gives $N-1$ linear homogeneous relations of the type

$$\sum_{k=1}^{N-1}(\alpha_{nk}-\gamma^2\beta_{nk})(c_k-c_N)=0,\quad (n=1,2,3,\ldots,N-1). \tag{86}$$

They will be compatible if and only if γ^2 is a root of the algebraic equation of degree $N-1$ in γ^2, obtained by placing the determinant of their coefficients equal to zero. That is, if

$$\delta_{nk}\equiv\alpha_{nk}-\gamma^2\beta_{nk} \tag{87}$$

then γ^2 must be the root of

$$\begin{vmatrix} \delta_{1,1} & \delta_{1,2} & \dots & \delta_{1,N-1} \\ \delta_{2,1} & \delta_{2,2} & \dots & \delta_{2,N-1} \\ \dots & \dots & \dots & \dots \\ \delta_{N-1,1} & \dots & \dots & \delta_{N-1,N-1} \end{vmatrix} = 0. \tag{88}$$

Since $\sum_{k=1}^{N} \delta_{nk} = 0$, one may, by adding all the other rows to the first row, and all the other columns to the first column, show that the equation is the same as if c_1, or any other c_n had been taken as a reference potential instead of c_N. Another form of this equation will be obtained if one makes *all* of the N equations (80), namely,

$$\gamma c_s - ipA(\infty) = \sum_{k=1}^{N} z_{sk} I_k, \quad (s=1, 2, 3, \dots, N),$$

compatible with *all* of the N equations (20), Section III:

$$\gamma^2 I_n = \sum_{s=1}^{N} a_{ns} \gamma c_s.$$

Since

$$\sum_{s=1}^{N} a_{ns} = 0, \quad (n=1, 2, 3, \dots, N),$$

this may be multiplied by $ipA(\infty)$ and subtracted from the above, giving

$$\gamma^2 I_n = \sum_{s=1}^{N} a_{ns}[\gamma c_s - ipA(\infty)], \quad (n=1, 2, 3, \dots, N). \tag{89}$$

Substituting in these equations the value of $\gamma c_s - ipA(\infty)$ from the above gives N linear homogeneous relations that must exist between the N currents, namely, where

$$\gamma^2 I_n = \sum_{k=1}^{N} z_{nk} I_k, \quad (n=1, 2, 3, \dots, N), \tag{90}$$

where

$$z_{nk} \equiv \sum_{s=1}^{N} a_{ns} z_{sk} \text{ so that } \sum_{n=1}^{N} z_{nk} = 0, \quad (k=1, 2, 3, \dots, N). \tag{91}$$

These equations will be compatible if and only if γ^2 is a root of

$$\Delta \equiv \begin{vmatrix} z_{11}-\gamma^2 & z_{12} & \dots & z_{1N} \\ z_{21} & z_{22}-\gamma^2 & \dots & z_{2N} \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ z_{N1} & z_{N2} & \dots & z_{NN}-\gamma^2 \end{vmatrix}$$

$$= -\gamma^2 \begin{vmatrix} z_{11}-\gamma^2 & z_{12} & \dots & z_{1N} \\ z_{21} & z_{22}-\gamma^2 & \dots & z_{2N} \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ z_{N-1,1} & z_{N-1,2} \dots & z_{N-1,N-1}-\gamma^2 & z_{N-1,N} \\ 1 & 1 \dots & 1 & 1 \end{vmatrix} = -\gamma^2 \sum_{k=1}^{N} \Delta_k = 0, \tag{92}$$

where Δ_k is the minor formed from Δ by suppressing the kth column and last row.

The second form of determinant is obtained by adding all the first $N-1$ rows of Δ to the last one and noting that $\sum_{n=1}^{N} z_{nk}=0\ (k=1, 2, 3, \ldots, N)$. The row whose elements are unity could thus be placed in any horizontal line of Δ. Thus $\Delta=0$ is the same equation as that obtained by combining the relation $\sum_{n=1}^{N} I_n=0$ with any $N-1$ of the relations (4). There will be $N-1$ values of γ^2 in addition to the value $\gamma^2=0$ which are roots of this equation. The latter leads to no possible solution, *i.e.*, a zero field everywhere.

If γ^2 is a root of $\Delta=0$ and any $N-1$ of the equations be solved for $N-1$ of the currents in terms of one of them, this leads to the relation

$$\frac{I_1}{\Delta_1}=\frac{I_2}{\Delta_2}=\ldots=\frac{I_N}{\Delta_N}=D, \tag{93}$$

or

$$I_n=D\Delta_n,\ (n=1,\ 2,\ 3, \ldots, N), \tag{94}$$

where D is an arbitrary constant. The set of equations

$$\gamma c_n-ipA(\infty)=\sum_{k=1}^{N} z_{nk}I_k=D\sum_{k=1}^{N} z_{nk}\Delta_k,\ (n=1,\ 2,\ 3, \ldots, N),$$

gives any potential difference

$$\gamma(c_n-c_N)=D\sum_{k=1}^{N}(z_{nk}-z_{Nk})\Delta_k=D\begin{vmatrix} z_{11}-\gamma^2 & z_{12}\ldots\ldots z_{1N} \\ z_{21}\ldots\ldots & \ldots\ldots z_{2N} \\ z_{n1,1}\ldots\ldots & \ldots\ldots z_{N-1,N} \\ z_{n1}-z_{N1}\ldots & \ldots z_{nN}-z_{NN} \end{vmatrix} \tag{95}$$

which shows that γ and the reactance of the line may be found without computing $A(\infty)$. Its value, however, may be found in terms of $\phi(\infty)$ by multiplying each equation of type (80) by $\phi_s(\infty)$ and summing with respect to s, noting that

$$\sum_{s=1}^{N}\phi_s(\infty)=1,\ \gamma\sum_{s=1}^{N}\phi_s(\infty)c_s=\gamma\phi(\infty);$$

this gives

$$E(\infty)=\gamma\phi(\infty)-ipA(\infty)=\sum_{n=1}^{N}\sum_{k=1}^{N}\phi_n(\infty)z_{nk}I_k=D\sum_{n=1}^{N}\sum_{k=1}^{N}\phi_n(\infty)z_{nk}\Delta_k. \tag{96}$$

Except for the fact that the terms z_{nk} are functions of the frequency, the equation (92) for γ is the same as in the case of linear circuits, and it reduces to the same form with vanishing frequency.

2. *Determination of Constants to fit Terminal Conditions.*

For each value of γ^2 which is a root of $\Delta=0$, there are two values $\pm\gamma$ corresponding to a forward and backward wave, each containing an arbitrary constant D. In general there will be $2N-2$ wave types and this number of constants D. It is evident that if the circuit equations involving the currents, potential differences and impedances of the terminal apparatus are written for all the independent modes of connection of the cylinders at the terminal $z=0$, and $z=l$, these with (94) and (95) will afford the necessary and sufficient number of relations to determine the $2N-2$ constants of type D and hence to co-pletely determine the field. There are $N-1$ independent modes of connecting the terminals of N cylinders at each end.

3. *Special Cases.*

(*a*) Case of Two Conductors:

$$\begin{aligned} &a_{11}=a_{22}=-a_{12}=-a_{21}=G_{12}+ipC_{12}, \\ &z_{12}=z_{21}\begin{cases} R_{12}=R_{21}, \\ L_{12}=L_{21}, \end{cases} \\ &z_{11}=-z_{21}=a_{11}z_{11}+a_{12}z_{21}=(G_{12}+ipC_{12})\ (z_{11}-z_{12}), \\ &z_{22}=-z_{12}=a_{21}z_{12}+a_{22}z_{22}=(G_{12}+ipC_{12})\ (z_{22}-z_{12}), \\ &\gamma^2=z_{11}-z_{12}=(G_{12}+ipC_{12})\ (z_{11}+z_{22}-2z_{12}), \\ &\gamma c_1-ipA(\infty)=z_{11}I_1+z_{12}I_2=(z_{11}-z_{12})I_1, \\ &\gamma c_2-ipA(\infty)=z_{21}I_1+z_{22}I_2=(z_{12}-z_{22})I_1, \\ &\gamma(c_1-c_2)=(z_{11}+z_{22}-2z_{12})I_1=(R+ipL)I_1, \end{aligned} \tag{97}$$

where R and L denote resistance and inductance of the *line* per unit length and are defined by

$$\begin{cases} R\equiv R_{11}+R_{22}-2R_{12}, \\ L\equiv L_{11}+L_{22}-2L_{12}, \end{cases}$$

$$\gamma^2=(G_{12}+ipC_{12})(R+ipL), \tag{98}$$

and where $\frac{1}{G_{12}}$ and C_{12} are the insulation resistance and capacity respectively between the two conductors per unit length.

(*b*) Uniform Proximity Effect.

If the sections of the conductors are all small compared to their distance apart, the general equation (36), Section IV, breaks up into N independent integral equations of the type:

$$\begin{aligned} &E(x_ny_n)+ip\iint M(x_ny_n,\xi_n\eta_n)\lambda(\xi_n\eta_n)E(\xi_n\eta_n)d\xi_n d\eta_n \\ &\quad=-ip\sum_k{}' M(x_ny_n,\xi_k\eta_k)\iint\lambda(\xi_k\eta_k)E(\xi_k\eta_k)d\xi_k d\eta_k+\gamma c_n-ipA(\infty) \end{aligned} \tag{99}$$

where $\sum_k{}'$ denotes that the term corresponding to $k=n$ is omitted from the summation. These terms may be written $ip\sum_k{}' L_{nk}I_k$ where

$$\begin{aligned} L_{nk}=L_{kn}&=M(x_ny_n,\xi_k\eta_k)=\text{a real constant} \\ &=-2\log r(x_ny_n,\xi_k\eta_k) \text{ if all materials are non-magnetic.} \end{aligned} \tag{100}$$

The integral on the left side of (99) is taken over the section of the nth conductor, the effect of the proximity of other conductors being uniform over this section and represented by the summation on the right side. The entire second member of (99) is constant, say D_n, over the section S_n. This equation may be called the "skin effect equation" since the proximity of other conductors does not affect the relative distribution of current in any one. Its solution will be of the form

$$E(x_n y_n) = D_n F_n(x_n y_n, ip). \tag{101}$$

Multiplying by $\lambda_n dx_n dy_n$ and integrating over the section S_n gives

$$I_n = D_n \iint \lambda(x_n y_n) F_n(x_n y_n, ip) dx_n dy_n. \tag{102}$$

Since this integral can never vanish (equation 58), the real functions of frequency R_{nn} and L_{nn} may be defined by

$$R_{nn} + ipL_{nn} = \frac{1}{\iint \lambda_n F_n(x_n y_n, ip) dS_n}. \tag{103}$$

If we also define R_{nk} to be zero if $n \neq k$ and replace D_n in (101) by its value

$$\gamma c_n - ipA(\infty) - ip\sum_k{}' L_{nk} I_k$$

it reduces to the general form (80) Section IV:

$$\gamma c_2 - ipA(\infty) = \sum_{k=1}^{N} (R_{nk} + ipL_{nk}) I_k = \sum_{k=1}^{N} z_{nk} I_k. \tag{104}$$

The completion of the problem is the same as that outlined in general in this section; the coefficients of leakage and capacity must be found in order to write out the determinant for γ^2.

VI. Energy Relations and Mean Values

Suppose ϕ, A, and E have been found in the forms $\phi' + i\phi''$, $A' + iA''$, $E' + iE''$ where ϕ', ϕ'', etc., are real functions of x, y and p. The instantaneous value of any of the above quantities is found by supplying the exponential factor $e^{ipt - \gamma z} = e^{-bz} \cdot e^{ip(t - \frac{z}{v})}$ and then taking the real part.

If $P = P' + iP''$ and $Q = Q' + iQ''$ are any two typical quantities of this kind, the time average of the product of the instantaneous (real) values of P and Q is found to be

$$e^{-2bz}\left(\frac{P'Q' + P''Q''}{2}\right).$$

The time averages which will be dealt with in this section presuppose a single type of wave corresponding to one value of γ only, and obviously do not apply when several types of wave coexist since it is the field components that are additive and not the energy.

Consider the medium between two planes perpendicular to the z axis at z and $z+dz$. Denote the time average of the rate of flow of energy through the plane $z=$constant, in the positive z direction by $e^{-2bz}\overline{F}$. There is no flow of energy in this direction in conductors; it is all in the dielectric and flows laterally into the conductors. Denote the mean rate of heating in all the conductors between these two planes by $e^{-2bz}\,dz\overline{W}_c$, that in the dielectric by $e^{-2bz}\,dz\overline{W}_d$. Also let $e^{-2bz}dz\overline{U}$ and $e^{-2bz}dz\overline{T}$ denote respectively the time average of the electrical energy in the dielectric, and of the electrokinetic energy between these two planes. The former is localized in the dielectric only.

The definitions of these quantities lead to the forms:

$$\overline{F}=\frac{1}{8\pi}\iint\left[\frac{\partial\phi}{\partial x}\cdot\frac{\partial A'}{\partial x}+\frac{\partial\phi}{\partial y}\cdot\frac{\partial A'}{dy}+\frac{\partial\phi''}{\partial x}\cdot\frac{\partial A''}{\partial x}+\frac{\partial\phi''}{\partial y}\cdot\frac{\partial A''}{\partial y}\right]dS,$$

$$\overline{W}_d=\iint\frac{\lambda}{2}\left[\left(\frac{\partial\phi'}{\partial x}\right)^2+\left(\frac{\partial\phi'}{\partial y}\right)^2+\left(\frac{\partial\phi''}{\partial x}\right)^2+\left(\frac{\partial\phi''}{\partial y}\right)^2\right]dS,$$

$$\overline{U}=\frac{1}{8\pi}\iint\frac{k}{2c^2}\left[\left(\frac{\partial\phi'}{\partial x}\right)^2+\left(\frac{\partial\phi'}{\partial y}\right)^2+\left(\frac{\partial\phi''}{\partial x}\right)^2+\left(\frac{\partial\phi''}{\partial y}\right)^2\right]dS,$$

where the integration is over all dielectric sections in the plane;

$$\overline{W}_c=\iint\lambda\left(\frac{E'^2+E''^2}{2}\right)ds$$

over all conducting sections;

$$T=\frac{1}{8\pi}\iint\frac{1}{2\mu}\left[\left(\frac{\partial A'}{\partial x}\right)^2+\left(\frac{\partial A'}{\partial y}\right)^2+\left(\frac{\partial A''}{\partial x}\right)^2+\left(\frac{\partial A''}{\partial y}\right)^2\right]dS \quad \text{(over the entire plane)}$$

$$=\frac{1}{2}\iint\lambda\left(\frac{E'A'+E''A''}{2}\right)dS$$

over all conducting sections.

The justification of these expressions of $\overline{W}_d$ and $\overline{U}$ lies in the fact that $\left(\lambda+\frac{ipk}{4\pi c^2}\right)E^2$ is infinitesimal compared to

$$\left(\lambda+\frac{ipk}{4\pi c^2}\right)\left[\left(\frac{\partial\phi}{\partial x}\right)^2+\left(\frac{\partial\phi}{\partial y}\right)^2\right].$$

The formula is correct to first approximations even in those cases where the constant $E(\infty)$ is not zero, for a second approximation would make E vanish at infinity without affecting the first approximation to current distribution in conductors, or to x,y components of leakage and capacity current in dielectrics.

The differential equations and boundary conditions which ϕ and A satisfy lead, by simple transformations of the above integrals, to the following values $\left(\text{if } \gamma \equiv b + \frac{ip}{V}\right)$:

$$\overline{F} = \frac{1}{2}\sum_{k=1}^{N}(c_k'I_k' + c_k''I_k''),$$

$$\overline{W}_d = \frac{1}{2}\sum_{k=1}^{N}[c_k'(G_k' - pQ_k'') + c_k''(G_k'' + pQ_k')],$$

$$\overline{U} = \frac{1}{4p}\sum_{k=1}^{N}[c_k'(G_k'' + pQ_k') - c_k''(G_k' - pQ_k'')],$$

$$\overline{W}_c = \frac{1}{2}\sum_{k=1}^{N}\left[b(c_k'I_k' + c_k''I_k'') + \frac{p}{V}(c_k'I_k' - c_k''I_k')\right],$$

$$\overline{T} = \frac{1}{4p}\sum_{k=1}^{N}\frac{p}{V}[c_k'I_k' + c_k''I_k'') - b(c_k'I_k'' - c_k''I_k'].$$

Since

$$\gamma I_k = g_k + ipq_k \begin{cases} bI_k' - \frac{p}{V}I_k'' = G_k' - pQ_k'', \\ \frac{p}{V}I_k' + bI_k'' = G_k'' + pQ_k', \end{cases}$$

the above expressions for $\overline{W}_d$ and $\overline{U}$ may be put in the forms

$$\overline{W}_d = \frac{1}{2}\sum_{k=1}^{N}\left[b(c_k'I_k' + c_k''I_k'') - \frac{p}{V}(c_k'I_k'' - c_k''I_k')\right],$$

$$\overline{U}_d = \frac{1}{4p}\sum_{k=1}^{N}\left[\frac{p}{V}(c_k'I_k' + c_k''I_k'') + b(c_k'I_k'' - c_k''I_k')\right].$$

From these expressions it is evident that

$$W_c - W_d = \frac{2p^2}{bV}(\overline{U} - \overline{T}),$$

$$\overline{F} = \frac{\overline{W}_c + \overline{W}_d}{2b} = V(\overline{U} + \overline{T}).$$

That is, the mean flow of energy $e^{-2bz}\overline{F}$ through the plane z = constant is equal to the phase velocity V, multiplied by the density per unit length along z of the total electric and magnetic energy $e^{-2bz}(\overline{U}+\overline{T})$. The last equation may also be put in the form

$$-\frac{d}{dz}(e^{-2bz}\overline{F})dz=e^{-2bz}(\overline{W}_c+\overline{W}_d)dz,$$

which states that the excess of mean energy flow into the medium between the planes z and $z+dz$ is equal to the mean rate of dissipation of energy between these planes.

It may be noted that the dissipation of energy $\overline{W}_d$ in the dielectric may be of the same order of magnitude as that in conductors $\overline{W}_c$. If these two should be equal, then the mean electrical and magnetic energies would be equal as in the case of homogeneous waves.

The relation

$$\gamma c_k-ipA(\infty)=\sum_{n=1}^{N}(R_{nk}+ipL_{nk})I_n$$

enables us to write

$$\overline{W}_c=\sum_{n=1}^{N}\sum_{k=1}^{N}R_{nk}\left(\frac{I_n'I_k'+I_n''I_k''}{2}\right),$$

$$\overline{T}=\frac{1}{2}\sum_{n=1}^{N}\sum_{k=1}^{N}L_{nk}\left(\frac{I_n'I_k'+I_n''I_k''}{2}\right).$$

The relation

$$\gamma I_k=g_k+ipq_k=\sum_{n=1}^{N-1}(G_{nk}+ipC_{nk})\,(c_N-c_n)$$

leads to

$$\overline{W}_d=\sum_{n=1}^{N-1}\sum_{k=1}^{N-1}G_{nk}\left[\frac{(c_n'-c_N')\,(c_N'-c_k')+(c_n''-c_N'')(c_N''-c_k'')}{2}\right],$$

$$\overline{U}=\frac{1}{2}\sum_{n=1}^{N-1}\sum_{k=1}^{N-1}C_{nk}\left[\frac{(c_n'-c_N')\,(c_N'-c_k')+(c_n''-c_N'')\,(c_N''-c_k'')}{2}\right].$$

If the new real quantities R_n and L_n be defined by the equation

$$R_n+ipL_n\equiv\frac{\sum_{k=1}^{N}z_{nk}\Delta_k}{\Delta_n}=\frac{\sum_{k=1}^{N}(R_{nk}+ipL_{nk})\Delta_k}{\Delta_n},$$

then

$$\gamma c_n-ipA(\infty)=(R_n+ipL_n)I_n=\sum(R_{nk}+ipL_{nk})I_k,$$

and this gives the simpler forms

$$\overline{W}_c=\sum_{n=1}^{N}R_n\left[\frac{I_n'^2+I_n''^2}{2}\right],$$

$$\overline{T}=\frac{1}{2}\sum_{n=1}^{N}L_n\left[\frac{I_n'^2+I_n''^2}{2}\right].$$

Such expressions may, however, be misleading on account of their simplicity. For example the heating in the nth conductor is not

$$R_n\left[\frac{I_n'^2+I_n''^2}{2}\right]$$

in general, but is

$$\iint \lambda(E'^2+E''^2)dS_n.$$

VII. Examples and Applications

1. *Case of Two Conductors With Circular Symmetry.*

The only interest that can be attached here to such a well-known problem is to show that the integral equation for say ω_1 does uniquely determine ω_1 and hence $\psi_1=\phi_1-\omega_1$ without reference to any boundary conditions. Hence the simplest case is taken of a circular cylinder of radius r_1 with a return conductor in the shape of a concentric shell of inner and outer radii r_2 and r_3 respectively, both conductors being non-magnetic and having the same conductivity, and the dielectric being air.

In this case if $h^2=-4\pi ip\lambda$ we have:

$r_1>r\geqq 0$	$r_2>r>r_1$	$r_3>r>r_2$	$r>r_3$
$\phi_1=1$	$\dfrac{\log \frac{r}{r_2}}{\log\frac{r_1}{r_2}}$	0	0
$\psi_1=1+C_1J_0(hr)$	$C_2+C_3\log r$	$C_4J_0(hr)+C_5K_0(hr)$	$C_6\log r$

These forms are required by the differential equation for ψ_1 (76a). Writing out the equations (76b) and (76c), the boundary conditions for ψ, and eliminating the constants C_2, C_3 and C_6, which correspond to the dielectric regions, let

$$P(x)\equiv J_0(x)-\left(\log\frac{x}{h}\right)xJ_0'(x),$$

$$Q(x)\equiv K_0(x)-\left(\log\frac{x}{h}\right)xK'(x).$$

This gives the three equations

$$-P(x_1)C_1+P(x_2)C_4+Q(x_2)C_5=1,$$

$$-x_1J_0'(x_1)C_1+x_2J_0'(x_2)C_4+x_2K_0'(x_2)C_5=0,$$

$$0C_1+P(x_3)C_4+Q(x_3)C_5=0.$$

The determinant of these three equations cannot vanish for any real value of the frequency, hence they determine C_1, C_4 and C_5.

To show that the integral equation for ω_1 leads to the same values of these three constants directly, express $M(rr')$ in the form

$$M(rr') = -2 \log R(rr') = -2\left[a_0(rr') - \sum_{n=1}^{\infty} \frac{a_n(rr') \cos n(\theta-\theta')}{n} \right],$$

where

$$\begin{aligned} a_0(rr') &= \log r \quad \text{if } r'<r, \\ &= \log r' \quad \text{if } r'>r, \\ a_n(rr') &= \left(\frac{r'}{r}\right)^n \text{ if } r'<r, \\ &= \left(\frac{r}{r'}\right)^n \text{ if } r'>r. \end{aligned}$$

Since ω_1 is a function of r only and not of θ in both conductors, the integral equation for ω_1 reduces to

$$\omega_1(r) + h_2 \int a_0(rr') r' \omega_1(r') dr' = \phi_1(r) = \begin{cases} 1 \text{ if } r<r_1, \\ 0 \text{ if } r>r_2, \end{cases}$$

where the integral is taken over both conductors. This equation has one solution and one only, and if we assume the form

$$\begin{aligned} \omega_1(r) &= -C_1 J_0(hr), & r<r_1, \\ &= -C_4 J_0(hr) - C_5 K_0(hr), & r_2<r<r_3, \end{aligned}$$

and substitute this form in the above equation, it requires that

$$\begin{aligned} &-P(x_1)C_1 + [P(x_2)-P(x_3)]C_4 + [Q(x_2)-Q(x_3)]C_5 = 1, \\ &-x_1 J_0(x_1) C_1 + x_2 J_0'(x_2) C_4 + x_2 K_0'(x_2) C_5 = 0, \\ &P(x_3) C_4 + Q(x_3) C_5 = 0. \end{aligned}$$

By adding the third of these to the first, the same set of equations for C_1, C_4, and C_5 is obtained as was found by satisfying the differential equations and boundary conditions. This shows that the integral equation will uniquely determine the field without any reference to boundary conditions.

It may be noted that $E(\infty) = \gamma\phi(\infty) - ipA(\infty)$ cannot vanish in this problem.

In evaluating the integrals to obtain these equations, use may be made of the fact that

$$\frac{1}{x}\frac{d}{dx}[xJ_0(x)] + J_0(x) = 0, \qquad xJ_0(x) = -\frac{d}{dx}[xJ_0'(x)],$$

$$x \log xJ_0(x) = -\frac{d}{dx}[x \log xJ_0'(x)] + \frac{dJ_0(x)}{dx},$$

with similar formulas for $K_0(x)$.

2. *Case of Two Circular, Cylindrical Conductors of Different Conductivity, Permeability, and Radius, Surrounded by a Homogeneous, Slightly Conducting Dielectric* (*See Fig. 2, p. 202*). *Mixed Method.*

In this example the mixed method of Section IV will be employed:

$$(106)\quad \begin{cases} h_1 \equiv \sqrt{2\pi p \mu_1 \lambda_1}\,(1-i), & x_1 = h_1 a_1 = \sqrt{\dfrac{2p\mu_1}{R_1(0)}}\,(1-i), \\ h_2 \equiv \sqrt{2\pi p \mu_2 \lambda_2}\,(1-i), & x_2 = h_2 a_2 = \sqrt{\dfrac{2p\mu_2}{R_2(0)}}\,(1-i). \end{cases}$$

Assume the forms:

$$(107)\quad \begin{cases} A = \dfrac{\gamma c_1}{ip} - I_1\left[C_0 J_0(h_1 r_1) + \sum\limits_{n=1}^{\infty} C_n J_n(h_1 r_1)\cos n\,\theta_1\right] \text{if } P(r_1\theta_1) \text{ is in No. 1,} \\ A = \dfrac{\gamma c_2}{ip} - I_2\left[D_0 J_0(h_2 r_2) + \sum\limits_{n=1}^{\infty} D_n J_n(h_2 r_2)\cos n\,\theta_2\right] \text{if } P(r_2\theta_2) \text{ is in No. 2.} \end{cases}$$

The equation which must be identically true when the point P is within either circle is (51):

$$(108)\quad A(\infty) = \frac{a_1}{2\pi}\int_o^{2\pi} d\theta_1'\left[A(a_1\theta_1')\frac{\partial}{\partial a_1}\log R(PP_1') - \frac{\log R(PP_1')}{\mu_1}\left(\frac{\partial A}{\partial a_1}\right)_{(a_1\theta_1')}\right]$$
$$+\frac{a_2}{2\pi}\int_o^{2\pi} d\theta_2'\left[A(a_2\theta_2')\frac{\partial}{\partial a_2}\log R(PP_2') - \frac{\log R(PP_2')}{\mu_2}\left(\frac{\partial A}{\partial a_2}\right)_{(a_2\theta_2')}\right].$$

In Case 1, P is inside circle No. 1,

$$\log R(PP_1') = \log a_1 - \sum_{k=1}^{\infty}\left(\frac{r_1}{a_1}\right)^k \frac{\cos k(\theta_1 - \theta_1')}{k},$$

$$\log R(PP_2') = \log \rho_2 - \sum_{k=1}^{\infty}\left(\frac{a_2}{\rho_2}\right)^k \frac{\cos k(\theta_2' - \alpha_2)}{k}.$$

In Case 2, P is within circle No. 2,

$$\log R(PP_1') = \log \rho_1 - \sum_{k=1}^{\infty}\left(\frac{a_1}{\rho_1}\right)^k \frac{\cos k(\theta_1' - \alpha_1)}{k},$$

$$\log R(PP_2') = \log a_2 - \sum_{k=1}^{\infty}\left(\frac{r_2}{a_2}\right)^k \frac{\cos k\,(\theta_2 - \theta_2')}{k}.$$

If the expressions (107) for A be used in (108) and the terms be written out for Case 1, the corresponding equation for Case 2 may be formed from this by a permutation of the subscripts 1 and 2 and of the constants C_n and D_n. In Case 1, the equation (108) becomes

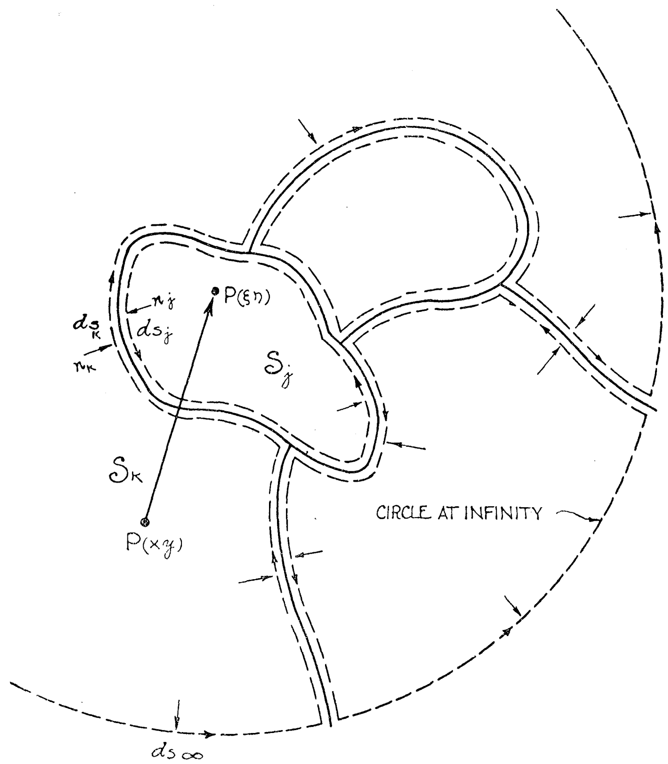


Fig. 1
Place section of cylindrical media (z=constant) showing positive directions of normals and line elements in each region.

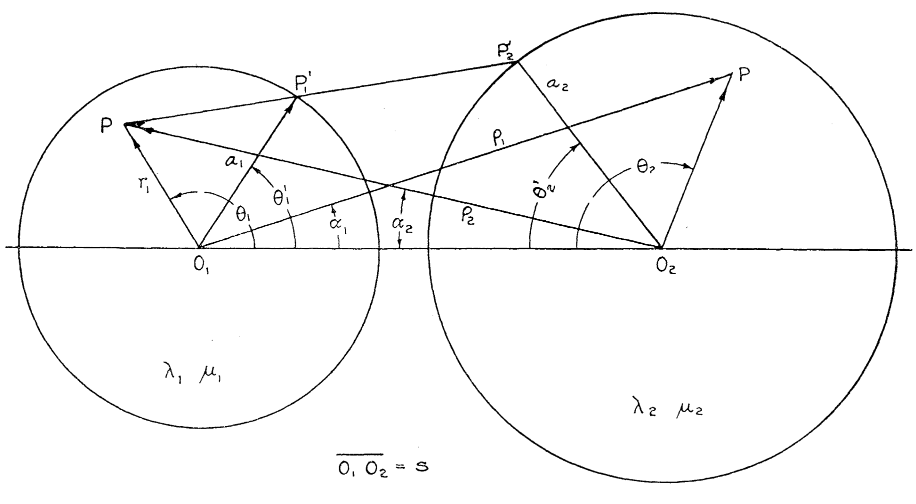


Fig. 2
Sections of two cylindrical conductors.

$$A(\infty)=\frac{I_1}{2\pi}\int_o^{2\pi}d\theta_1'\left\{\left[\frac{\gamma c_1}{ipI_1}-C_0J_0(x_1)-\sum_{n=1}^{\infty}C_nJ_n(x_1)\cos n\,\theta_1'\right].\right.$$

$$.\left[1+\sum_{k=1}^{\infty}\left(\frac{r_1}{a_1}\right)^k\cos k(\theta_1-\theta_1')\right]$$

$$+\frac{x_1}{\mu_1}\left[C_0J_0'(x_1)+\sum_{n=1}^{\infty}C_nJ_n'(x_1)\cos n\,\theta_1'\right].$$

$$\left..\left[\log a_1-\sum_{k=1}^{\infty}\left(\frac{r_1}{a_1}\right)^k\frac{\cos k(\theta_1-\theta_1')}{k}\right]\right\}$$

$$+\frac{I_2}{2\pi}\int_o^{2\pi}d\theta_2'\left\{\left[\frac{\gamma c_2}{I_2}-D_0J_0(x_2)-\sum_{n=1}^{\infty}D_nJ_n(x_2)\cos n\,\theta_2'\right].\right.$$

$$.\left[-\sum_{k=1}^{\infty}\left(\frac{a_2}{\rho_2}\right)^k\cos k(\theta_2'-\alpha_2)\right]$$

$$+\frac{x_2}{\mu_2}\left[D_0J_0'(x_2)+\sum_{n=1}^{\infty}D_nJ_n'(x_2)\cos n\,\theta_2'\right].$$

$$\left..\left[\log \rho_2-\sum_{k=1}^{\infty}\left(\frac{a_2}{\rho_2}\right)^k\frac{\cos k(\theta_2'-\alpha_2)}{k}\right]\right\};$$

or, after integrating with respect to θ_1' and θ_2',

$$A(\infty)=\frac{\gamma c_1}{ip}-I_1C_0\left[J_0(x_1)-\frac{x_1J_0'(x_1)\log a_1}{\mu_1}\right]$$

(109)
$$-\frac{I_1}{2}\sum_{n=1}^{\infty}C_n\left[J_n(x_1)+\frac{x_1}{\mu_1}\frac{J_n'(x_1)}{n}\right]\left(\frac{r_1}{a_1}\right)^n\cos n\theta_1$$

$$+I_2D_0\frac{x_2J_0'(x_2)\log\rho_2}{\mu_2}+\frac{I_2}{2}\sum_{k=1}^{\infty}D_k\left[J_k(x_2)-\frac{x_2J_k'(x_2)}{\mu_2 k}\right]\left(\frac{a_2}{\rho_2}\right)^k\cos k\alpha_2.$$

Introducing into (109) the expansions of the triangle

$$\log\rho_2=\log s-\sum_{n=1}^{\infty}\left(\frac{a_1}{s}\right)^n\left(\frac{r_1}{a_1}\right)^n\frac{\cos n\theta_1}{n},$$

$$\left(\frac{a_2}{\rho_2}\right)^k\cos k\alpha_2=\left(\frac{a_2}{s}\right)^k\left\{1+\sum_{n=1}^{\infty}\frac{n+k-1!}{n!\;k-1!}\left(\frac{a_1}{s}\right)^n\left(\frac{r_1}{a_1}\right)^n\cos n\,\theta_1\right\},$$

we obtain, after inverting the order of summation of the double sum,

$$0=\frac{\gamma c_1-ipA(\infty)}{ip}-I_1C_0\left[J_0(x_1)-\frac{x_1J_0'(x_1)\log a_1}{\mu_1}\right]+I_2D_0\frac{x_2J'(x_2)}{\mu_2}\log s$$

$$+\frac{I_2}{2}\sum_{k=1}^{\infty}D_k\left[J_k(x_2)-\frac{x_2J_k'(x_2)}{\mu_2k}\right]\left(\frac{a_2}{s}\right)^k$$

$$-\sum_{n=1}^{\infty}\left(\frac{r_1}{a}\right)^n\cos n\theta_1\left\{\frac{I_1C_n}{2}\left[J_n(x_1)+\frac{x_1J_n'(x_1)}{\mu_1n}+\frac{I_2D_0x_2J_0'(x_2)}{\mu_2n}\left(\frac{a_1}{s}\right)^n\right.\right.$$

$$\left.\left.-I_2\sum_{k=1}^{\infty}\frac{n+1-1!}{n!\,k-1!}\left(\frac{a_1}{s}\right)^n\left(\frac{a_2}{s}\right)^k\left[J_k(x_2)-\frac{x_2J_k(x_2)}{\mu_2k}\right]\frac{D_k}{2}\right\}.$$

Since this must be identically true whatever the values of r_1,θ_1, the coefficient of r_1^n must vanish for every value of $n=0, 1, 2, 3,\ldots\infty$. If for brevity one writes

$$(110)\qquad\begin{cases}Q_n\equiv-\dfrac{D_n}{4\mu_2}\left(\dfrac{a_2}{s}\right)^n\left[\mu_2nJ_n(x_2)-x_2J_n'(x_2)\right],\\[2ex] P_n\equiv-\dfrac{C_n}{4\mu_1}\left(\dfrac{a_1}{s}\right)^n\left[\mu_1nJ_n(x_1)-x_1J_n'(x_1)\right],\end{cases}$$

$$(n=0, 1, 2, 3,\ldots\infty),$$

this identity gives the set of equations

$$(111)\begin{cases}-I_2\displaystyle\sum_{k=1}^{\infty}\frac{Q_k}{k}+2I_1P_0\left[\log a_1-\frac{\mu_1J_0(x_1)}{x_1J_0'(x_1)}\right]=-2I_2Q_0\log s-\frac{\gamma c_1-ipA(\infty)}{2ip},\\[2ex] I_2\displaystyle\sum_{k=1}^{\infty}\frac{n+k-1!}{n-1!\,k!}Q_k-\left(\frac{s}{a_1}\right)^{2n}\frac{\mu_1nJ_n(x_1)+x_1J_n'(x_1)}{\mu_1nJ_n(x_1)+x_1J_n'(x_1)}.P_nI_1=-2Q_0I_2,\end{cases}$$

$$(n=1, 2, 3,\ldots\infty).$$

The corresponding set

$$(112)\begin{cases}-I_1\displaystyle\sum_{k=1}^{\infty}\frac{P_k}{k}+I_22Q_0\left[\log a_2-\frac{\mu_2J_0(x_2)}{x_2J_0'(x_2)}\right.=-2I_1P_0\log s-\frac{\gamma c_2-ipA(\infty)}{2ip},\\[2ex] I_1\displaystyle\sum_{k=1}^{\infty}\frac{n+k-1!}{n-1!\,k!}P_k-\left(\frac{s}{a_2}\right)^{2n}\frac{\mu_2nJ_n(x_2)+x_2J_n'(x_2)}{\mu_2nJ_n(x_2)-x_2J_n'(x_2)}Q_nI_2=-2P_0I_1,\end{cases}$$

$$(n=1, 2, 3,\ldots\infty),$$

is obtained from this identity in Case 2 where the point P is within the 2nd circle, by permuting the subscripts and constants. Since $E=\gamma c_1-ipA$ in cylinder No. 1

$$I_1 = ipI_1C_0 2\pi\lambda_1 \int_o^{a_1} rJ_0(h_1 r)dr = \frac{I_1 C_0 x_1 J_0'(x_1)}{2\mu_1} = 2I_1P_0,$$

or

$$2P_0 = +1.$$

Similarly

$$I_2 = ipI_2D_0 2\pi\lambda_2 \int_o^{a_2} rJ_0(h_2 r)dr = \frac{I_2 D_0 x_2 J_0'(x_2)}{2\mu_2} = +I_2 2Q_0.$$

Hence

$$2P_0 = 2Q_0 = +1.$$

Placing these values of P_0 and Q_0 in (111) and (112), and placing $I_2 = -I_1$ gives

$$(113)\qquad \begin{cases} \displaystyle\sum_{k=1}^{\infty} \frac{n+k-1!}{n-1!\ k!} Q_k + \left(\frac{s}{a_1}\right)^{2n} \frac{\mu_1 n J_n(x_1) + x_1 J_n'(x_1)}{\mu_1 n J_n(x_1) - x_1 J_n'(x_1)} P_n = -1, \\ \displaystyle\sum_{k=1}^{\infty} \frac{n+k-1!}{n-1!\ k!} P_k + \left(\frac{s}{a_2}\right)^{2n} \frac{\mu_2 n J_n(x_2) + x_2 J_n'(x_2)}{\mu_2 n J_n(x_2) - x_2 J_n'(x_2)} Q_n = -1, \end{cases}$$

$$(n = 1, 2, 3, \ldots \infty).$$

The two equations for $n = 0$ become

$$(114)\qquad \begin{cases} \gamma c_1 - ipA(\infty) = 2ipI_1 \left[\log \dfrac{s}{a_1} + \dfrac{\mu_1 J_0(x_1)}{x_1 J_0'(x_1)} - \displaystyle\sum_{k=1}^{\infty} \frac{Q_k}{k} \right], \\ \gamma c_2 - ipA(\infty) = -2ipI_1 \left[\log \dfrac{s}{a_2} + \dfrac{\mu_2 J_0(x_2)}{x_2 J_0'(x_2)} - \displaystyle\sum_{k=1}^{\infty} \frac{P_k}{k} \right], \end{cases}$$

or, by subtraction

$$(115)\qquad \gamma(c_1 - c_2) = 2ipI_1 \left\{ \frac{\mu_1 J_0(x_1)}{x_1 J_0'(x_1)} + \frac{\mu_2 J_0(x_2)}{x_2 J_0'(x_2)} + \log \frac{s^2}{a_1 a_2} - \sum_{k=1}^{\infty} \left[\frac{P_k + Q_k}{k} \right] \right\}.$$

The second member of this equation will be known when P_k and Q_k for $k = 1, 2, 3, \ldots \infty$, have been found as solutions of the equations (113). The equation of conservation of electricity is

$$(116)\qquad \gamma I_1 = (G_{12} + ipC_{12})\ (c_1 - c_2),$$

where $\dfrac{1}{G_{12}}$ and C_{12} are respectively the insulation resistance and capacity between the two cylinders per unit length. The value of γ found by eliminating $\dfrac{(c_1 - c_2)}{I_1}$ between (115) and (116) is

$$(117)\qquad \gamma^2 = (G_{12} + ipC_{12}) 2ip \left\{ \frac{\mu_1 J_0(x_1)}{x_1 J_0'(x_1)} + \frac{\mu_2 J_0(x_2)}{x_2 J_0'(x_2)} + \log \frac{s^2}{a_1 a_2} - \sum_{k=1}^{\infty} \left[\frac{P_k + Q_k}{k} \right] \right\}.$$

The resistance R and inductance L per unit length of the line are then given by equation (98),

$$(118)\quad R+ipL=\frac{\gamma^2}{G+ipC}=2ip\left\{\frac{\mu_1 J_0(x_1)}{x_1 J_0'(x_1)}+\frac{\mu_2 J_0(x_2)}{x_2 J_0'(x_2)}+\log\frac{s^2}{a_1a_2}-\sum_{n=1}^{\infty}\left[\frac{P_n+Q_n}{n}\right]\right\}.$$

The insulation resistance $\frac{1}{G}$ and capacity C between the two wires per unit length are given by

$$(119)\quad G=\frac{4\pi\lambda d}{2\log\frac{1}{b_1b_2}},\qquad C=\frac{k}{2c^2\log\frac{1}{b_1b_2}},\qquad c=3(10)^{10}$$

in electromagnetic e.g.s. units, λ_d being the conductivity of the dielectric in these units and k its dielectric constant in c.g.s. electrostatic units. The pure numbers b_1 and b_2 are both positive and less than 1 and satisfy the equations

$$(120)\quad \begin{cases} a_1b_1(s-a_2b_2)=a_1^2, \\ a_2b_2(s-a_1b_1)=a_1^2, \end{cases}$$

s being the distance between the centres of the circles. The distance of the image point in No. 1 from its centre is a_1b_1 and a_2b_2 is the distance of the image point in No. 2 from the centre of the second circle. The explicit formulas for b_1 and b_2 are

$$(121)\quad \begin{cases} b_1=\dfrac{s^2+a_1^2-a_2^2}{2sa_1}-\sqrt{\left(\dfrac{s^2+a_1^2-a_2^2}{2sa_1}\right)^2-1}, \\ b_2=\dfrac{s^2+a_2^2-a_1^2}{2sa_1}-\sqrt{\left(\dfrac{s^2+a_2^2-a_1^2}{2sa_2}\right)^2-1}, \end{cases}$$

which are both real, positive, and less than 1 (when the circles are external to each other), but b_1 and b_2 both approach the value 1 when the cylinders approach contact, and both approach the value zero when they are widely separated. In case $a_1=a_2$ then $b_1=b_2$. If the plus sign were taken in front of the radical, this would give the reciprocal of the value of b_1 given above, which is the other root of the quadratic equation to determine b_1, namely,

$$(122)\quad b_1^2-\left(\frac{s^2+a_1^2+a_2^2}{sa_1}\right)b_1+1=0.$$

The equation for b_2 is obtained by interchange of subscripts.

In abandoning the strict method of integral equations and assuming that the solution of the problem may be found in the form of a series (the Fourier-Bessel expansion) one cannot be sure of the existence of a solution of this form, or if a solution is found, there is no assurance that it is unique. Curiously enough, the infinite set of equations (113) which have been obtained for the coefficients $P_1,P_2,\ldots,Q_1,Q_2,\ldots$ admit of two solutions, and that one must be rejected as unphysical which does not satisfy the integral equation. This fact is brought out in the following derivation of asymptotic or high frequency formulae for the alternating current resistance and inductance.

In case the two cylinders have the same radius, conductivity, and permeability $b_1=b_2$, and $Q_n=P_n$. This is a problem in current distribution which has been treated by a number of mathematicians, the earliest being Mie*. It has been successfully solved for low frequencies by Curtis†, using a Fourier-Maclaurin expansion. The regular method of integral equations has also been used in a paper by Maneback‡, whose results are also limited in their application. The most thorough treatment, however, is that of Carson§, who has made arithmetical computations from the infinite set of linear equations, and whose results are quite unrestricted as to spacing of the wires or frequency. The problem having circular symmetry has also been solved in a general manner by Carson and Gilbert||.

The remainder of this paper will be devoted to the derivation of high frequency formulae for the alternating current resistance R, inductance L, and the attenuation b and phase velocity v, which hold for any dimensions or spacing of the two unequal cylinders.

(*a*) *First and second approximations at high frequency. Asymptotic formulas for R, L, and γ with any spacing.*
For high frequency

$$\frac{\mu_1 n J_n(x_1)+x_1 J_n'(x_1)}{\mu_1 n J_n(x_1)-x_1 J_n'(x_1)} \sim -1+\frac{2i\mu_1}{x_1}n+\text{higher powers of }\frac{1}{x}$$

$$\sim -\left\{1+(1-i)n\sqrt{\frac{\mu_1 R_1(0)}{2p}}\right\}$$

and

$$\frac{\mu_2 n J_n(x_2)+x_2 J_n'(x_2)}{\mu_2 n J_n(x_2)-x_2 J_n'(x_2)} \sim -\left\{1+(1-i)n\sqrt{\frac{\mu_2 R_2(0)}{2p}}\right\}.$$

The equations (113) become

$$(123)\quad \begin{cases} 1+\sum_{k=1}^{\infty}\frac{\overline{n+k-1}!}{\overline{n-1}!\,k!}P_k-\left(\frac{s}{a_2}\right)^{2n}Q_n\left\{1+n(1-i)\sqrt{\frac{\mu_2 R_2(0)}{2p}}\right\}=0, \\ 1+\sum_{k=1}^{\infty}\frac{\overline{n+k-1}!}{\overline{n-1}!\,k!}Q_k-\left(\frac{s}{a_1}\right)^{2n}P_n\left\{1+n(1-i)\sqrt{\frac{\mu_1 R_1(0)}{2p}}\right\}=0. \end{cases}$$

*G. Mie. Annalen der Physik II, 1900, pp. 201-249.

†H. L. Curtis. Bureau of Standards Scientific Paper No. 374.

‡Charles Maneback. Journal of Math. and Phys. Massachusetts Institute of Technology. Vol. I, No. 3, April, 1922; pp. 123-124.

§John R. Carson. Phil. Mag., vol. XLI, April, 1921.

||J. R. Carson and J. J. Gilbert. *Transmission Characteristics of the Submarine Cable.* Journal of the Franklin Institute. December 1921; pp. 705-735.

See also H. B. Dwight, Trans. A.I.E.E., 1923, p. 850.

Assume

$$(124)\qquad \begin{cases} P_n = P_n^{(0)} + \dfrac{1}{\sqrt{2p}} P_n^{(1)}, \\ Q_n = Q_n^{(0)} + \dfrac{1}{\sqrt{2p}} Q_n^{(1)}. \end{cases}$$

Substituting these forms in the above set of equations and equating like powers of $\frac{1}{\sqrt{2p}}$ gives

$$(125)\qquad \begin{cases} 1 + \displaystyle\sum_{k=1}^{\infty} \frac{n+k-1!}{n-1!\ k!} P_k^{(0)} = \left(\frac{s}{a_2}\right)^{2n} Q_n^{(0)}, \\ 1 + \displaystyle\sum_{k=1}^{\infty} \frac{n+k-1!}{n-1!\ k!} Q_k^{(0)} = \left(\frac{s}{a_1}\right)^{2n} P_n^{(0)}, \end{cases} \qquad (n=1,\ 2,\ 3, \ldots \infty),$$

and

$$(126)\qquad \begin{cases} \displaystyle\sum_{k=1}^{\infty} \frac{n+k-1!}{n-1!\ k!} P_k^{(1)} - \left(\frac{s}{a_2}\right)^{2n} Q_n^{(1)} = n\left(\frac{s}{a_2}\right)^{2n} Q_n^{(0)} (1-i)\sqrt{\mu_2 R_2(0)}, \\ \displaystyle\sum_{k=1}^{\infty} \frac{n+k-1!}{n-1!\ k!} Q_k^{(1)} - \left(\frac{s}{a_1}\right)^{2n} P_n^{(1)} = n\left(\frac{s}{a_1}\right)^{2n} P_n^{(0)} (1-i)\sqrt{\mu_1 R_1(0)}. \end{cases}$$

The equations (125) can be solved exactly, and the values of $P_n^{(0)}$ and $Q_n^{(0)}$ then substituted in (126) and the resulting equations solved exactly as follows:

Exact solution of the first system of equations (125).

Since

$$1 + \sum_{k=1}^{\infty} \frac{n+k-1!}{n-1!\ k!} z^k = \frac{1}{(1-z)^n} \text{ if } |z| < 1,$$

the equations (125) for $P_n^{(0)}$ and $Q_n^{(0)}$ suggest the forms

$$P_n^{(0)} = z_1^n, \qquad Q_n^{(0)} = z_2^n, \quad \text{where} \quad \begin{matrix} |z_1| < 1, \\ |z_2| < 1, \end{matrix}$$

and where z_1 and z_2 are to be determined. Substituting in the equations (125) gives

$$\begin{cases} 1 + \displaystyle\sum_{k=1}^{\infty} \frac{n+k-1!}{n-1!\ k!} z_1^k = \frac{1}{(1-z_1)^n} = \left(\frac{s}{a_2}\right)^{2n} z_2^n, \\ 1 + \displaystyle\sum_{k=1}^{\infty} \frac{n+k-1!}{n-1!\ k!} z_2^k = \frac{1}{(1-z_2)^n} = \left(\frac{s}{a_1}\right)^{2n} z_1^n, \end{cases} \qquad (n=1,\ 2,\ 3, \ldots\ldots \infty),$$

which shows that the 2∞ equations will all be satisfied if z_1 and z_2 can be found each numerically less than 1, satisfying the two equations

$$(127)\qquad \left(\frac{sz_1}{a_1}\right)\left[s-a_2\left(\frac{sz_2}{a_2}\right)\right]=a_1,\ \left(\frac{sz_2}{a_2}\right)\left[s-a_1\left(\frac{sz_1}{a_1}\right)\right]=a_2.$$

By reference to the equations (120) which determine b_1 and b_2 it is evident that

$$z_1=\frac{a_1b_1}{s}\ ,\ z_2=\frac{a_2b_2}{s}$$

where b_1 and b_2 are roots of the corresponding quadratic equations. In order that z_1 and z_2 shall be less than 1, b_1 and b_2 must both be less than 1, or both greater than 1. Since the product of the two roots of each equation is one, it follows that there are two solutions for the above system of equations (125) given by

$$(128)\qquad \begin{cases} P_n^{(0)}=\left(\dfrac{ab_1}{s}\right)^n, \\ Q_n^{(0)}=\left(\dfrac{ab_2}{s}\right)^n, \end{cases} \quad \text{or} \quad \begin{cases} P_n^{(0)}=\left(\dfrac{a}{sb_1}\right)^n, \\ Q_n^{(0)}=\left(\dfrac{a}{sb_2}\right)^n. \end{cases}$$

It will appear presently that the set corresponding to b_1 and b_2 both less than unity leads to positive expressions for the resistance and inductance, while the other set leads to the same numerical values but with a negative sign. Using these values of $P_n^{(0)}$ and $Q_n^{(0)}$ in (126) leads to the following equations for the $P_n^{(1)}$ and $Q_n^{(1)}$.

Exact solution of the second system of equations (126).

To find $P_n^{(1)}$ and $Q_n^{(1)}$ as solutions of

$$(129)\qquad \begin{cases} \displaystyle\sum_{k=1}^{\infty}\frac{n+k-1!}{n!\ k!}P_k^{(1)}-\left(\frac{s}{a_2}\right)^{2n}\frac{Q_n^{(1)}}{n}=\left(\frac{sb_2}{a_n}\right)^n(1-i)\sqrt{\mu_2R_2(0)}, \\ \displaystyle\sum_{k=1}^{\infty}\frac{n+k-1!}{n!\ k!}Q_n^{(1)}-\left(\frac{s}{a_1}\right)^{2n}\frac{P_n^{(1)}}{n}=\left(\frac{sb_1}{a_1}\right)^n(1-i)\sqrt{\mu_1R_1(0)}. \end{cases}$$

Assume

$$P_n^{(1)}=\alpha nz_1^n \text{ and } Q_n^{(1)}=\beta nz_2^n \text{ where } \begin{matrix}|z_1|<1,\\ |z_2|<1,\end{matrix}$$

and substitute in above equations (129). This gives

$$\begin{cases} \displaystyle\alpha z_1\sum_{k=1}^{\infty}\frac{n+k-1!}{n!\ k-1!}z_1^{k-1}-\beta\left(\frac{s}{a_2}\right)^{2n}z_2^n=\left(\frac{sb_2}{a_2}\right)^n(1-i)\sqrt{\mu_2R_2(0)}, \\ \displaystyle\beta z_2\sum_{k=1}^{\infty}\frac{n+k-1!}{n!\ k-1!}z_1^{k-1}-\alpha\left(\frac{s}{a_1}\right)^{2n}z_1^n=\left(\frac{sb_1}{a_1}\right)^n(1-i)\sqrt{\mu_1R_1(0)}. \end{cases}$$

Or, since by the binominal theorem

$$\sum_{k=1}^{\infty}\frac{n+k-1!}{n!\ k-1!}z_1^{k-1}=\frac{1}{(1-z_1)^{n+1}},$$

$$\frac{\alpha z_1}{1-z_1}\frac{1}{(1-z_1)^n}-\beta\left(\frac{s^2z_2}{a_2^2}\right)^n=\left(\frac{sb_2}{a_2}\right)^n(1-i)\sqrt{\mu_2R_2(0)},$$

$$\frac{\beta z_2}{1-z_2}\frac{1}{(1-z_2)^n}-\alpha\left(\frac{s^2z_1}{a_1^2}\right)^n=\left(\frac{sb_1}{a_1}\right)^n(1-i)\sqrt{\mu_1R_1(0)}.$$

If these equations are to hold for every positive integral value of n, this requires that α and β have certain values, and that

$$z_2=\frac{a_2b_2}{s}=\frac{\left(\frac{a_2}{s}\right)^2}{1-z_1},\quad z_1=\frac{a_1b_1}{s}=\frac{\left(\frac{a_1}{s}\right)^2}{1-z_2}, \tag{130}$$

where z_1, z_2, b_1, and b_2 must be the same as in the preceding case.

Since

$$\frac{z_1}{1-z_1}=\frac{a_1b_1b_2}{a_2},\qquad \frac{z_2}{1-z_2}=\frac{a_2b_1b_2}{a_1}$$

the constants α and β must be so chosen as to satisfy

$$\begin{cases}-a_1\alpha+a_2b_1b_2\beta=(1-i)a_1\sqrt{\mu_1R_1(0)},\\ a_1b_1b_2\alpha-a_2\beta=(1-i)a_2\sqrt{\mu_2R_2(0)},\end{cases} \tag{131}$$

which being solved give

$$\begin{cases}a_1\alpha=-\dfrac{(1-i)}{(1-b_1^2b_2^2)}[a_1\sqrt{\mu_1R_1(0)}+b_1b_2a_2\sqrt{\mu_2R_2(0)}],\\ a_2\beta=-\dfrac{(1-i)}{(1-b_1^2b_2^2)}[b_1b_2a_1\sqrt{\mu_1R_1(0)}+a_2\sqrt{\mu_2R_2(0)}.\end{cases} \tag{132}$$

Introducing the results of the first and second stages of the approximation, namely,

$$\begin{cases}P_n=P_n^{(0)}+\dfrac{1}{\sqrt{2p}}P_n^{(1)}=\left(\dfrac{a_1b_1}{s}\right)^n\left(1+\dfrac{n\alpha}{\sqrt{2p}}\right),\\ Q_n=Q_n^{(0)}+\dfrac{1}{\sqrt{2p}}P_n^{(1)}=\left(\dfrac{a_2b_2}{s}\right)^n\left(1+\dfrac{n\beta}{\sqrt{2p}}\right)\end{cases} \tag{133}$$

into the equation (118) for R and L leads to two infinite series which are summable. For

$$\log\frac{s}{a_1}-\sum_{n=1}^{\infty}\frac{P_n}{n}=\log\frac{s}{a_1}-\sum_{n=1}^{\infty}\frac{1}{n}\left(\frac{a_1b_1}{s}\right)-\frac{\alpha}{\sqrt{2p}}\sum_{n=1}^{\infty}\left(\frac{a_1b_1}{s}\right)^n$$

$$=\log\frac{s}{a_1}+\log\left(1-\frac{a_1b_1}{s}\right)-\frac{\alpha}{\sqrt{2p}}\left(\frac{1}{1-\frac{a_1b_1}{s}}-1\right)$$

$$=\log\frac{a_2}{a_1b_2}-\frac{\alpha}{\sqrt{2p}}\frac{a_1b_1b_2}{a_2}.$$

Similarly

$$\log\frac{s}{a_2}-\sum_{n=1}^{\infty}\frac{Q_n}{n}=\log\frac{a_1}{a_2b_1}-\frac{\beta}{\sqrt{2p}}\frac{a_2b_1b_2}{a_1}$$

and

$$\log\frac{s^2}{a_1a_2}-\sum_{n=1}^{\infty}\frac{P_n+Q_n}{n}=\log\frac{1}{b_1b_2}-\frac{b_1b_2}{\sqrt{2p}}\left(\frac{\alpha a_1}{a_2}+\frac{\beta a_2}{a_1}\right).$$

Substituting in (118) gives

$$R+ipL=2ip\left\{\frac{\mu_1J_0(x_1)}{x_1J_0'(x_1)}+\frac{\mu_2J_0(x_2)}{x_2J_0'(x_2)}+\log\frac{1}{b_1b_2}\right.$$
$$\left.+\frac{(1-i)b_1b_2}{(1-b_1^2b_2^2)\sqrt{2p}}\left[\frac{a_1}{a_2}\sqrt{\mu_1R_1(0)}+\frac{a_2}{a_1}\sqrt{\mu_2R_2(0)}+b_1b_2(\sqrt{\mu_1R_1(0)}+\sqrt{\mu_2R_2(0)})\right]\right\}. \tag{134}$$

But to the same degree of approximation as used in obtaining P_n and Q_n, that is neglecting $\frac{1}{p}$ compared to $\frac{1}{\sqrt{p}}$,

$$\frac{\mu_1J_0(x_1)}{x_1J_0'(x_1)}=-\frac{i\mu_1}{x_1}=\frac{(1-i)}{2}\sqrt{\frac{\mu_1R_1(0)}{2p}},$$

and

$$\frac{\mu_2J_0(x_2)}{x_2J'(x_2)}=-\frac{(1-i)}{2}\sqrt{\frac{\mu_2R_2(0)}{2p}}.$$

Using these expressions in the preceding equation (134) gives, upon equating reals and imaginaries,

$$R=\frac{1}{2(1-b_1^2b_2^2)}\left[\left(1+2b_1b_2\frac{a_1}{a_2}+b_1^2b_2^2\right)\sqrt{2p\mu_1R_1(0)}\right.$$
$$\left.+\left(1+2b_1b_2\frac{a_2}{a_1}+b_1^2b_2^2\right)\sqrt{2p\mu_2R_2(0)}\right]$$
$$=\frac{1+b_1b_2}{1-b_1b_2}\left(\frac{1}{a_1}+\frac{1}{a_2}\right)\sqrt{\frac{\mu f}{\lambda}},\quad (\text{if } \mu_1=\mu_2=\mu,\ \lambda_1=\lambda_2=\lambda), \tag{135}$$

$$L=2\log\frac{1}{b_1b_2}+\frac{R}{p}=\frac{c^2C}{k}+\frac{R}{p}. \tag{136}$$

The attenuation b and phase velocity V are given by

$$\gamma \equiv b + ipV = \sqrt{(R+ipL)(G+ipC)},$$

or

$$(137) \qquad \begin{cases} 2b^2 = \sqrt{(R^2+p^2L^2)(G^2+p^2C^2)} + (RG - p^2LC), \\ 2\,\dfrac{p^2}{V^2} = \sqrt{(R^2+p^2L^2)(G^2+p^2C^2)} - (RG - p^2LC). \end{cases}$$

The ratio $\dfrac{G}{pC}$, *i.e.* $\dfrac{4\pi\lambda dc^2}{pk}$, is generally negligible, so that (137) reduces to

$$(138) \qquad V = \frac{1}{\sqrt{LC}}, \quad b = \frac{RCV}{2}.$$

The positive quantity b_1b_2 which appears in these formulae may be found as the root of the quadratic equation

$$(139) \qquad (b_1b_2)^2 - \left(\frac{s^2 - a_1^2 - a_2^2}{a_1a_2}\right) b_1b_2 + 1 = 0$$

which is less than 1. The other root, being the reciprocal of this will be greater than 1, and if used in these formulae would give the same numerical values of R and L but both negative in sign.

Special Case.

The circles have equal radii, $b_2 = b_1$,

$$(140) \qquad R = \left\{\frac{\sqrt{\mu_1 R_1(0)} + \sqrt{\mu_2 R_2(0)}}{2}\right\} \sqrt{\frac{2p}{1 - \left(\dfrac{2a}{s}\right)^2}},$$

$$(141) \qquad L = 4 \log \frac{1}{b_1} + \frac{R}{p},$$

where b_1 is that root of

$$(142) \qquad b_1^2 - \left(\frac{s}{a_1}\right) b_1 + 1 = 0$$

which is less than 1, or

$$b_1 = \frac{s}{2a_1} - \sqrt{\left(\frac{s}{2a_1}\right)^2 - 1}.$$

The asymptotic expressions for alternating current resistance and inductance of two cylinders here given are believed to be new. They have proven useful in the Radio Section of this Bureau for investigating the precision of measurements of short waves on wires, corresponding to a frequency of $(10)^7$ cycles per second*.

*A. Hund. Scientific Paper No. 491.

(*b*) *Special Case. Circular Cylinder Parallel to Semi-infinite Conductor of Finite Conductivity.*

If we place $s=d+a_2$ and holding d fixed let a_2 become infinite, this approaches the case of a cylinder of radius a_1 whose centre is at a distance d from the plane boundary of a semi-infinite medium of finite conductivity. This is the case of a single cylinder a height d above the earth, the latter being the return conductor. In this case

$$b_1 b_2 = \frac{d-\sqrt{d^2-a_1^2}}{a_1}\,.$$

Formulae (135) to (138) then simplify, and they show the manner in which the conductivity of the ground λ_2 affects the propagation along a horizontal antenna at radio frequencies. In case $\mu_1=\mu_2=1$, formula (135) gives for the resistance per unit length of the circuit made of a horizontal antenna and the ground, where d is large compared to a_1,

$$R=\left(\frac{1}{a_1\sqrt{\lambda_1}}+\frac{1}{d\sqrt{\lambda_2}}\right)\sqrt{f}\,, \tag{143}$$

where f is the frequency, a_1 and λ_1 the radius and conductivity of the antenna, and d and λ_2 its height above ground and the conductivity of the latter respectively.

As a numerical example of the use of (135) to (139) consider a two-conductor cable of copper wires of radii $a_1=.3$ cm., $a_2=.5$ cm., with a distance $s=1$ cm. between centres and surrounded by a large amount of insulating material whose dielectric constant is $k=4$ electrostatic c.g.s. units. Suppose that the electrical conductivity of this material λ_d is less than about 10^{-20} electromagnetic c.g.s. units, and the frequency f is 10^5 cycles per second. For this value of the frequency (or higher values) the insulation leakage G has no appreciable effect, for

$$\frac{G}{pC}=\frac{4\pi c^2\lambda_d}{2\pi fk}=\frac{4\pi 2^3\times 10^{20}\times 10^{-20}}{2\pi\times 4\times 10^5}=4\times 10^{-5}$$

which is negligible compared to unity and this indicates the error involved by neglecting G.

The direct current resistance of the wires per unit length in electromagnetic c.g.s. units is

$$R_1(0)=\frac{1}{\pi a_1^2\lambda}=\frac{1}{\pi(.3)^2.0006}=5.9\times(10)^3,$$

$$R_2(0)=\frac{1}{\pi a_2^2\lambda}=\frac{1}{\pi(.5)^2.0006}=2.1\times(10)^3.$$

The direct current resistance of the line per unit length is

$$R_1(0)+R_2(0)=8\times 10^3,$$

$$\sqrt{2pR_1(0)} = \frac{2}{a_1}\sqrt{\frac{f}{\lambda}} = \frac{2}{a_1}\times 1.3\times 10^4,$$

$$\sqrt{2pR_2(0)} = \frac{2}{a_2}\sqrt{\frac{f}{\lambda}} = \frac{2}{a_2}\times 1.3\times 10^4.$$

The equation (135) for the alternating current resistance of the line per unit length is

$$R=\frac{1+b_1b_2}{1-b_1b_2}\left(\frac{1}{.3}+\frac{1}{.5}\right)1.3\times 10^4=6.93\frac{1+b_1b_2}{1-b_1b_2}\times 10^4,$$

(electromagnetic c.g.s. units).

The numerical value of b_1b_2 is found as that root of the quadratic equation (139)—

$$(b_1b_2)^2-\frac{1-.3^2-.5^2}{.3\times .5}(b_1b_2)+1=0,$$

which is less than 1. The two roots are $b_1b_2=.24$ or 4.16. The former value must be used, and it gives

$$R=6.93\left(\frac{1.24}{.76}\right)\times 10^4=113\times 10^3$$

as the alternating current resistance of the line per unit length in electromagnetic c.g.s. units. This is 14 times the direct current resistance. After R has been computed, the alternating current inductance L of the line per unit length may be computed by (136):

$$L=2\log_e\frac{1}{.24}+\frac{113\times 10^3}{2\pi 10^5}=2.85+.18=3.03$$

(electromagnetic c.g.s. units per cm.).

The capacity of the line per unit length is given by (119):

$$C=\frac{4}{3^2\times 10^{20}\times 2.85}=1.56\times 10^{-21},$$

(electromagnetic c.g.s. units).

The phase velocity V is next found by means of (138):

$$V=\frac{1}{\sqrt{LC}}=\frac{1}{\sqrt{3.03\times 1.56\times 10^{-21}}}=1.45\times 10^{10}$$

which is about half the velocity of light in space. However, the velocity of light in the medium which is assumed to have a dielectric capacity $k=4$ is

$$V_0=\frac{c}{\sqrt{k}}=\frac{3}{\sqrt{4}}+10^{10}=1.50\times 10^{10},$$

which shows that V is very nearly V_0. In this problem the dielectric whose specific inductive capacity is 4 electrostatic c.g.s. units is assumed to fill the space external to the wires.

Finally the attenuation constant b is found by (137):

$$b=\frac{RCV}{2}=\frac{113\times10^3\times1.56\times10^{-21}\times1.45\times10^{10}}{2}=1.28\times10^{-6}.$$

The current I is of the form

$$I=I_0e^{-bz}\cos 2\pi f\left(t-\frac{z}{V}\right),$$

and the wave-length is

$$\frac{V}{f}=\frac{1.45\times10^{10}}{10^5}=1450 \text{ meters}.$$

VIII. Summary

In a single type of wave, the field components are proportional to $e^{ipt-\gamma z}$ where the propagation constant γ is $b+\frac{ip}{v}$, the real constants b and v being the attenuation and phase velocity respectively. They are functions of the frequency. There are $N-1$ possible values of γ^2 and $2N-2$ values of γ or types of waves for a system of N cylindrical conductors. The electromagnetic field is in general the super-position of the fields corresponding to each type and these are derivable from a complex scalar potential ϕ, and a vector potential A. On account of the vast difference in the order of magnitude of the electrical conductivities of a conductor and of a dielectric, certain approximations can be made in general which are valid from the lowest to the highest or radio frequencies. These lead to the conclusion that the x- and y-components of the vector potential are negligible.

Beginning with the differential equations and boundary conditions which ϕ and A must satisfy, the existence of a solution ϕ has been proven, and an integral representation of it obtained and its properties studied by constructing a symmetrical auxiliary function of two points $G(xy\xi\eta)$ which may be regarded as a generalization of Green's function. By its means, the unknown charges on the boundaries between different dielectrics are eliminated from consideration and ϕ is determined at all points in terms of its values on the conducting sections. Certain constants G_{nk} and C_{nk} are derived which in general are complex functions of the frequency, and are coefficients of leakage and capacity respectively. The conservation of electricity for each conductor takes the form, for $n=1, 2, 3, \ldots, N$,

$$\gamma I_n=\sum_{k=1}^{N}(G_{nk}+ipC_{nk})c_k \text{ where } \begin{cases}G_{nk}=G_{kn},\\ C_{nk}=C_{kn},\end{cases}$$

and

$$\sum_{k=1}^{N}G_{nk}=\sum_{k=1}^{N}C_{nk}=0.$$

In a similar manner an integral equation has been obtained for the vector potential A by devising a symmetrical function of two points $M(xy, \xi\eta)$ which may be called a magnetic flux function. The existence and uniqueness of a solution of this equation has been proven and methods of solving it discussed. The form and properties of the solution are studied. Its value at all points in conductors may be found without reference to the dielectric. The function M thus eliminates from consideration the dielectric as well as the surface distribution of magnetism existing at the boundaries between different magnetic materials. The real coefficients of resistance and inductance R_{nk} and L_{nk} are derived which have the symmetrical property $R_{nk}=R_{kn}$, $L_{nk}=L_{kn}$, and in terms of these the definition of current leads to the form

$$\gamma c_n - ipA(\infty) = \sum_{k=1}^{N} (R_{nk}+ipL_{nk})I_k, \quad (n=1, 2, 3, \ldots, N).$$

These equations are rendered compatible with those for the conservation of electricity by choosing γ a root of a certain determinant which leads to an algebraic equation of degree $N-1$ in γ^2. It is shown that there is but one arbitrary constant for each type of propagation and all these may be determined when the terminal apparatus at both ends of the line are given. Thus the attenuation, phase velocity, and reactance of the line may be found. Formulae are also developed for the heating in conductors, and in dielectric, and for the mean electrical and electrokinetic energy of the system in terms of the coefficients above mentioned. Application is made to a pair of circular cylindrical conductors of unequal radii, conductivity, and permeability, and high frequency formulae derived for the resistance R and inductance L of the line per unit length, as well as for the attenuation b and phase velocity V. These are believed to be new.

CONTENTS

INDEX OF PRINCIPAL SYMBOLS

$E(E_x, E_y, E_z)$ = electric vector.
$H(H_x, H_y, H_z)$ = magnetic vector.
$B(B_x, B_y, B_z)$ = magnetic induction.
f = frequency.
p = $2\pi f$.
γ = propagation constant $= b + \frac{ip}{V}$, $i = \sqrt{-1}$.
b = attenuation constant.
V = phase velocity.
c = $3(10)^{10}$ = the ratio of the two c.g.s. electrical units.
μ = magnetic permeability.
k = dielectric constant.
λ = electrical conductivity.

All constants and vectors measures in c.g.s. electromagnetic units except k which is measured in electrostatic c.g.s. units so that

k = 1 to 5 for most dielectrics.

μ = 1 for non-magnetic metals.

ϕ = complex scalar potential.

$A(A_x A_y A_z)$ = complex vector potential.

NOTE—Since A_z is the principal component considered A_z is written as A where this can be done without misunderstanding. Similarly, E is used for E_z where the meaning is plain.

$c_1, c_2, \ldots, c_r$ = complex constant values of ϕ upon conducting sections Nos. 1, 2, 3, ..., N.

$a_j \equiv \lambda_j + \dfrac{ipk_j}{4\pi c^2}$ = complex conductivity of jth material.

ds_j = element of arc of natural boundary of section S_j of homogeneous material (conductor or dielectric).

n_j = normal to this boundary curve, see Fig. 1.

Q_n = complete free charge per unit length upon nth conducting group.

G_n = leakage current (complex) from nth conducting group per unit length.

G_{nk} and C_{nk} = real coefficients of leakage and capacity respectively, defined by (26).

$\mathbf{I}_n$ = z-component of total conduction current through the section S_n of the nth conducting group.

$G(xy,\xi\eta)$ = a generalized Green's function, $M(xy,\xi\eta)$ = a magnetic flux function — defined where introduced.

R_{nk} and L_{nk} = coefficients of resistance and inductance; defined by (81).

R and L = the alternating current resistance and inductance per unit length respectively for a simple return circuit of two conductors. (Equations 135 and 136).

PERTURBATIONS IN THE ORBITS OF ELECTRONS

BY PROFESSOR G. GIANFRANCESCHI,
President of the Pontifical Academy of Science "The Nuovi Lincei", Rome, Italy.

The motion of an electron around the nucleus is similar to the motion of a planet around the sun. We will assume the force that maintains the electron in its orbit to be the electric force, and to act in accord with Coulomb's Law. The equations of a moving electron are therefore

$$(1) \qquad \frac{d^2x}{dt^2} + \frac{ux}{r^3} = 0, \qquad \frac{d^2y}{dt^2} + \frac{uy}{r^3} = 0,$$

u being the constant of attraction.

The emission of light-vibrations is certainly dependent on the orbital motion of electrons, whatever may be the theory of that emission. The Zeeman and Stark effects show that a magnetic or an electric field influences that motion, so that the lines of emission are split into a certain number of components. In order to explain the Zeeman effect, Lorentz has shown that the magnetic field introduces a central force in addition to the electric force which he supposes is retaining the electron in its orbit.

It is possible to solve the problem also on the hypothesis of Coulomb's force, assuming the magnetic or electric force as a perturbing force. We can suppose this additional force to be in the plane of the orbit, that is to say, we can consider only its component in the plane. Also we can choose the x-axis in the same direction as the component X of the perturbing force. The equations of motion then become

$$(2) \qquad \frac{d^2x}{dt^2} + \frac{ux}{r^3} = X, \qquad \frac{d^2y}{dt^2} + \frac{uy}{r^3} = 0,$$

or in the polar system,

$$(3) \qquad \frac{d^2r}{dt^2} - r\left(\frac{d\theta}{dt}\right)^2 = -\frac{\mu}{r^2} + X\cos\theta, \qquad \frac{1}{r}\frac{d}{dt}\left(r^2\frac{d\theta}{dt}\right) = -X\sin\theta.$$

Professor Armellini of Rome has published a paper on the perturbations of planetary orbits by the attraction of the galactic. His method consists in determining successively the variations of the characteristic quantities subjected to the perturbing force. The method can be applied to the electronic orbits.

In order to solve only the proposed problem we may consider the variation of the vector r in the simplest case of circular orbits. Let r_0 be the radius of the primitive unperturbated orbit, and ρ the variation due to the perturbing force X. Let us introduce the expression $r_0+\rho$ into the equations and compare the resulting equations with the original ones. Then it is easy to deduce a differential equation in ρ, and from that to obtain the value of ρ. This differential equation in ρ is

$$\frac{d^2\rho}{dt^2}+n^2\rho=3X\cos nt, \tag{4}$$

n being the frequency which corresponds to the average periodic time. We can integrate the equation for any hypothesis.

If we consider the frequency n to be the same in the first as in the second member of the èquation we obtain

$$\rho=\frac{3Xt\sin nt}{2n} \tag{5^{(a)}}$$

and if n in the first member is regarded as a primitive constant frequency n_0 and in the second as a variable n the result is

$$\rho=\frac{3X\cos nt}{n^2-n_0{}^2} \tag{5^{(b)}}$$

Assuming ρ to be an infinitesimal quantity of the first order, and knowing the expression for n, we may write

$$n=n_0\left(1-\frac{3\rho}{2r_0}\right). \tag{6}$$

By means of (5) and (6) we can determine the variation of the frequency n caused by the perturbing force X, and from that variation can also be obtained λ the wave-length, and the consequent distribution of lines.

If we assume the X force to be the component of an electric field of force we can deduce the Stark effect.

If, on the contrary, X is the magnetic force, we have the Zeeman effect. In this case we can very easily proceed to the resulting expression for n. The magnetic force of a field H is always directed along the vector r, and its intensity is $\frac{e}{mc}(vH)$, a well known expression. Then the components in equations (3) are

$$X\cos\theta=\pm\frac{e}{mc}(vH),\quad X\sin\theta=0. \tag{7}$$

The sign $+$ or $-$ depends on the sense of the rotating electron. The value of ρ as determined from (4) is

$$\rho^1=\pm\frac{evH}{mcn^2}, \tag{8}$$

and consequently the frequency n as function of the primitive n_0 is

$$n = n_0 \pm \frac{3}{2} \frac{eH}{mc}. \tag{9}$$

If on the contrary we assume the central force to be an electric force, we can substitute for μ/r^2 in equations (3) a force proportional to the distance, that is fr. It is then very easy to obtain for ρ the expression:

$$\rho = \pm \frac{evH}{4fmc}, \tag{10}$$

and therefore

$$n = n_0 \pm \frac{1}{2} \frac{eH}{mc}. \tag{11}$$

as in the Lorentz formula.

ON THE QUANTIZATION OF CERTAIN ORBITS

By Professor Arthur W. Conway,
University College, Dublin, Irish Free State.

When an electron moves in a field of force we have in many cases three integrals, the energy integral which gives a constant value, C, for the sum of the kinetic and potential energies, and two angular momentum integrals. By the aid of these integrals and by a suitable choice of co-ordinates (not necessarily unique) it is sometimes possible to express the Action Integral $\sum\int pdq$, where q is a co-ordinate and p the corresponding momentum, in the form $f_1(q_1)+f_2(q_2)+f_3(q_3)$. The variables are then said to be "separable". These functions will, in general, be periodic functions of the variables and the Sommerfeld-Bohr quantization can be summed up in the statement that the periods of these functions are all integral multiples n_1, n_2, n_3, of the quantum constant h. We can thus express C in terms of three integers and the quantity C/h is called a term of the line-spectrum. Various empirical formulae for the terms are given, the usual being of the type $R/(n+\alpha)^2$ where R is a constant (the Rydberg constant), n an integer and α a function of n. By taking the law of force to be that of the inverse square, e^2/r^2, we get the hydrogen type R/r^2 and by taking more complicated central forces $e^2/r^2(1+c_1/r+c_2/r^2+\ldots)$. Sommerfeld* has obtained expressions which correspond to the term formulae of Rydberg and Ritz. In this paper we consider a type of force which is no longer central and we get expressions which are analogous to a very recent type of formula, that of Hicks.

Assume an atomic system which consists of an electron of mass m and charge $-e$ and a nucleus of mass m' and charge e, the co-ordinates being (x, y, z) and (x', y', z') respectively. We shall consider two cases:

I. The Lagrangian Function is

$$\tfrac{1}{2}m(\dot{x}^2+\dot{y}^2+\dot{z}^2)+\tfrac{1}{2}m'(\dot{x}'^2+\dot{y}'^2+\dot{z}'^2)-U(\dot{x}\dot{x}'+\dot{y}\dot{y}'+\dot{z}\dot{z}')+e^2/r$$

where $U=ke^2/r$, k being a constant, and $r^2=(x-x')^2+(y-y')^2+(z-z')^2$.

II. The Lagrangian Function is

$$\tfrac{1}{2}m(\dot{x}^2+\dot{y}^2+\dot{z}^2)+\tfrac{1}{2}I\omega^2+\lambda\omega\sin^2\theta\dot{\phi}/r+e^2/r$$

where the co-ordinates of m are (x, y, z) or (r, θ, ϕ) referred to a fixed point which we can consider the centroid of the nucleus and ω is the angular velocity of the nucleus about an axis, which is fixed in space, and I is an inertia coefficient.

**Atombau und Spectralinien*, pp. 721 *et seq.*

In the first case we get at once the three momentum integrals of the type $(m-U)\dot{x}+(m'-U)\dot{x}'=\text{const.}$, etc. If the system has no linear momentum the right hand sides of these equations are zero and putting $x-x'=\xi$; etc., the equations of motion, after some reductions, take the form

$$\frac{d}{dt}(M\dot{\xi})=\frac{\delta w}{\delta\xi}+\frac{\mu\mu' v^2}{(\mu+\mu')^2}\frac{\delta U}{\delta\xi};\ \text{etc.}$$

where

$$w=e^2/r,\ \mu=m-U,\ \mu'=m'-U,\ v^2=(\dot{x}-\dot{x}')^2+(\dot{y}-\dot{y}')^2+(\dot{z}-\dot{z}')^2$$

and

$$M=(mm'-U^2)/(m+m'-2U).$$

We get at once the angular momentum integrals $M(\xi\dot{\eta}-\dot{\eta}\zeta)=\text{const.}$, etc.

In a similar way we get the activity equation

$$\Sigma\,\dot{\xi}\frac{d}{dt}(M\dot{\xi})=\frac{dw}{dt}+\frac{\mu\mu'}{\mu+\mu' v^2}\frac{dU}{dt}.$$

On noticing that $\dfrac{dM}{dt}=2\dfrac{\mu\mu'}{(\mu+\mu')^2}\dfrac{dU}{dt}$ we get the energy equation $\frac{1}{2}Mv^2=w-C$. It may be noticed that the form of this is independent of U which, however, occurs in M. Taking now the axes so that $\zeta=0$, which is always possible on account of the angular momentum integrals, we have then using polar co-ordinates,

$$M^2\dot{r}^2=2e^2M/r-2CM-A^2/r^2,\quad Mr^2\dot{\theta}=A.$$

So far we have assumed nothing as to the size of the constant k occurring in U. The shortest way of describing its magnitude is to state that M can be expanded in positive powers of k.

To a first approximation we therefore have

$$M^2\dot{r}^2=B_1/r-C_1-A_1/r^2$$

where

$$A_1=A^2-\frac{4e^2kM_0}{M_1},\ B_1=2e^2M_0\left(1-\frac{2k}{M_1}\right),\ C_1=2CM_0,$$

writing for shortness

$$M_0=mm'/(m+m'),\ M_1=m+m'.$$

We get as usual the phase integral

$$2\int M\dot{r}\,dr=\pi B_1/\sqrt{C_1}-2\pi\sqrt{A_1}.$$

If we put nh for the value of this phase constant and if we put A equal to $n'h$ we get, putting $R=2\pi e^2M_0/h^3$,

$$\sqrt{\frac{Rh}{C}}=n+n'-\frac{2hRK}{n'M_1}+\frac{K}{M_1}\sqrt{Rhc},$$

or

$$C=\frac{Rh}{\left\{n+n'-\frac{4kh^2R}{n'M_1}+\frac{kh^2R}{M_1(n+n')}\right\}^2},$$

a formula which is of the type of Hicks' formula.

This formula involves only the total quantum number $n+n'$ and there is no distinction between the latitude and longitude quantum numbers n'' and n''' only their sum n' being involved. In fact we have selected axes such that $n''=0$. By taking the terms (i) and (iii), however, the quantum numbers get separated and a different set of terms arise all of which have the same total azimuthal quantum number. It is somewhat analogous to a dynamical system which has a number of its periods equal and these periods become separated when a slight change is made in the system.

The Lagrangian Function now becomes

$$\tfrac{1}{2}m\left\{\dot{r}^2+r^2(\dot{\theta}^2+\dot{\phi}^2\sin^2\theta)\right\}+\tfrac{1}{2}I\omega^2+\frac{\lambda\omega\dot{\phi}\sin^2\theta}{r}+\frac{e^2}{r}.$$

We easily get the following three integrals (neglecting λ^2)

$$mr^2\sin^2\theta\dot{\phi}+\frac{\lambda\omega\sin^2\theta}{r}=\alpha_1,$$

$$m^2r^4\dot{\theta}^2=\alpha_2^2-\alpha_1^2/\sin^2\theta,$$

$$I\omega+\frac{\lambda\dot{\phi}\sin^2\theta}{r}=\alpha_3,$$

and the energy integral

$$\tfrac{1}{2}m\dot{r}^2+\frac{\alpha_2}{2mr^2}+\frac{\alpha_3}{2I}-\frac{\lambda\alpha_1\alpha_3}{Imr^3}-\frac{e^2}{r}=-E$$

where E, the Energy, is constant.

On quantizing these equations we get:

$$2\pi\alpha_1=n_1h+\frac{\lambda\alpha_1\alpha_3}{mI}\int\frac{dt}{r^3},$$

$$2\pi\alpha_2=2\pi\alpha_1+n_2h,$$

$$2\pi\alpha_3=n_3h+\lambda\int\frac{\dot{\phi}\sin^2\theta}{r}d\chi,\text{ where }\dot{\chi}=\omega,$$

$$2\pi^2me^4/h^2(E+\alpha_3/2I)=n_2+n_4$$

Thus, in the usual notation,

$$-E=\frac{\alpha_3}{2I}-\frac{Rh}{(n_2+n_4)^2}.$$

The orbits are therefore to the first power of λ degenerate.

THE GRAVITATIONAL FIELD IN A CURVED SPACE OF AN ELECTRICAL SPHERE IN WHICH THE DENSITY OF MATTER IS VARIABLE

By Dr. H. L. Vanderlinden,
Royal Observatory of Belgium, Uccle, Belgium.

I. Introduction

It is proposed in this paper to determine in a curved space the gravitational field of a sphere composed of a fluid of variable density; furthermore we suppose that the surface of the sphere is covered with an infinitesimally thin layer of electricity. The solution is obtained exactly in—and outside the sphere. The present considerations lead to the generalization of the existing solutions in the case of the sphere.

1. *The field with spherical symmetry.* We suppose that the field in the curved space is produced by a sphere with radius $r=r_a$, the surface of which is covered with an electrostatic charge. The centre of the sphere is chosen as origin of the coordinates. The field will have the properties of spherical symmetry, and supposing that it is stationary, it will in polar coordinates r, θ, ϕ be characterized by the quadratic form:

$$\delta s^2 \equiv A\,\delta r^2 + B(\delta\theta^2 + \sin^2\theta\,\delta\phi^2) + C\delta t^2, \tag{1}$$

where A, B, C are only functions of r.

Adopting Schwarzschild's notation

$$x_1 = r^3/3,\ x_2 = -\cos\theta,\ x_3 = \phi,\ x_4 = t, \tag{2}$$

the quadratic form (1) can be written:

$$\delta s^2 = -f_1\delta x_1^2 - f_2\delta x_2^2/(1-x_2^2) - f_2(1-x_2^2)\,\delta x_3^2 + f_4\delta t^2. \tag{3}$$

The coordinate r may always be chosen in such a way that the determinant of the potentials is constant, *e.g.*, the square of the velocity c of light *in vacuo*:

$$-g = f_1 f_2^2 f_4 = c^2. \tag{4}$$

The problem now consists in determining f_1, f_2, f_4 by means of Einstein's field equations.

2. *The field equations.* The ten general field equations are

$$G_{\alpha\beta}-\lambda g_{\alpha\beta}=-\kappa(T_{\alpha\beta}-\tfrac{1}{2}g_{\alpha\beta}T), \qquad (\alpha, \beta=1, 2, 3, 4), \tag{5}$$

where λ is the cosmological constant and κ a universal constant closely connected with Gauss' gravitational constant.

Because of the constancy of g, the symmetrical quantities $G_{\alpha\beta}$ are defined as follows:

$$G_{\alpha\beta}=\sum_{\sigma}\left[-\frac{\partial\{\alpha\beta,\sigma\}}{\partial x_\sigma}+\sum_{\tau}\{\beta\tau,\sigma\}\{\alpha\sigma,\tau\}\right]=G_{\beta\alpha}. \tag{6}$$

The Christoffel symbols are

$$\{\alpha\beta,\sigma\}=\tfrac{1}{2}\sum_{\tau}g^{\sigma\tau}\left(\frac{\partial g_{\sigma\tau}}{\partial x_\beta}+\frac{\partial g_{\beta\tau}}{\partial x_\alpha}-\frac{\partial g_{\alpha\beta}}{\partial x_\tau}\right), \tag{7}$$

$g^{\sigma\tau}$ being the complement of $g_{\sigma\tau}$ in the determinant g. In this particular case the symbols (7) which are not identically zero, have the following values:

$$\{11,1\}=\tfrac{1}{2}\frac{1}{f_1}\frac{\partial f_1}{\partial x_1}, \quad \{22,1\}=-\tfrac{1}{2}\frac{1}{f_1}\frac{\partial f_2}{\partial x_1}\cdot\frac{1}{1-x_2^2}, \quad \{33,1\}=-\tfrac{1}{2}\frac{1}{f_1}\frac{\partial f_2}{\partial x_1}(1-x_2^2),$$

$$\{44,1\}=\tfrac{1}{2}\frac{1}{f_1}\frac{\partial f_4}{\partial x_1}, \quad \{21,2\}=\tfrac{1}{2}\frac{1}{f_2}\frac{\partial f_2}{\partial x_1}, \quad \{22,2\}=\frac{x_2}{1-x_2^2}, \quad \{33,2\}=x_2(1-x_2^2),$$

$$\{31,3\}=\tfrac{1}{2}\frac{1}{f_2}\frac{\partial f_2}{\partial x_1}, \quad \{32,3\}=-\frac{x_2}{1-x_2^2}, \quad \{41,4\}=\tfrac{1}{2}\frac{1}{f_4}\frac{\partial f_4}{\partial x_1}.$$

Introducing these values in (6), we obtain those of $G_{\alpha\beta}$; for reasons of symmetry, we can limit our calculations to the plane $x_2=0$, and henceforth we will write simply x for x_1. We have:

$$\begin{cases} G_{11}=-\tfrac{1}{2}\dfrac{d}{dx}\left(\dfrac{1}{f_1}\dfrac{df_1}{dx}\right)+\tfrac{1}{4}\dfrac{1}{f_1^2}\left(\dfrac{df_1}{dx}\right)^2+\tfrac{1}{2}\dfrac{1}{f_2^2}\left(\dfrac{df_2}{dx}\right)^2+\tfrac{1}{4}\dfrac{1}{f_4^2}\left(\dfrac{df_4}{dx}\right)^2, \\ G_{22}=G_{33}=\tfrac{1}{2}\dfrac{d}{dx}\left(\dfrac{1}{f_1}\dfrac{df_1}{dx}\right)-1-\tfrac{1}{2}\dfrac{1}{f_1f_2}\left(\dfrac{df_2}{dx}\right)^2, \\ G_{44}=-\tfrac{1}{2}\dfrac{d}{dx}\left(\dfrac{1}{f_1}\dfrac{df_4}{dx}\right)+\tfrac{1}{2}\dfrac{1}{f_1f_4}\left(\dfrac{df_4}{dx}\right)^2. \end{cases} \tag{8}$$

All other quantities $G_{\alpha\beta}$ are zero.

We pass now to the determination of the tensor components $T_{\alpha\beta}$ and the scalar T. Because of the presence of matter and electricity, we will have to consider two different kinds of stresses in the second member of (5), *e.g.*, the stress-energy tensor of matter $\overset{m}{T}_{\alpha\beta}$ and the electromagnetic stresses $\overset{e}{T}_{\alpha\beta}$.

The second member of equations (5) can thus be written

$$\left(\overset{m}{T}_{\alpha\beta}-\tfrac{1}{2}g_{\alpha\beta}\overset{m}{T}\right)+\left(\overset{e}{T}_{\alpha\beta}-\tfrac{1}{2}g_{\alpha\beta}\overset{e}{T}\right). \tag{9}$$

We suppose that the fluid inside the sphere and the electricity which covers its surface are at rest; then

$$(10)\quad \begin{cases} u^a = \dfrac{dx_a}{ds} = 0, \quad (a=1, 2, 3), \quad u^4 = (g_{44})^{-1/2} = (f_4)^{-1/2}, \\ u_a = \sum\limits_{\beta} u^a g_{a\beta} = 0, \begin{pmatrix} a=1, 2, 3 \\ \beta = 1, 2, 3, 4 \end{pmatrix}, \; u_4 = (g_{44})^{1/2} = (f_4)^{1/2}. \end{cases}$$

(*a*) At every point the fluid is characterized by its pressure p and the density μ, which are both functions of x (or r). The mixed tensors are

$$(11)\quad \begin{cases} \overset{m}{T}{}^a_a = -pc, \quad (a=1, 2, 3), \quad \overset{m}{T}{}^4_4 = \mu, \\ \overset{m}{T} = \sum\limits_{\beta} T^\beta_\beta = \mu - 3pc, \quad (\beta = 1, 2, 3, 4). \end{cases}$$

The other components vanish. The values of the symmetrical stresses follow at once

$$(12)\quad \overset{m}{T}_{11} = \overset{m}{T}{}^1_1 g_{11} = f_1 pc, \quad \overset{m}{T}_{22} = \overset{m}{T}_{33} = f_2 pc, \quad \overset{m}{T}_{44} = f_4 \mu.$$

(*b*) The electrified sphere produces an electromagnetic field; this is determined by the four electromagnetic potentials Φ_a. In the electrostatic case $\Phi_1 = \Phi_2 = \Phi_3 = 0$; $\Phi_4 = \Phi$ remains and is a function of x or r.

The theory of the electromagnetic field furnishes generally†:

$$(13)\quad \overset{e}{T}_{a\beta} = \overset{e}{T}_{\beta a} = -\sqrt{-g} \sum_a \sum_b g^{ab} [M_{aa} M_{\beta b} + M^*_{aa} M^*_{\beta b}] - \frac{L\sigma}{2} u_a u_\beta$$

and

$$L = \sum_\nu u^\nu \Phi_\nu, \qquad (a, b, \nu = 1, 2, 3, 4).$$

In this case,

$$(14)\quad L\sigma = \sigma u^4 \Phi = \rho \Phi,$$

ρ being the electric density, and

$$(15)\quad M^*_{14} = -\frac{d\Phi}{dx} = -M^*_{41}\ ;$$

the other quantities $M^*_{a\beta}$ vanish.

Hence

$$(16)\quad M_{23} = -\frac{1}{c} f_2^2 \frac{d\Phi}{dx}\ ;$$

consequently:

†Th. de Donder. *Premiers Compléments à la Gravifique einsteinienne.* Compl. I. Ann. Observ. Roy. de Belgique. Gauthier-Villars, Paris, 1922.

$$(17)\qquad \begin{cases} \overset{e}{T}_{11} = -\dfrac{1}{2c} f_1 f_2^2 \left(\dfrac{d\Phi}{dx}\right)^2, \\ \overset{e}{T}_{22} = \overset{e}{T}_{33} = \dfrac{1}{2c} f_2^3 \left(\dfrac{d\Phi}{dx}\right)^2, \\ \overset{e}{T}_{44} = \dfrac{1}{2c} f_2^2 f_4 \left(\dfrac{d\Phi}{dx}\right)^2 - \dfrac{\rho\Phi}{2} f_4 \end{cases}$$

and for the Maxwellian stresses we have:

$$(18)\qquad \begin{cases} \overset{e}{T}{}_1^1 = -\overset{e}{T}{}_2^2 = -\overset{e}{T}{}_3^3 = \dfrac{1}{2c} f_2^2 \left(\dfrac{d\Phi}{dx}\right)^2, \ \overset{e}{T}{}_4^4 = \dfrac{1}{2c} f_2^2 \left(\dfrac{d\Phi}{dx}\right)^2 - \dfrac{\rho\Phi}{2}, \\ \overset{e}{T} = \sum_a \overset{e}{T}{}_a^a = -\dfrac{\rho\Phi}{2}. \end{cases}$$

Let us define these tensors as functions of the electric density. Therefore we have to remember that the equations of the electromagnetic field are reduced in this case to the single one

$$M^4 = \frac{\partial M_{12}}{\partial x_3} + \frac{\partial M_{23}}{\partial x_1} + \frac{\partial M_{31}}{\partial x_2} = \rho,$$

which leads to the relation

$$(19)\qquad -\frac{1}{c}\frac{d}{dx}\left(f_2^2 \frac{d\Phi}{dx}\right) = \rho.$$

This gives, on integrating between 0 and x or 0 and r:

$$(20)\qquad \frac{4\pi}{c}\left[f_2^2 \frac{d\Phi}{dx}\right]_0^x = -4\pi \int_0^r \rho r^2 dr = -e(r).$$

Hence

$$(21)\qquad \frac{d\Phi}{dx} = -c\frac{e(r)}{4\pi f_2^2},$$

where $e(r)$ is the charge on the sphere with radius r.

Outside the sphere $r = r_a$, *e.g.*, when $r \geqq r_a$, $e(r)$ is constant and equal to the total charge e on the sphere; here we have

$$(22)\qquad \left(\frac{d\Phi}{dx}\right)_{r \geqq r_a} = -c\frac{e}{4\pi f_2^2}.$$

In virtue of (21) the symmetrical tensors (17) have the form:

$$(23)\quad \begin{cases} \overset{e}{T}_{11} = -c\dfrac{f_1}{f_2^2}\dfrac{e^2(r)}{32\pi^2}, \\ \overset{e}{T}_{22} = \overset{e}{T}_{33} = c\dfrac{1}{f_2}\dfrac{e^2(r)}{32\pi^2}, \\ \overset{e}{T}_{44} = c\dfrac{f_4}{f_2^2}\dfrac{e^2(r)}{32\pi^2} - \dfrac{\rho\Phi}{2}f_4. \end{cases}$$

Putting

$$(24)\quad \epsilon^2(r) = \kappa c\frac{e^2(r)}{32\pi^2},$$

equations (5) of the gravitational field of the electrical sphere become:

$$(25)\quad \begin{cases} G_{11} + \lambda f_1 = -\dfrac{\kappa}{2}f_1(\mu - pc) + \dfrac{f_1}{f_2^2}\epsilon^2(r) + \dfrac{\kappa}{4}f_1\rho\Phi, \\ G_{22} + \lambda f_2 = -\dfrac{\kappa}{2}f_2(\mu - pc) - \dfrac{1}{f_2}\epsilon^2(r) + \dfrac{\kappa}{4}f_2\rho\Phi, \\ G_{44} - \lambda f_4 = -\dfrac{\kappa}{2}f_4(\mu + 3pc) - \dfrac{f_4}{f_2^2}\epsilon^2(r) + \dfrac{\kappa}{4}f_4\rho\Phi, \end{cases}$$

where the quantities $G_{\alpha\alpha}$ have the values (8).

It is proposed to determine the potentials f_1, f_2, f_4 in such a way, that in the *interior* solution there occur no singularities, and that the functions f and their first derivatives are continuous at the surface of the sphere. At the surface the pressure p must be zero. *Outside* the sphere μ and ρ are also zero, but $e(r)$ is equal to the charge e of the sphere. Because of the fact that electricity is only supposed to be spread over the surface, the terms in ρ, Φ, $e(r)$, *e.g.*, $\overset{e}{T}_{\alpha\alpha}$ and $\overset{e}{T}$ vanish *inside* the sphere.

II. The Exterior Field

3. *Outside* the sphere

$$\mu = p = \rho = 0,$$

and putting

$$e(r) = e, \quad \epsilon(r) = \epsilon,$$

the field equations are

$$(26)\quad \begin{cases} (a) \qquad G_{11} + \lambda f_1 = \dfrac{f_1}{f_2^2}\epsilon^2, \\ (b) \qquad G_{22} + \lambda f_2 = -\dfrac{1}{f_2}\epsilon^2, \\ (c) \qquad G_{44} - \lambda f_4 = -\dfrac{f_4}{f_2^2}\epsilon^2. \end{cases}$$

Multiplying these equations resp. by -2, $+4\frac{f_1}{f_2}$ and $-2\frac{f_1}{f_4}$ and making the combinations $a+b+c$, $a+c$, we obtain

$$4\frac{f_1}{f_2}-\frac{1}{f_2^2}\left(\frac{df_2}{dx}\right)^2-\frac{2}{f_2 f_4}\frac{df_2}{dx}\frac{df_4}{dx}-4\lambda f_1-4\frac{f_1}{f_2^2}\epsilon^2=0, \tag{27}$$

$$2\frac{d}{dx}\left(\frac{1}{f_2}\frac{df_2}{dx}\right)+\frac{3}{f_2^2}\left(\frac{df_2}{dx}\right)^2=0. \tag{28}$$

Equation (28) gives on two consecutive integrations the value of f_2,

$$f_2=k(3x+\beta)^{2/3},$$

where k and β are two integration constants.

It may be remarked that for a vanishing sphere, the field must go over into a field of Minkowski, from which we deduce that $k=1$ and consequently

$$f_2=(3x+\beta)^{2/3}=R^2; \tag{29}$$

the constant β is to be determined later by the conditions of continuity at the origin and at the surface.

In order to obtain the values of f_4 and f_1, we use the following transformation of variables:

$$f_4=\frac{c^2\zeta}{R},\ f_1=\frac{1}{R^3\zeta}. \tag{30}$$

Equation (29) gives immediately

$$R^2\frac{dR}{dx}=1. \tag{31}$$

According to (29, 30 and 31), equation (27) can be transformed into

$$\frac{d\zeta}{dx}=1-\epsilon^2R^{-2}-\lambda R^2.$$

On integrating we obtain the value of ζ:

$$\zeta+\alpha=R+\frac{\epsilon^2}{R^2}-\frac{\lambda}{3}R^3, \tag{32}$$

α being a constant which will be determined later.

The *exterior* field is now completely determined and the potentials f are (29, 30, 32):

$$\begin{cases} f_2=R^2, \\ f_4=c^2\left(1-\frac{\alpha}{R}+\frac{\epsilon^2}{R^2}-\frac{\lambda}{3}R^2\right), \\ f_1=R^{-4}\left(1-\frac{\alpha}{R}+\frac{\epsilon^2}{R^2}-\frac{\lambda}{3}R^2\right)^{-1}. \end{cases} \tag{33}$$

III. The Interior Field

4. *Inside* the sphere, all quantities $\overset{e}{T}_{\alpha\alpha}$ and $\overset{e}{T}$ are zero; hence the field equations are (25):

$$(34)\qquad \begin{cases} G_{11}+\lambda f_1 = -\dfrac{\kappa}{2} f_1(\mu - pc), \\ G_{22}+\lambda f_2 = -\dfrac{\kappa}{2} f_2(\mu - pc), \\ G_{44}-\lambda f_4 = -\dfrac{\kappa}{2} f_4(\mu + 3pc). \end{cases}$$

The determinant relation (4) becomes on logarithmic derivation

$$(35)\qquad \frac{1}{f_1}\frac{df_1}{dx}+\frac{2}{f_2}\frac{df_2}{dx}+\frac{1}{f_4}\frac{df_4}{dx}=0$$

and furthermore we have the equation of equilibrium furnished by the theorem of energy and momentum:

$$(36)\qquad -c\frac{dp}{dx}=\frac{\mu+pc}{2}\frac{1}{f_4}\frac{df_4}{dx}.$$

The problem consists now in determining the functions f_1, f_2, f_4 and p by means of the preceding equations.

We put†

$$(37)\qquad \gamma=(\mu+pc)\sqrt{f_4},$$

γ being still an unknown function of x.

From this relation we deduce

$$(38)\qquad f_4=\left(\frac{d\gamma}{dx}\Big/\frac{d\mu}{dx}\right)^2.$$

Using here the same combination as in the case of the exterior field (27, 28), equations (34) are transformed into:

$$(39)\qquad 4\frac{f_1}{f_2}-\left(\frac{1}{f_2}\frac{df_2}{dx}\right)^2-\frac{2}{f_2 f_4}\frac{df_2}{dx}\frac{df_4}{dx}+4\kappa f_1 pc-4\lambda f_1=0,$$

$$(40)\qquad 2\frac{d}{dx}\left(\frac{1}{f_2}\frac{df_2}{dx}\right)+\frac{3}{f_2^2}\left(\frac{df_2}{dx}\right)^2+2\kappa f_1(\mu+pc)=0.$$

For the integration of these equations we use the transformation:

$$(41)\qquad f_2=r'^2,\ f_4=\frac{c^2\psi}{r'},\ f_1=\frac{1}{r'^3\psi},$$

†Th. de Donder. Compl. IV. *Gravifique einsteinienne.*

and equations (39, 40) become:

$$r'^2 \frac{dr'}{dx} \cdot \frac{d\psi}{dx} = r'^{-2} + \kappa c^{-1} \gamma \psi^{-1/2} r'^{1/2} - \kappa\mu - \lambda, \tag{42}$$

$$2\psi \frac{d}{dx}\left(r'^2 \frac{dr'}{dx}\right) = -\kappa c^{-1} \gamma \psi^{-1/2} r'^{1/2}. \tag{43}$$

On addition of these equations and multiplication of the result by $r'^2 \frac{dr'}{dx}$, we obtain by integration

$$\psi\left(r'^2 \frac{dr'}{dx}\right)^2 = r' - \frac{\lambda}{3} r'^3 - \kappa \int_0^{r'} \mu r'^2 dr' + \delta, \tag{44}$$

δ being an integration constant.

In multiplying equation (43) by $r'^2 \frac{dr'}{dx}$ and dividing the result by the power 3/2 of (44), we obtain the differential equation:

$$\frac{2\frac{d}{dx}\left(r'^2 \frac{dr'}{dx}\right)}{\left(r'^2 \frac{dr'}{dx}\right)^2} = -\frac{\kappa c^{-1} \gamma r'^{5/2}}{\left(r' - \frac{\lambda}{3} r'^3 - \kappa \int_0^{r'} \mu r'^2 dr' + \delta\right)^{3/2}} \frac{dr'}{dx}$$

and thence, by integration:

$$\frac{2}{r'^2 \frac{dr'}{dx}} - \frac{2}{\left(r'^2 \frac{dr'}{dx}\right)_a} = -\frac{\kappa}{c} \int_{r'}^{r_a'} \frac{\gamma r'^{5/2} dr'}{\left(r' - \frac{\lambda}{3} r'^3 - \kappa \int_0^{r'} \mu r'^2 dr' + \delta\right)^{3/2}}. \tag{45}$$

5. *The conditions of continuity at the surface and at the origin of the sphere.* At the surface, we put

$$r = r_a,\ R = R_a,\ r' = r',\ x = x_a,\ \psi = \psi_a, \text{ etc.}$$

The integration constants will be determined by the conditions that there be no singularities inside the sphere, and that the potentials f (29, 33, 41) and their first derivatives be continuous at the surface.

From the potentials we have

$$\begin{cases} R_a = r_a', \\ \beta = r_a'^3 - r_a^3, \\ \alpha = r_a' + \frac{\epsilon^2}{r_a'} - \frac{\lambda}{3} r_a'^3 - \psi_a, \end{cases} \tag{46}$$

and from their derivatives

$$(47)\qquad \begin{cases} \left(\dfrac{dr'}{dx}\right)_a = \dfrac{1}{r_a'^2}, \\ \left(\dfrac{d\psi}{dr'}\right)_a = 1 - \dfrac{\epsilon^2}{r_a'^2} - \lambda r_a'^2. \end{cases}$$

From (44) and (47) we deduce

$$(48)\qquad \psi_a = r_a' - \frac{\lambda}{3} r_a'^3 - \kappa \int_0^{r_a'} \mu r'^2 dr' + \delta$$

and from (37), (41) and (48), remarking that $p=0$ at the surface,

$$(49)\qquad \gamma_a = c\mu_a \sqrt{\frac{\psi_a}{r_a'}}.$$

In virtue of (31) and (47) equation (45) becomes

$$(50)\qquad \frac{R^2 dR}{r'^2 dr'} = 1 - \frac{\kappa}{2c} \int_{r'}^{r_a} \frac{\gamma r'^{5/2} dr'}{\left(r' - \dfrac{\lambda}{3} r'^3 - \kappa \displaystyle\int_0^{r'} \mu r'^2 dr' + \delta\right)^{3/2}}$$

On integrating between r' and r_a', we obtain by (46):

$$(51)\qquad R^3 = r'^3 + \frac{3\kappa}{2c} \int_{r'}^{r_a'} r'^2 dr' \int_{r'}^{r_a'} \frac{\gamma r'^{5/2} dr'}{\left(r' - \dfrac{\lambda}{3} r'^3 - \kappa \displaystyle\int_0^{r'} \mu r'^2 dr' + \delta\right)^{3/2}}.$$

In this way, the conditions at the surface are fulfilled. We have still to determine the constants β and δ by the continuity at the origin. There we have $x=r=0$ and we have also to admit that $r'=0$. Hence by (51) we obtain β(29):

$$(52)\qquad R_0^3 = \beta = \frac{3\kappa}{2c} \int_0^{r_a'} r'^2 dr' \int_{r'}^{r_a'} \frac{\gamma r'^{5/2} dr'}{\left(r' - \dfrac{\lambda}{3} r'^3 - \kappa \displaystyle\int_0^{r'} \mu r'^2 dr' + \delta\right)^{3/2}}.$$

The constant δ is determined by the hypothesis that the pressure and the density are finite and positive at the centre. Consequently f_4 there is finite and different from zero and according to Schwarzschild* $\delta = 0$.

The constant α is now known also and has the value:

$$(53)\qquad \alpha = \frac{\epsilon^2}{r_a'} + \kappa \int_0^{r_a'} \mu r'^2 dr',$$

which contains two parts respectively due to electricity and matter.

*K. Schwarzschild. Sitzungsber. Akad. Berlin, March 1916.

6. *Transformation of variables.* Writing

$$\omega = r' - \frac{\lambda}{3} r'^3 - \kappa \int_0^{r'} \mu r'^2 dr', \tag{54}$$

we obtain, according to (50), the generalization of the transformation of DeDonder:

$$r^3 = 3x = r'^3 - \frac{3\kappa}{2c} \int_0^{r'} r'^2 dr' \int_{r'}^{r'_a} \frac{\gamma r'^{5/2} dr'}{\omega^{3/2}}, \tag{55}$$

which determines the relation between the variables r and r' used respectively outside and inside the sphere. This relation must be satisfied at the surface of the sphere:

$$r_a^3 = r_a'^3 - \frac{3\kappa}{2c} \int_0^{r'_a} r'^2 dr' \int_{r'}^{r'_a} \frac{\gamma r'^{5/2} dr'}{\omega^{3/2}}. \tag{56}$$

Substituting in (56) the values of β (46) and (52), we find that equation (55) is identically satisfied at the surface. That means that the variable r' can be chosen in such a way that

$$\beta = 0 \tag{57}$$

and hence (46) and (2, 29) give at once

$$r_a = r'_a = R_a \tag{58}$$

and

$$R = r, \tag{59}$$

so that outside the sphere we may use the variable r.

7. *The potentials.* Inside the sphere, the functions f are determined as follows (41, 44, 50):

$$\begin{cases} f_2 = r'^2, \\ f_4 = c^2 \dfrac{\omega}{r'} \left(1 - \dfrac{\kappa}{2c} \displaystyle\int_{r'}^{r'_a} \dfrac{\gamma r'^{5/2} dr'}{\omega^{3/2}} \right)^2, \\ f_1 = \dfrac{c^2}{f_4 f_2^2}. \end{cases} \tag{60}$$

The potential f_4 (and also f_1) will be completely determined when the function γ is known. Therefore we introduce (60) in (38) and obtain

$$\frac{d\gamma}{dx} = c\omega^{1/2} r'^{-1/2} \left(1 - \frac{\kappa}{2c} \int_{r'}^{r'_a} \frac{\gamma r'^{5/2} dr'}{\omega^{3/2}} \right) \frac{d\mu}{dx}. \tag{61}$$

On derivation with respect to r', we get the generalization of the linear differential equation of M. Brillouin†:

†M. Brillouin. Comptes Rendus Acad. Sciences, Paris. 19 Juin 1922.

$$\frac{d}{dr'}\left(\frac{d\gamma}{dx}\Big/\frac{d\mu}{dx}\cdot\sqrt{\frac{r'}{\omega}}\right)=\frac{\kappa}{2}\cdot\frac{r'^{5/2}}{\omega^{3/2}}\gamma. \tag{62}$$

The integration of this equation will introduce two constants, which will be determined by the conditions at the surface. By (49, 38, 48, 60) we have at once these conditions:

$$\gamma_a=c\mu_a\sqrt{\frac{\omega_a}{r'_a}},\quad \left(\frac{d\gamma}{dx}\right)_a=c\left(\frac{d\mu}{dx}\right)_a\sqrt{\frac{\omega_a}{r'_a}}. \tag{63}$$

8. *The pressure.* Equations (37, 38) give

$$p=-\frac{1}{c}\left(\mu-\gamma\frac{d\mu}{dx}\Big/\frac{d\gamma}{dx}\right), \tag{64}$$

where γ and $\dfrac{d\gamma}{dx}$ are determined by (62) and (63).

We see that the pressure vanishes at the surface.

IV. The Solution

The complete solution, which includes all the particular cases of the sphere investigated up to the present time, is the following:

(*a*) *Outside* the sphere, the field is characterized in polar coordinates r, θ, ϕ by (2, 3, 34, 59):

$$\delta s^2\equiv-\frac{1}{\nu}\delta r^2-r^2(\delta\theta^2+\sin^2\theta\,\delta\phi^2)+c^2\nu\,\delta t^2 \tag{65}$$

where

$$\nu\equiv\left(1-\frac{a}{r}+\frac{\epsilon^2}{r^2}-\frac{\lambda}{3}r^2\right)=\left(1-\frac{2fM}{c^2r}+\frac{f.e^2}{4\pi c^4.r^2}-\frac{\lambda}{3}r^2\right), \tag{66}$$

f being the Gaussian constant and M the mass of the sphere.

(*b*) *Inside* the sphere, the field is determined in coordinates r', θ, ϕ by (2, 3, 60, 53, 54):

$$\delta s^2\equiv-\frac{\delta r'^2}{1-\dfrac{\kappa}{r'}\displaystyle\int_0^{r'}\mu r'^2dr'-\dfrac{\lambda}{3}r'^2}-r'^2(\delta\theta^2+\sin^2\theta\,\delta\phi^2)+f_4\delta t^2. \tag{67}$$

9. *Remark.* If in (65), we abolish the sphere, *e.g.*, if we put $a=\epsilon=0$, we obtain the solution for empty space

$$\delta s^2\equiv-\frac{\delta r^2}{1-\dfrac{\lambda}{3}r^2}-r^2(\delta\theta^2+\sin^2\theta\,\delta\phi^2)+c^2\left(1-\frac{\lambda}{3}r^2\right)\delta t^2. \tag{68}$$

Putting

$$r\equiv\left(\frac{\lambda}{3}\right)^{-1/2}\sin\chi,$$

equation (67) becomes

$$\delta s^2 \equiv -\frac{3}{\lambda}(\delta\chi^2+\sin^2\chi\delta\theta^2+\sin^2\chi\sin^2\theta\delta\phi^2)+c^2\cos^2\chi\delta t^2.$$

This is DeSitter's† quadratic form for a curved empty space. The geometrical part

$$\delta\sigma^2 \equiv \frac{3}{\lambda}(\delta\chi^2+\sin^2\chi\delta\theta^2+\sin^2\chi\sin^2\theta\delta\phi^2)$$

is the characteristic line-element of the geometry in spherical space, the radius of curvature being $\sqrt{\frac{3}{\lambda}}$.

10. *The motion of a particle in the field of the charged sphere.* We will investigate the motion of a particle having both inertia mass and charge in the field (65). The track of such a particle is determined by the general equations

$$F^i \equiv F^i_{(m)}+F^i_{(e)}=0, \quad (i=1, 2, 3, 4), \tag{69}$$

which express that the total "force" is zero and which are derived from the equations of conservation of energy and momentum.

We have

$$\begin{cases} F^i_{(m)} \equiv \mu A^i, \\ F^i_{(e)} \equiv \dfrac{L\sigma}{2}A^i-\sigma\sum\limits_\alpha\sum\limits_\beta u^\alpha M^*_{\alpha\beta}g^{\beta i}, \end{cases} \tag{70}$$

and the contravariant acceleration vector

$$A^i \equiv \frac{d^2x_i}{ds^2}+\sum_\alpha\sum_\beta\{\alpha\beta,i\}\frac{dx_\alpha}{ds}\frac{dx_\beta}{ds}. \tag{71}$$

Hence the general equations of motion of the charged particle are‡

$$\left(\mu-\frac{L\sigma}{2}\right)A^i=\sigma\sum_\alpha\sum_\beta u^\alpha M^*_{\alpha\beta}g^{\beta i} \tag{72}$$

in which $\left(\mu-\frac{L\sigma}{2}\right)$ and σ are both multipliers, the ratio of which is invariant. In the case where we suppose that there is no "substantial" action due to electricity, the term in $\frac{L\sigma}{2}$ in the first member vanishes and the equations of motion are simply

$$\mu A^i=\sigma\sum_\alpha\sum_\beta u^\alpha M^*_{\alpha\beta}g^{\beta i}$$

where μ/σ is still invariant.

†W. de Sitter. Monthly Notices, Royal Ast. Soc., November 1917.
A. S. Eddington. *The Mathematical theory of Relativity*, p. 161. Cambridge, 1923.
‡Th. de Donder. Compl. II. *Gravifique einsteinienne.*

The Christoffel symbols which do not vanish have the values

$$\{11, 1\} = -\tfrac{1}{2}\frac{\nu'}{\nu}, \ \{22, 1\} = -r\nu, \ \{33, 1\} = -r\nu \sin^2\theta, \ \{44, 1\} = \frac{c^2}{2}\nu'\nu,$$

$$\{41, 4\} = \tfrac{1}{2}\frac{\nu'}{\nu}, \ \{21, 2\} = \{31, 3\} = \frac{1}{r}, \ \{33, 2\} = -\sin\theta\cos\theta, \ \{32, 3\} = \operatorname{cotg}\theta;$$

the accent denotes a derivative with respect to r.

The quantities in the second member of (72) are known from (14), (15), (22), (29), (59), (65). Furthermore in order to restore our ordinary units, the ratio $\left(\mu - \frac{L\sigma}{2}\right) : \sigma$ is equivalent to $\delta E'/\delta e'$, where

$$\int \delta E' \equiv \int \left(c^2 \delta m' - \frac{\Phi}{2}\delta e'\right) \equiv E',$$

$$\int \delta e' \equiv \int \rho\, \delta x_1 \delta x_2 \delta x_3 \equiv e',$$

m' and e' denoting the mass and the charge of the particle.

Hence the equations of motion are:

$$(73)\begin{cases} (a)\ \dfrac{d^2r}{ds^2} - \dfrac{1}{2}\dfrac{\nu'}{\nu}\left(\dfrac{dr}{ds}\right)^2 - r\nu\left(\dfrac{d\theta}{ds}\right)^2 - r\nu\sin^2\theta\left(\dfrac{d\phi}{ds}\right)^2 + \dfrac{c^2}{2}\nu'\nu\left(\dfrac{dt}{ds}\right)^2 = \dfrac{c\nu ee'}{4\pi r^2 E'}\dfrac{dt}{ds}, \\ (b)\ \dfrac{d^2\theta}{ds^2} + \dfrac{2}{r}\cdot\dfrac{dr}{ds}\cdot\dfrac{d\theta}{ds} - \sin\theta\cos\theta\left(\dfrac{d\phi}{ds}\right)^2 = 0, \\ (c)\ \dfrac{d^2\phi}{ds^2} + \dfrac{2}{r}\dfrac{dr}{ds}\cdot\dfrac{d\phi}{ds} + 2\operatorname{cotg}\theta\dfrac{d\theta}{ds}\dfrac{d\phi}{ds} = 0, \\ (d)\ \dfrac{d^2t}{ds^2} + \dfrac{\nu'}{\nu}\dfrac{dr}{ds}\cdot\dfrac{dt}{ds} = \dfrac{ee'}{4\pi r^2 c . E'\nu}\dfrac{dr}{ds}. \end{cases}$$

Choosing the coordinates in such a way that initially the particle moves in the plane $\theta = \frac{\pi}{2}$, then $\cos\theta = 0$, $\frac{d\theta}{ds} = 0$ initially and (73 b), $\frac{d^2\theta}{ds^2} = 0$; the particle remains moving in that plane. The three remaining equations may be simplified. By (73 c) and (73 d), we find at once the invariants or integrals:

$$(74) \qquad r^2\frac{d\phi}{ds} = \beta,$$

$$(75) \qquad c^2\nu\frac{dt}{ds} = \gamma - \frac{c . ee'}{4\pi r^2 E'},$$

β and γ representing two constants of integration.

Eliminating e, e', E', between equations (73), we find the third integral:

$$(76) \qquad -\frac{1}{\nu}\left(\frac{dr}{ds}\right)^2 - r^2\left(\frac{d\phi}{ds}\right)^2 + c^2\nu\left(\frac{dt}{ds}\right)^2 = 1.$$

Eliminating ds and dt between the three last equations, and putting $\frac{1}{r} \equiv u$, we obtain the differential equation of the geometrical orbits of the particle:

$$(77) \qquad \beta^2\left(\frac{du}{d\phi}\right)^2 = -\left(1 - au + \epsilon^2 u^2 - \frac{\lambda}{3}u^{-2}\right)(1+\beta^2u^2) + \left(\frac{\gamma}{c} - \frac{ee'}{4\pi E'}u^2\right)^2.$$

Neglecting the curvature of the universe ($\lambda = 0$) and the substantial action of electricity, then $E' = c^2 m'$ and equation (77) reduces at once to the one obtained by G. B. Jeffery† for the motion of a charged particle round an atomic nucleus; in this case u can be expressed as an elliptic function of ϕ.

11. *The velocity of the particle.* Eliminating ds between (74), (75), (76), we have the relation which determines the velocity of the particle in the field; it may be written as follows:

$$(78) \quad \left(\frac{dr}{dt}\right)^2 + r^2\left(\frac{d\phi}{dt}\right)^2 = c^2\nu^2\left\{1 - \left[\nu - \beta^2\left(\frac{a}{r^3} - \frac{\epsilon^2}{r^4} + \frac{\lambda}{3}\right)\right]\left(\frac{\gamma}{c} - \frac{ee'}{4\pi r^2 E'}\right)^{-2}\right\}.$$

By (74) and (75), the equation of areas is

$$(79) \qquad r^2\frac{d\phi}{dt} = \beta c^2 \nu\left(\gamma - \frac{c.ee'}{4\pi r^2 E'}\right)^{-1}.$$

12. *Particular case of the uncharged sphere and particle.* Now we have to put $\epsilon = e = e' = 0$ in ν(66) and the equations of motion are geodesics of the field. The differential equation of the geometrical orbits of the particle will be (77):

$$\left(\frac{du}{d\phi}\right)^2 = \frac{1}{\beta^2}\left(\frac{\gamma^2}{c^2} - 1\right) + \frac{\lambda}{3} + \frac{a}{\beta^2}u - u^2 + au^3 + \frac{\lambda u^{-2}}{3\beta^2},$$

or after differentiation,

$$(80) \qquad \frac{d^2u}{d\phi^2} + u = \frac{1}{2}\frac{a}{\beta^2} + \frac{3}{2}au^2 - \frac{\lambda u^{-3}}{3\beta^2}.$$

Neglecting for a moment the curvature of the universe, the corresponding equation determines the Einsteinian orbit of a particle with moving perihelion. The Newtonian orbit under the influence of gravitation is obtained by considering only the term $\frac{1}{2}\frac{a}{\beta^2}$ in the second member of (80). Taking into account only the last term in (80), we have the orbit of a particle in De Sitter's empty space. This orbit is the same as if it were produced by a repulsive force acting pro-

†G. B. Jeffery. Proc. Royal Soc., London, 99, p. 129, 1921.

portionally to the distance from the origin. Hence this "repulsion" will neutralize the Newtonian gravitation† at a distance r from the origin given by (80):

$$\frac{\alpha}{2} = \frac{\lambda}{3} u^{-3},$$

or

(81) $$r^3 = \frac{3}{2}\frac{\alpha}{\lambda},$$

where $\frac{\alpha}{2}$ is the gravitational mass of the sphere (1, 47 km. in the case of the Sun).

13. *The equation of living forces or of energy.* Equation (78) becomes (66):

(82) $$\left(\frac{dr}{dt}\right)^2 + r^2\left(\frac{d\phi}{dt}\right)^2 = c^2\left(1-\frac{\alpha}{r}-\frac{\lambda}{3}r^2\right)^2\left[1-\frac{c^2}{\gamma^2}\left(1-\frac{\alpha}{r}-\frac{\lambda}{3}r^2\right)+\frac{\beta^2c^2}{\gamma^2}\left(\frac{\alpha}{r^3}+\frac{\lambda}{3}\right)\right],$$

which determines the total velocity of the particle.

In the first member, we have the expression of the "Euclidean" velocity. Hence, by comparison with classical mechanics, we are led to consider

$$c^2\left(1-\frac{c^2}{\gamma^2}+\frac{\beta^2c^2\lambda}{3\gamma^2}\right) \equiv h$$

as the constant of living forces.

Furthermore, by (79) and (66), the equation of areas is

(83) $$r^2\frac{d\phi}{dt} = \frac{\beta c^2}{\gamma}\left(1-\frac{\alpha}{r}-\frac{\lambda}{3}r^2\right).$$

The "radial velocity" of a particle ($d\phi=0$) is determined (82, 83) by:

(84) $$\left(\frac{dr}{dt}\right)^2 = c^2\left(1-\frac{\alpha}{r}-\frac{\lambda}{3}r^2\right)^2\left[1-\frac{c^2}{\gamma^2}\left(1-\frac{\alpha}{r}-\frac{\lambda}{3}r^2\right)\right].$$

From this equation, or also from (73), we deduce the "radial acceleration"

(85) $$\frac{d^2r}{dt^2} = c^2\left(\frac{\alpha}{2r^2}-\frac{\lambda}{3}r\right)\left(1-\frac{\alpha}{r}-\frac{\lambda}{3}r^2\right)\left[2-\frac{3c^2}{\gamma^2}\left(1-\frac{\alpha}{r}-\frac{\lambda}{3}r^2\right)\right];$$

this vanishes at the neutral point (81), where the velocity (84) is then constant.

14. *The tracks of light-rays.* The equation of the track of a light-ray is determined by the condition $ds=0$. In order to obtain this equation, we introduce an auxiliary parameter l, which is invariant for all transformations of coordinates.‡ The calculations are analogous to those performed in the case of the particle. Now we have the invariants:

†Compare L. Silberstein. Monthly Notices, Royal Ast. Soc., 84, p. 364, 1924.
‡Th. de Donder. *Gravifique einsteinienne;* equation (293).

$$(86)\qquad \begin{cases} -\dfrac{1}{\nu}\left(\dfrac{dr}{dl}\right)^2 - r^2\left(\dfrac{d\phi}{dl}\right)^2 + c^2\nu\left(\dfrac{dt}{dl}\right)^2 = 0, \\ r^2\dfrac{d\phi}{dl} = \beta, \\ c^2\left(1-\dfrac{\alpha}{r}+\dfrac{\epsilon^2}{r^2}-\dfrac{\lambda}{3}r^2\right)\dfrac{dt}{dl} = \gamma. \end{cases}$$

Eliminating dl and dt ,we find the differential equation of the light-track

$$(87)\qquad \frac{d^2u}{d\phi^2} + u = \frac{3}{2}\,\alpha u^2 - 2\epsilon^2 u^3.$$

Neglecting the charge of the sphere ($\epsilon = 0$), we have the equation ordinarily used in the case of the field of the Sun. Abolishing the sphere ($\alpha = \epsilon = 0$), the light-tracks are straight lines in deSitter's empty space†. Equation (87) shows that the charge tends to counteract the effect of gravitation in the deflection of a light-ray. There would be compensation at (66):

$$(88)\qquad r = \frac{e^2}{6\pi c^2 M}.$$

The velocity of light is given (86) by the relation:

$$(89)\qquad \left(\frac{dr}{dt}\right)^2 + r^2\left(\frac{d\phi}{dt}\right)^2\left(1-\frac{\alpha}{r}+\frac{\epsilon^2}{r^2}-\frac{\lambda}{3}r^2\right) = c^2\left(1-\frac{\alpha}{r}+\frac{\epsilon^2}{r^2}-\frac{\lambda}{3}r^2\right)^2;$$

the two last equations (86) give furthermore the condition

$$(90)\qquad r^2\frac{d\phi}{dt} = \frac{\beta c^2}{\gamma}\left(1-\frac{\alpha}{r}+\frac{\epsilon^2}{r^2}-\frac{\lambda}{3}r^2\right),$$

an equation which is analogous to the equation of areas (79) and (83) in the case of a particle.

†A. S. Eddington. *l.c.*, p. 163.

THE GRAVITATIONAL FIELD OF n MOVING PARTICLES IN THE THEORY OF RELATIVITY

BY DR. H. C. LEVINSON,
Chicago, Illinois, U.S.A.

The development of the Theory of Relativity in the direction of Dynamical Astronomy requires the solution of Einstein's field equations of gravitation, subject to physical conditions that correspond to the various systems that are possible. This paper attempts to give a solution, without approximations, for the case of n particles in motion. This solution is in the form of power series in the masses. Formulae are given for the computation of as many of the coefficients as may be desired. The solution provides, furthermore, the means of studying the types of terms that are present in the equations of motion, a study which is essential for the consideration of periodic motion.

Consider an observer at rest in a system of coordinates x_1, x_2, x_3, x_4, of which the first three are rectangular space coordinates, the last a time coordinate. In this system are n material points, with the finite masses $m_1, \ldots, m_n$, which move in any assigned manner. The motions are expressible, relative to the observer's system, as functions of the time coordinate x_4. The space coordinates of the n particles at the time x_4, will be denoted by

$$x_\alpha^{(i)} = \phi_\alpha^{(i)}(x_4), \quad (\alpha = 1, 2, 3;\ i = 1, \ldots, n).$$

The subscripts refer to the coordinates, the superscripts to the masses. The gravitational field produced in this system is defined by a linear element

$$ds^2 = \sum_{\mu=1}^{4} \sum_{\nu=1}^{4} g_{\mu\nu}\, dx_\mu\, dx_\nu. \tag{1}$$

The coefficients $g_{\mu\nu}$ of the form are particular solutions of the field equations

$$G_{\mu\nu} = -\sum_{\alpha=1}^{4} \frac{\partial}{\partial x_\alpha} \left\{ \begin{matrix} \mu\nu \\ \alpha \end{matrix} \right\} + \sum_{\alpha=1}^{4} \sum_{\beta=1}^{4} \left\{ \begin{matrix} \mu\nu \\ \beta \end{matrix} \right\} \left\{ \begin{matrix} \nu\beta \\ \alpha \end{matrix} \right\} + \frac{\partial^2}{\partial x_\mu\, \partial x_\nu} \log \sqrt{-g}$$
$$- \sum_{\alpha=1}^{4} \left\{ \begin{matrix} \mu\nu \\ \alpha \end{matrix} \right\} \frac{\partial}{\partial x_\alpha} \log \sqrt{-g} = 0, \quad (\mu, \nu = 1, \ldots, 4), \tag{2}$$

such that a set of physical conditions at the boundaries, which we may denote by B, are satisfied. When the $g_{\mu\nu}$ are thus determined, the equations of motion of a material point in the system are explicitly known.

This problem represents the actual problem of n bodies when the $\phi_\alpha^{(i)}(x_4)$ are so determined that each particle moves in a geodesic of the system. The determination of these functions in this way evidently requires the simultaneous solution of the $4n$ equations of motion, to which we are led if we write the four equations of motion of each particle. This system of equations reduces to the system of $3n$ equations of the Newtonian dynamics, if the approximation corresponding to this case is introduced.

The point of view of this paper is to consider the $g_{\mu\nu}$ as expansible in convergent power series in the quantities $m_1, \ldots, m_n$. When these power series are substituted in the equations $G_{\mu\nu}=0$, and the coefficients of the various powers of the m_i are equated to zero, the equations to be solved become linear. The solutions permit the calculation of any desired number of coefficients of the power series. The $g_{\mu\nu}$ may thus be considered as determined. Between the solution given and any other there must exist a transformation of coordinates.

The gravitational field of n moving spheres has been treated by Einstein,* and by Droste,* both papers giving approximate solutions which are sufficiently accurate for numerical purposes. It will be shown that neither of these results admits of an application to the case of n particles that gives all of the terms of the first order in the masses.

The $g_{\mu\nu}$ of our problem have the form:

$$g_{\mu\nu}=A_{\mu\nu}^{(0,\ldots,0)}+\sum_{a_1=0}^{\infty}\cdots\sum_{a_n=0}^{\infty}A_{\mu\nu}^{(a_1,\ldots,a_n)}m_1^{a_1}\cdots m_n^{a_n}, \tag{3}$$

$$(\mu,\nu=1,\ldots,4),\quad (a_1,\ldots,a_n\neq 0,\ldots,0).$$

The $A_{\mu\nu}^{(a_1,\ldots,a_n)}$ are functions of the variables $x_1, \ldots, x_4$.

Before considering the form of the boundary conditions B, a few preliminaries are necessary. We shall consider our time variable x_4 to be related to the measured time t by

$$x_4=ict,$$

where $i=\sqrt{-1}$ and c = velocity of light in vacuo. Then the linear element of a Galilean field has the form

$$ds^2=-\sum_{a=1}^{4}dx_a^{\,2}=-\sum_{\mu=1}^{4}\sum_{\nu=1}^{4}\delta_{\mu\nu}dx_\mu dx_\nu, \tag{4}$$

where

$$\delta_{\mu\nu}=\begin{cases}1 & (\mu=\nu),\\ 0 & (\mu\neq\nu).\end{cases}$$

The field of one particle of mass m, fixed at the origin of coordinates, is given by the well-known Schwarzschild form

$$ds^2=-\sum_{a=1}^{3}dx_a'^2-\frac{\psi}{r'^2}\left(\sum_{a=1}^{3}x_a'\,dx_a'\right)^2+\left(-1+\frac{2m}{r'}\right)dx_4^2,$$

where

$$r'=\sqrt{\sum_{a=1}^{3}x_a'^2},\qquad \psi=\frac{2m}{r'-2m}.$$

*See de Sitter; Monthly Notices, Royal Ast. Soc., 1916, vol. LXXVI, No. 9.

A simpler form is obtained if we apply the transformation

$$x_i' = \left(1+\frac{m}{2r}\right)x_i, \quad (i=1,\ 2,\ 3).$$

Then

$$ds^2 = -\left(1+\frac{m}{2r}\right)^4 \sum_{\alpha=1}^{3} dx_\alpha^2 - \frac{\left(1-\frac{m}{2r}\right)^2}{\left(1+\frac{m}{2r}\right)^2} dx_4^2, \tag{5}$$

where

$$r=\sqrt{\sum_{\alpha=1}^{3} x_\alpha^2}.$$

If the particle is at the fixed point $(d_1,\ d_2,\ d_3)$, the field is found by the transformation

$$x_i = y_i - d_i, \quad (i=1,\ 2,\ 3).$$

After this transformation is applied to (5), if we change the notation of our variables from y_i back to x_i, the form is exactly that of (5), with r_0 substituted for r, where

$$r_0=\sqrt{\sum_{\alpha=1}^{3} (x_i-d_i)^2}.$$

The conditions B, which define the physical problem, may be written as follows:

(*a*) The $g_{\mu\nu}$ are continuous functions in any domain of space D_1, which contains none of the particles, and can be expanded as power series in the m_i, convergent in some domain D contained in the domain D_1. The $g_{\mu\nu}$ admit n singular points, corresponding to the space positions of the n particles.

(*b*) If all of the masses m_i approach zero, the $g_{\mu\nu}$ approach the values of a Galilean field, that is, the values given by (4).

(*c*) If any one of the space variables x_1, x_2, x_3 increases without limit, the $g_{\mu\nu}$ approach the values given by (4).

(*d*) For x_4 = a constant, if the coordinates of all of the particles are made to approach the fixed values $(d_1,\ d_2,\ d_3)$ respectively, the $g_{\mu\nu}$ approach those of the field of a fixed particle at the point $(d_1,\ d_2,\ d_3)$, that is, the values given by (5), in which m is replaced by $\sum_{i=1}^{n} m_i$, and r is replaced by r_0.

The condition (*b*) determines at once the values of the functions $A_{\mu\nu}^{(0,\dots,0)}$ as the following:

$$A_{\mu\nu}^{(0,\dots,0)} = -\delta_{\mu\nu}, \quad (\mu,\nu=1,\ \dots,\ 4).$$

If we had not taken the variable x_4 as purely imaginary, the perfect symmetry in the four variables would have been destroyed at this point. It would then be impossible to write the equations in brief form. The expressions for the $g_{\mu\nu}$ now become

$$(6)\qquad g_{\mu\nu} = -\delta_{\mu\nu} + \sum_{\alpha_1=0}^{\infty} \cdots \sum_{\alpha_n=0}^{\infty} A_{\mu\nu}^{(\alpha_1,\ldots,\alpha_n)} m_1^{\alpha_1} \ldots m_n^{\alpha_n},$$

$$(\mu,\nu=1,\ldots,4),\qquad (\alpha_1,\ldots,\alpha_n \neq 0,\ldots,0).$$

We now substitute these values in the equations (2), and equate to zero each coefficient of the m_i. We denote the resulting coefficient of $m_1^{\alpha_1}\ldots m_n^{\alpha_n}$ by $G_{\mu\nu}^{(\alpha_1,\ldots,\alpha_n)}$. We shall say that the equations $G_{\mu\nu}^{(\alpha_1,\ldots,\alpha_n)}=0$ are of order k if

$$\sum_{i=1}^{n} \alpha_i = k.$$

There will be as many independent sets of equations of order k as there are solutions in positive integers, including zero, of this equation. The equations (2) are then replaced by the set

$$(7)\qquad G_{\mu\nu}^{(\alpha_1\ldots\alpha_n)} = -\tfrac{1}{2}\sum_{\alpha=1}^{4}\left(\frac{\partial^2 A_{\mu\nu}^{(\alpha_1\ldots\alpha_n)}}{\partial x_\alpha^2} + \frac{\partial^2 A_{\alpha\alpha}^{(\alpha_1\ldots\alpha_n)}}{\partial x_\mu \partial x_\nu} - \frac{\partial^2 A_{\mu\alpha}^{(\alpha_1\ldots\alpha_n)}}{\partial x_\nu \partial x_\alpha} - \frac{\partial^2 A_{\nu\alpha}^{(\alpha_1\ldots\alpha_n)}}{\partial x_\mu \partial x_\alpha}\right)$$
$$+\psi_{\mu\nu}^{(\alpha_1\ldots\alpha_n)} = 0,\qquad (\mu,\nu=1,\ldots,4).$$

The $\psi_{\mu\nu}^{(\alpha_1,\ldots,\alpha_n)}$ denote functions of the $A_{\mu\nu}^{(\alpha_1,\ldots,\alpha_n)}$ of order less than k. For $k=1$ these functions are all zero. If, then, the equations (7) are solved step by step, commencing with $k=1$, the $\psi_{\mu\nu}^{(\alpha_1,\ldots,\alpha_n)}$, for each k, are known functions.

It is well known that of the ten field equations of gravitation, six only are independent, corresponding to the indetermination of the four variables. In the case of the problem treated here, this is equivalent to the existence of four identical relations among the equations of each set of each order. In fact, it is found that the explicit relations are

$$(8)\qquad \sum_{\alpha=1}^{4}\frac{\partial}{\partial x_i} G_{\alpha\alpha}^{(\alpha_1\ldots\alpha_n)} - 2\sum_{\beta=1}^{4}\frac{\partial}{\partial x_\beta} G_{i\beta}^{(\alpha_1\ldots\alpha_n)} = \sum_{\alpha=1}^{4}\frac{\partial}{\partial x_i}\psi_{\alpha\alpha}^{(\alpha_1\ldots\alpha_n)} - 2\sum_{\beta=1}^{4}\frac{\partial}{\partial x_\beta}\psi_{i\beta}^{(\alpha_1\ldots\alpha_n)} = 0,$$

$$(i=1,\ldots,4).$$

These identities permit us to assign four arbitrary relations among the $A_{\mu\nu}^{(\alpha_1,\ldots,\alpha_n)}$, provided that the solution thus obtained from (7) satisfies identically the four assigned relations. The boundary conditions B must, of course, also be satisfied. The choice of these relations must in addition be such that the equations may be easily integrated. The following choice satisfies all of these conditions:

$$(9)\qquad \left\{\begin{aligned} & A_{11}^{(\alpha_1,\ldots,\alpha_n)} = A_{22}^{(\alpha_1,\ldots,\alpha_n)} = A_{33}^{(\alpha_1,\ldots,\alpha_n)} \equiv A^{(\alpha_1,\ldots,\alpha_n)}, \\ & \sum_{\alpha=1}^{3}\frac{\partial A_{\alpha 4}^{(\alpha_1,\ldots,\alpha_n)}}{\partial x_\alpha} = 0, \\ & \frac{\partial^2 A_{12}^{(\alpha_1,\ldots,\alpha_n)}}{\partial x_1 \partial x_2} + \frac{\partial^2 A_{13}^{(\alpha_1,\ldots,\alpha_n)}}{\partial x_1 \partial x_3} + \frac{\partial^2 A_{23}^{(\alpha_1,\ldots,\alpha_n)}}{\partial x_2 \partial x_3} = 0. \end{aligned}\right.$$

Using the Laplacian notation

$$\Delta F=\sum_{\alpha=1}^{3}\frac{\partial^2}{\partial x_\alpha{}^2}F,$$

and introducing (9) into the equations (7), we are able to write them in the form

$$(10)\left\{\begin{aligned}
&\Delta A_{\mu\mu}^{(\alpha_1,\ldots,\alpha_n)}=\tfrac{1}{2}\sum_{\alpha=1}^{3}\psi_{\alpha\alpha}^{(\alpha_1,\ldots,\alpha_n)}-\tfrac{1}{2}\psi_{44}^{(\alpha_1,\ldots,\alpha_n)}\equiv\chi_{\mu\mu}^{(\alpha_1,\ldots,\alpha_n)},\quad(\mu=1,2,3),\\
&\Delta A_{44}^{(\alpha_1,\ldots,\alpha_n)}=-3\frac{\partial^2 A^{(\alpha_1,\ldots,\alpha_n)}}{\partial x_4{}^2}+2\psi_{44}^{(\alpha_1,\ldots,\alpha_n)}\equiv\chi_{44}^{(\alpha_1,\ldots,\alpha_n)},\\
&\Delta A_{\mu 4}^{(\alpha_1,\ldots,\alpha_n)}+\frac{\partial^2 A_{\mu 4}^{(\alpha_1,\ldots,\alpha_n)}}{\partial x_4{}^2}=-\tfrac{3}{2}\frac{\partial^2 A^{(\alpha_1,\ldots,\alpha_n)}}{\partial x_\mu\,\partial x_4}+\tfrac{1}{2}\frac{\partial^2 A_{44}^{(\alpha_1,\ldots,\alpha_n)}}{\partial x_\mu\,\partial x_4}\\
&+2\psi_{\mu 4}^{(\alpha_1,\ldots,\alpha_n)}+\tfrac{1}{2}\int\frac{\partial}{\partial x_4}\left(\frac{\partial^2 A^{(\alpha_1,\ldots,\alpha_n)}}{\partial x_4{}^2}+\Delta A^{(\alpha_1,\ldots,\alpha_n)}-2\psi_{\mu\mu}^{(\alpha_1,\ldots,\alpha_n)}\right)dx_\mu\\
&\qquad\equiv\chi_{\mu 4}^{(\alpha_1,\ldots,\alpha_n)},\quad(\mu=1,2,3),\\
&\frac{\partial^2 A_{12}^{(\alpha_1,\ldots,\alpha_n)}}{\partial x_1\partial x_2}=\frac{\partial^2 A_{34}^{(\alpha_1,\ldots,\alpha_n)}}{\partial x_3\partial x_4}-\tfrac{1}{2}\Delta A^{(\alpha_1,\ldots,\alpha_n)}-\tfrac{1}{2}\frac{\partial^2 A^{(\alpha_1,\ldots,\alpha_n)}}{\partial x_3{}^2}\\
&\qquad-\tfrac{1}{2}\frac{\partial^2 A^{(\alpha_1,\ldots,\alpha_n)}}{\partial x_4{}^2}-\tfrac{1}{2}\frac{\partial^2 A_{44}^{(\alpha_1,\ldots,\alpha_n)}}{\partial x_3{}^2}+\psi_{33}^{(\alpha_1,\ldots,\alpha_n)}\equiv\chi_{12}^{(\alpha_1,\ldots,\alpha_n)},\\
&\frac{\partial^2 A_{13}^{(\alpha_1,\ldots,\alpha_n)}}{\partial x_1\partial x_3}=\frac{\partial^2 A_{24}^{(\alpha_1,\ldots,\alpha_n)}}{\partial x_2\partial x_4}-\tfrac{1}{2}\Delta A^{(\alpha_1,\ldots,\alpha_n)}-\tfrac{1}{2}\frac{\partial^2 A^{(\alpha_1,\ldots,\alpha_n)}}{\partial x_2{}^2}\\
&\qquad-\tfrac{1}{2}\frac{\partial^2 A^{(\alpha_1,\ldots,\alpha_n)}}{\partial x_4{}^2}-\tfrac{1}{2}\frac{\partial^2 A_{44}^{(\alpha_1,\ldots,\alpha_n)}}{\partial x_2{}^2}+\psi_{22}^{(\alpha_1,\ldots,\alpha_n)}\equiv\chi_{13}^{(\alpha_1,\ldots,\alpha_n)},\\
&\frac{\partial^2 A_{23}^{(\alpha_1,\ldots,\alpha_n)}}{\partial x_2\partial x_3}=\frac{\partial^2 A_{14}^{(\alpha_1,\ldots,\alpha_n)}}{\partial x_1\partial x_4}-\tfrac{1}{2}\Delta A^{(\alpha_1,\ldots,\alpha_n)}-\tfrac{1}{2}\frac{\partial^2 A^{(\alpha_1,\ldots,\alpha_n)}}{\partial x_1{}^2}\\
&\qquad-\tfrac{1}{2}\frac{\partial^2 A^{(\alpha_1,\ldots,\alpha_n)}}{\partial x_4{}^2}-\tfrac{1}{2}\frac{\partial^2 A_{44}^{(\alpha_1,\ldots,\alpha_n)}}{\partial x_1{}^2}+\psi_{11}^{(\alpha_1,\ldots,\alpha_n)}\equiv\chi_{23}^{(\alpha_1,\ldots,\alpha_n)}.
\end{aligned}\right.$$

The right side of each of these equations, represented by the notation $\chi_{\mu\nu}^{(\alpha_1,\ldots,\alpha_n)}$, is a function of the $\psi_{\mu\nu}^{(\alpha_1,\ldots,\alpha_n)}$, which are known for each k, and of those of the $A_{\mu\nu}^{(\alpha_1,\ldots,\alpha_n)}$ which are determined by preceding equations. These equations may be solved in the order written, for each set of α_i for each k, beginning with $k=1$.

Solutions of equations of the type of (10) that satisfy the conditions of our problem are well known. Before writing the solutions we return to a real time variable by putting

$$x_4=ict.$$

It will be noticed that the right sides of the fifth, sixth, and seventh equations of (10) now contain the factor $-i/c$. Thus the $A_{\mu 4}^{(a_1,\dots,a_n)}$ $(\mu=1, 2, 3)$ are purely imaginary. The right sides of all of the other equations then become real, and all of the $A_{\mu\nu}^{(a_1,\dots,a_n)}$, except the three mentioned above, are real quantities. When the solutions are substituted in the power series (3), the $g_{\mu 4}$ $(\mu=1, 2, 3)$ are purely imaginary, the other $g_{\mu\nu}$ being real. But when these values are substituted in the quadratic form (1), the coefficients $g_{\mu 4}(\mu=1, 2, 3)$ are multiplied by $dx_\mu dx_4 = icdx_\mu dt$, so that the linear element is entirely real.

Solutions of equations (10) are

$$
(11)\left\{\begin{aligned}
&A_{\mu\mu}^{(a_1,\dots,a_n)} = -\frac{1}{4\pi}\int_D \frac{\chi_{\mu\mu}^{(a_1,\dots,a_n)}(\xi, \eta, \zeta, t)\,d\xi d\eta d\zeta}{r}, \quad (\mu=1,\dots,4),\\
&A_{\mu 4}^{(a_1,\dots,a_n)} = -\frac{1}{4\pi}\int_D \frac{\chi_{\mu 4}^{(a_1,\dots,a_n)}(\xi, \eta, \zeta, t-r/c)\,d\xi d\eta d\zeta}{r}, \quad (\mu=1, 2, 3),\\
&A_{\mu\nu}^{(a_1,\dots,a_n)} = \iint \chi_{\mu\nu}^{(a_1,\dots,a_n)}\,dx_\mu dx_\nu, \qquad (\mu, \nu=1, 2, 3;\ \mu\neq\nu),
\end{aligned}\right.
$$

where

$$r=\sqrt{(x_1-\xi)^2+(x_2-\eta)^2+(x_3-\zeta)^2}.$$

The formulae (11) give a complete solution of our problem, for they determine successively all of the coefficients of the power series (3), thus determining the values of the $g_{\mu\nu}$, which in turn determine the linear element (1).

It remains to show that the solutions (11) satisfy identically the arbitrary relations (9). The two relations of the first line are obviously satisfied. To prove that the third is also satisfied we proceed as follows: The relation in question is

$$\sum_{\alpha=1}^{3} \frac{\partial A_{\alpha 4}^{(a_1,\dots,a_n)}}{\partial x_\alpha} = 0.$$

From the fifth, sixth, and seventh equations of (10) we find that

$$\sum_{\beta=1}^{3} \frac{\partial}{\partial x_\beta} \sum_{\alpha=1}^{4} \frac{\partial^2 A_{\beta 4}^{(a_1,\dots,a_n)}}{\partial x_\alpha^2} = -\sum_{\alpha=1}^{3} \frac{\partial \psi_{\alpha\alpha}^{(a_1,\dots,a_n)}}{\partial x_4} + \frac{\partial \psi_{44}^{(a_1,\dots,a_n)}}{\partial x_4}$$

$$+2\sum_{\alpha=1}^{3} \frac{\partial \psi_{\alpha 4}^{(a_1,\dots,a_n)}}{\partial x_\alpha} = 0.$$

The expression on the right is exactly the fourth of the identities (8), and is therefore identically equal to zero. Interchanging the order of the derivatives we now have

$$\sum_{\beta=1}^{3} \sum_{\alpha=1}^{4} \frac{\partial^2}{\partial x_\alpha{}^2} \frac{\partial A_{\beta 4}^{(a_1,\dots,a_n)}}{\partial x_\beta} = 0.$$

This result shows that the equations (10) are entirely compatible with the arbitrary relation in question. To show that (11) are also compatible with it, we need only apply the same operation by which we passed from the fifth, sixth, and seventh equations of (10) to the corresponding ones of (11). We integrate the equation, then, over the domain D, the time t being retarded to $t-\dfrac{r}{c}$. Performing this operation we get

$$\sum_{\beta=1}^{3} \frac{\partial A_{\beta 4}^{(\alpha_1, \ldots, \alpha_n)}}{\partial x_\beta} = 0.$$

This is precisely the relation that we wished to show was satisfied. Similarly, the proof that the last relation of (9) is satisfied follows at once from the last three equations of (10). If we add these equations, the sum of the right sides reduces exactly to the fourth of the identities (8). We thus obtain precisely the relation we wished to deduce. Our solution therefore satisfies all of the relations (9).

In order to be assured that the solution represented by (11) will satisfy all of the boundary conditions B, and in particular the condition (d), it only remains to choose correctly the particular solutions of the first order, that is, for $k=1$. Select any set, such as $A_{\mu\nu}^{(1,0,\ldots,0)}$. As remarked before, all of the $\psi_{\mu\nu}^{(\alpha_1,\ldots,\alpha_n)}$ for $k=1$ are equal to zero. The first three equations of (11) in this case are

$$\Delta A_{\mu\mu}^{(1,0,\ldots,0)} = 0, \quad (\mu = 1, 2, 3).$$

Using the notation

$$r_i = \sqrt{\sum_{\alpha=1}^{3} (x_\alpha - \phi_\alpha^{(i)})^2}, \quad (i = 1, \ldots, n)$$

the required solution is

$$A_{\mu\mu}^{(1,0,\ldots,0)} = A^{(1,0,\ldots,0)} = -\frac{2}{r_1}.$$

Introducing this value into the fourth equation of (11) it becomes

$$\Delta A_{44}^{(1,0,\ldots,0)} = 6 \frac{\partial^2}{\partial x_4^2} \frac{1}{r_1} \equiv \omega_{44}(x_1, x_2, x_3, x_4).$$

The required solution is

$$A_{44}^{(1,0,\ldots,0)} = \frac{2}{r_1} - \frac{1}{4\pi} \int_D \frac{\omega_{44}(\xi, \eta, \zeta, x_4)\, d\xi d\eta d\zeta}{r}.$$

When these results are substituted in the last six equations of (11), all of the right sides are such that the physical conditions are satisfied. Any other set of $A_{\mu\nu}^{(\alpha_1,\ldots,\alpha_n)}$ for $k=1$, as $A_{\mu\nu}^{(0,\ldots,0,\alpha_\lambda,0,\ldots,0)}$, where $\alpha_\lambda = 1$, is determined at once from the above results by writing r_λ for r_1, and $\phi_\alpha^{(\lambda)}$ for $\phi_\alpha^{(1)}$. A similar symmetry persists for all orders. It is necessary to determine only one set of the $A_{\mu\nu}^{(\alpha_1,\ldots,\alpha_n)}$ for each value of k. The other sets are then deduced by permuting indices.

To calculate the $A_{\mu\nu}^{(\alpha_1, \ldots, \alpha_n)}$ of the second order, $k=2$, we compute first the values of the $\psi_{\mu\nu}^{(\alpha_1, \ldots, \alpha_n)}(k=2)$, as defined by equations (7), in terms of the first order solutions. Then equations (11) give the second order values. This process can be extended indefinitely, giving the values of the $A_{\mu\nu}^{(\alpha_1, \ldots, \alpha_n)}$ for any desired value of k. The substitution of these results successively in (3) and (1) furnishes the complete solution.

The relations (9) were chosen arbitrarily, subject to satisfying a certain set of conditions. Another choice which also satisfies the conditions, and which leads to an interesting but slightly more complicated form of solution is the following:

$$(12)\quad \begin{cases} \dfrac{\partial A_{\mu\mu}^{(\alpha_1, \ldots, \alpha_n)}}{\partial x_4} = \dfrac{\partial A_{\mu 4}^{(\alpha_1, \ldots, \alpha_n)}}{\partial x_\mu}, \quad (\mu=1, 2, 3), \\ A_{11}^{(\alpha_1, \ldots, \alpha_n)} = \text{arbitrary function of } x_1, \ldots, x_4. \end{cases}$$

The equations (7) may then be written in the form

$$(13)\quad \begin{cases} \Delta A_{\mu\mu}^{(\alpha_1, \ldots, \alpha_n)} + \dfrac{\partial^2 A_{\mu\mu}^{(\alpha_1, \ldots, \alpha_n)}}{\partial x_4^2} = X_{\mu\mu}^{(\alpha_1, \ldots, \alpha_n)}, \quad (\mu=2, 3), \\ \Delta A_{44}^{(\alpha_1, \ldots, \alpha_n)} = X_{44}^{(\alpha_1, \ldots, \alpha_n)}, \\ \dfrac{\partial^2 A_{\mu\nu}^{(\alpha_1, \ldots, \alpha_n)}}{\partial x_\mu \partial x_\nu} = X_{\mu\nu}^{(\alpha_1, \ldots, \alpha_n)}, \quad (\mu, \nu=1, 2, 3;\ \mu \neq \nu), \end{cases}$$

where the $X_{\mu\nu}^{(\alpha_1, \ldots, \alpha_n)}$ have properties analagous to those of the $\chi_{\mu\nu}^{(\alpha_1, \ldots \alpha_n)}$ of (10). These equations, with (12), determine in order all of the $A_{\mu\nu}^{(\alpha_1, \ldots, \alpha_n)}$ for each value of k, commencing with $k=1$. The forms of equations involved are the same as those of (10). It is therefore unnecessary to write the form of the solution.

A few words will be added on the subject of the approximate solutions of the field of n spheres, given by Einstein and Droste, to which reference was made above. The field equations in this case have the general form

$$G_{\mu\nu} = -\kappa\,(T_{\mu\nu} - \tfrac{1}{2} g_{\mu\nu} T) = F_{\mu\nu},$$

where $T_{\mu\nu}$ is the tensor of matter. If we put

$$g_{\mu\nu} = -\delta_{\mu\nu} + \gamma_{\mu\nu},$$

and neglect all powers of the $\gamma_{\mu\nu}$ above the first, the expressions for the $G_{\mu\nu}$ in terms of the $\gamma_{\mu\nu}$ become identical with those for the $G_{\mu\nu}^{(\alpha_1, \ldots, \alpha_n)}$ in terms of the $A_{\mu\nu}^{(\alpha_1, \ldots, \alpha_n)}$, given in (7), the $\psi_{\mu\nu}^{(\alpha_1, \ldots, \alpha_n)}$ being considered equal to zero, since they represent the terms of higher order. The complete solution of the equations in this form should then correspond to the terms of the first order in the masses. It will be shown that the results of Einstein and Droste are not applicable to the simpler case of n particles in such a way as to give all of the terms of the first order. Droste neglects at the outset, in fact, certain of the first order terms in the equations, such that his results are numerically accurate only for small velocities of the bodies (relative to that of light). Einstein retains all terms of the first order. He chooses four arbitrary relations among the $\gamma_{\mu\nu}$, corresponding to the relations (9) of this paper. But Einstein has not shown, so far as I am aware, to what approximation the solutions which he thus deduces satisfy these four arbitrary relations. In fact, if we were to choose these same four relations in the problem of n particles, we would obtain solutions satisfied by certain sets of values of the unknown functions, which fail to satisfy the four relations, or the original equations.

In his Princeton Lectures Einstein deduces his solution in a very simple manner. The result is

$$\Delta\gamma_{\mu\nu} + \frac{\partial^2\gamma_{\mu\nu}}{\partial x_4{}^2} = F_{\mu\nu}, \quad (\mu,\ \nu = 1,\ \ldots, 4), \tag{14}$$

and the solution

$$\gamma_{\mu\nu} = \int \frac{F_{\mu\nu}(\xi,\ \eta,\ \zeta,\ t - r/c)\,d\xi d\eta d\zeta}{r}, \quad (\mu,\ \nu = 1,\ \ldots, 4).$$

The four arbitrary relations chosen are

$$\sum_{\alpha=1}^{4} \left[\frac{\partial\gamma_{\mu\alpha}}{\partial x_\alpha} - \frac{1}{2}\frac{\partial\gamma_{\alpha\alpha}}{\partial x_\mu}\right] = 0, \quad (\mu = 1,\ 2,\ 3,\ 4). \tag{15}$$

If we choose these same relations in the case of n particles, the equations corresponding to (14) are

$$\Delta\gamma_{\mu\nu} + \frac{\partial^2\gamma_{\mu\nu}}{\partial x_4{}^2} = 0, \quad (\mu,\ \nu = 1, \ldots, 4). \tag{16}$$

These equations determine a mathematical solution of the problem (neglecting for a moment the physical conditions) only if every particular solution of them satisfies also the original equations; or, what comes to the same thing, the relations (15). For otherwise we must adjoin to (16) the set (15), making a simultaneous set of fourteen equations to consider. But evidently not every particular solution of (16) satisfies (15), for (16) admits, for example, a solution in which all of the $\gamma_{\mu\nu}$ but one are zero, which does not satisfy (15). To obtain a solution of our problem, then, it would be necessary to find out which, if any, of the solutions of (16) satisfy (15), that is, to solve (15) and (16) simultaneously. It is only after this has been accomplished that the question of satisfying the physical conditions arises.

It is probable that in the case of n spheres the solutions of (14) satisfy (15) to within a close approximation, but we cannot obtain from these results all of the terms of the first order in the masses for the field of n moving particles.

THE RED SHIFT OF THE SOLAR LINES AND RELATIVITY

BY DR. CHARLES E. ST. JOHN,
Mount Wilson Observatory, California, U.S.A.

The differences in wave-length between the centre of the sun's disk and the vacuum arc for 331 iron lines are used as the basis of a discussion, from the relativity point of view, of the displacements of the solar lines.

In favour of relativity is the general displacement to the red and against it is the discordance between the magnitudes of the observed displacements and the deductions from the theory. The accompanying table illustrates what has long been recognized, namely, that both the stronger lines and the weaker lines give displacements at the centre of the disk that differ systematically from the calculated values, columns 3 and 7, while for lines of medium intensity the displacements are in substantial agreement with the calculated values. The stronger lines show displacements about 50 per cent. greater and the weaker lines about 30 per cent. smaller than the displacements predicted by the theory of general relativity.

TABLE I

COMPARISON OF DISPLACEMENT (SUN-VACUUM) OBSERVED, AND CALCULATED FROM RELATIVITY THEORY

Group	No. of Lines	Solar Intensity	Mean Wave-length	Calculated	Observed	Obs.-Cal.		
a	17	12	3826	0.008	0.012	+0.004	0.3 km.	Down
b	24	14	3821	0.008	0.0112	+0.0032	0.25 "	"
c	10	10.4	4308	0.0091	0.0113	+0.0022	0.16 "	"
a	10	6.2	5443	0.0115	0.0112	−0.0003		
b	131	4.8	4758	0.0100	0.0084	−0.0016	0.1 "	Up
d	106	4.5	4763	0.0100	0.0069	−0.0031	0.2 "	"
a	33	3.3	4957	0.0105	0.0074	−0.0031	0.2 "	"

This discrepancy between observed and calculated displacements has long stood in the way of an interpretation based upon general relativity. It is the purpose of the present discussion to suggest a harmonizing interpretation.

The key to the solution offered is that the deviations from theory shown by the strong and weak lines are to be correlated not with line-intensity but with the levels in the sun's atmosphere at which the lines are produced. This is clearly evident from the behaviour of lines of enhanced and normal titanium. Correlation with levels is shown in the accompanying tabulation where lines at heights in the flash spectra of 6000, 1300 and 435 km. show respectively absolute displacements of +0.015, +0.0112 and 0.0054A and deviations of +0.007, +0.002 and −0.0034A, from the displacements calculated from relativity theory. That the deviations are not associated with line-intensity is shown by comparing the displacements and deviations for enhanced and normal lines of approximately equal intensities, which show respectively absolute displacements of +0.0112 and +0.0054A and deviations of +0.002 and −0.0054 from the calculated values.

RED DISPLACEMENT AND LEVEL

	No.	Mean Int.	Sun-Vac.	Obs.-Cal.	Height
Enh. *Ti*...........	2	11	+0.015A	+0.007A	6.000 km.
Enh. *Ti*...........	8	4.6	+0.0112	+0.002	1.300 "
Normal *Ti*.........	5	4.2	+0.0054	−0.0034	435 "

The low pressure in the reversing layer removes at one stroke consideration of pressure shift and of ray-curving due to anomalous refraction. The Zeeman effect of the general magnetic field tends only to produce a slight but symmetrical widening of the lines. The Stark effect has not been found in the sun although looked for under conditions apparently favourable to its appearance.

There remain to be considered Doppler and general relativity, neither of which alone is capable of producing red displacements of the solar lines in agreement with the observations. These two in combination, however, offer the simplest interpretation of the displacements at the centre of the disk. In column 8 of Table I are shown, reading from top to bottom, the magnitudes and the directions of the radial movements in the solar vapours that would account for the deviations from the Einstein Theory for high and low-level lines.

These velocities, upward near the photosphere and downward at high elevations, appear to result from spectrographic integrations over very extended areas of the solar surface, millions of square miles. If currents are upward over the bright hot granules and downward over the larger and cooler interspaces the integrated Doppler effect would be a slight widening of the spectral line greater on the violet than on the red edge. The decrease of upward velocity with elevation brings about a balanced state at the level of lines of medium intensity. Above the level of equilibrium the cooler downward drifting vapours become of increasing influence and the integrated effect is an asymmetry on the red edge increasing with increase of elevation.

The low pressure and the greater displacement to the red of high-level lines in the sun are confirmed by the low pressure found in Sirius, Procyon and Arcturus and by the larger red displacements of high-level in comparison with the displacements of low-level lines in the stellar atmospheres. The velocities of the solar vapours required to account for the deviations from the predictions of relativity are of an order of magnitude consistent with the stellar observations.

At the sun's limb where the radial components of velocity vanish, the red displacements of all lines exceed the relativity displacements by small amounts. This excess, the real limb effect from the relativity point of view, may be interpreted as the effect of molecular scattering in accordance with the Rayleigh-

Schuster formulae. Professor Julius has called attention to the fact that in general the refractive power is greater on the red than on the violet side of an absorption line by twice the refractive power of the solar atmosphere. As the coefficient of scattering increases as the square of the refractive power, the differential scattering tends to widen the lines on the red edge. The short paths through layers of low density at the centre of the disk would account for the absence of differential widening at the centre while the greatly lengthened paths at the limb would furnish conditions favourable to the differential effect.

The conclusion is that three major causes are producing the regular differences between solar and terrestrial wave-lengths and that it is possible to disentangle their effects. The causes appear to be the slowing up of the atomic clock in the sun to an amount predicted by the theory of generalized relativity, radial velocities of moderate cosmic magnitude and in probable directions, and differential scattering in the longer paths traversed through the solar atmosphere by light coming from the limb of the sun. The first obtains for all lines in all parts of the sun, the second appears regularly and continuously, downward at very high and upward at the very low levels, while the third manifests itself in the so-called limb-effect.

A GENERALIZATION OF ELECTRODYNAMICS, CONSISTENT WITH RESTRICTED RELATIVITY AND AFFORDING A POSSIBLE EXPLANATION OF THE EARTH'S MAGNETIC AND GRAVITATIONAL FIELDS, AND THE MAINTENANCE OF THE EARTH'S CHARGE*

By Professor W. F. G. Swann†,

Director of the Sloane Laboratory, Yale University, New Haven, Connecticut, U.S.A.

Introduction

Any theory of the earth's magnetic field which invokes the rotation of the earth as a cause will present that field as a function of a, the radius of the sphere and ω, the angular velocity. A sphere with a current density i of the form $i = AD\omega^m r^n$, where r is the distance from the axis of rotation, and D is the density of the sphere, gives rise to a magnetic field H_z at the pole given by‡

$$H_z = AD\omega^m a^{n+1} F(n) \tag{1}$$

where $F(n)$ is a function of n only.

A large class of theories lead to the equivalent of a current density proportional to $D\omega r$, and so give rise at the surface to a field proportional to $D\omega a^2$. This form of expression has the advantage that on its basis, assuming the magnetic field of the earth to be 0.5 gauss at the earth's pole, we should calculate the known value 50 gauss at the corresponding place on the sun. Moreover, the form of the expression is such as to result in the conclusion that it would be impossible to give to small spheres of laboratory size such angular velocities as would correspond to measurable magnetic fields. These two tests, prediction of the correct ratio of the magnetic fields of the earth and sun, and prediction of a magnetic field of magnitude not, at least, easily measurable in the case of a small sphere set into rotation with high speed, constitute conditions which any theory of terrestrial magnetism must satisfy. Some possibilities which suggest themselves as theories lead to fields at the

*Presented at the International Mathematical Congress, Toronto, August, 1924. Also at the Washington Meeting of the American Physical Society, December 29-31, 1924. A digest of the paper is also given in the writer's article *The Origin of the Earth's Electric and Magnetic Phenomena*, Jour. Franklin Inst., Vol. 201, pp. 143-176, 1926, and more recently it has been published in full in Phil. Mag., Vol. 3, pp. 1088-1136, 1927.

†Now Director of the Bartol Research Foundation of the Franklin Institute.

‡See Appendix, Note 1. We shall relegate to an appendix pieces of analysis and subsidiary discussions which, if included in the main part of the article, would interfere with the continuity of thought.

surface proportional to $D\omega^3 a^2$. This is the case, for example, with those in which the centrifugal force* of the earth's rotation is involved as the primary agency responsible for bringing about conditions which give rise to the field. Quite apart from the fact that such theories, when worked out on any reasonable assumptions give fields insignificant in absolute values, they would predict for the sun a field much smaller than that of the earth; and for a copper sphere of radius 10 centimetres, rotating at 100 revolutions per second, they would predict for the field at the surface a value over a hundred thousand times that of the earth. This fact is sufficient to rule such theories out of court.

Although a theory giving rise to a real or apparent current density of the form $D\omega r$ satisfies the conditions well as regards comparison with the field of the sun and that of a small body rotating at high speed, it will turn out that such a theory is unsatisfactory for other reasons. In the first place it makes the effective current density at an element of matter proportional to its linear velocity, and otherwise independent of its angular velocity, so that such theories would presumably lead, or at any rate suggest, in general, magnetic fields as a result of the translation of matter with uniform rectilinear velocity, a consequence which would lead to undesirable conclusions as regards magnetic fields due to the translatory motion of the planets, and which has, moreover, been demonstrated as non-existent by the experiments of H. A. Wilson†. Moreover, as we shall see later, if we attempt to correlate in a rational way, under a scheme of modified electrodynamics, a situation in which neutral matter can by its motion give rise to a magnetic field according to the same law as that which in classical theory governs the magnetic field due to moving electricity, we shall evolve the unwelcome conclusion that what we desire to regard as neutral matter will also give rise to an electric field. Even though our definitions may have been formulated in such a way that the matter may be regarded technically as neutral, the analysis will refuse to be deceived by our devices, and in giving us our magnetic field due to the rotation of neutral matter it will make that "neutral matter" present us with an electric field as large as the field we should have obtained in classical electrodynamics, if we had imagined the magnetic field produced by the rotation of a charged earth.

If we could provide for an apparent current density proportional to $D\omega^4 r^3$, the magnetic field at the surface of the sphere would come out proportional to $D\omega^4 a^4$, as may be seen from (1); and, within the limits of the experimental data, an expression of this form would serve almost as well as one of the form $D\omega a^2$ cited above, for the purpose of giving the correct ratio of the magnetic fields o the earth and sun‡. For a sphere of 10 cm. radius, and with the earth's density, rotating 100 times per second, it would give at the surface on the pole a magnetic intensity 3.4×10^{-4} of the corresponding intensity for a sphere of the earth's size rotating with the earth's angular velocity. A field amounting to 3.4×10^{-4}

*See W. F. G. Swann: *On the Magnetic Electric Fields which Spontaneously arise in the Vicinity of a Rotating Conducting Sphere*, Terr. Mag., Vol. 22, p. 154, 1917.

†See H. A. Wilson: Proc. Royal Soc., Vol. 104, pp. 451-455, 1923.

‡In fact, the ratio of $\omega^3 a^2$ for the sun to that for the earth is equal to 1/1.5, and is thus not very far different from unity.

of the earth's intensity is not inconceivably beyond the limits of experimental detection, but the difficulty of detecting it would be very great when the high angular velocity associated with its production is taken into account.

Of all possible expressions of a form proportional to D and to integral powers of ω and r for the apparent current density at a distance r from the axis of rotation, that proportional to $D\omega^4r^3$ is the only one, other than the expression* $D\omega r$, which would be in harmony with the relative magnetic fields of the sun and earth and with the smallness of the magnetic field to be obtained by the rotation of a small body at high speed. For comparison of the fields of the sun and earth would necessitate that in any additional factor $\omega^m r^n$ which we might attach to $D\omega^4r^3$, m would be greater than n, while the avoidance of production of an unallowably large magnetic field for a small body rotating at high speed of, say, 1000 revolutions per second would necessitate $m \leqslant n$.

In Table I are collected a few figures illustrating the foregoing remarks and giving, for various types of law as to the variation of apparent current density with r and ω, the relative values of the magnetic intensity H_z at the pole for the cases of the sun, the earth, and a sphere of the earth's density, but of radius 10 cm., rotating at 100 revolutions per second. The relative values of a, ω, and ω^3a^2 are also given. The significance attaching to the quantity J will appear later.

TABLE I

	Relative Values			Remarks
	S ○	E ○	100 rev./sec. radius 10 cm. ○	
a	7.0×10^9	6.4×10^7	1	
ω	1	25.5	2.2×10^8	
ω^3a^2	1	1.4	2.2×10^5	
H_z	58	0.5	1.0×10^{-9}	$i \propto Dr\omega$; $H_z \propto D\omega a^2$
H_z	0.1	0.5	7.8×10^4	$i \propto Dr\omega^2$; $H_z \propto D\omega^3a^2$
H_z	42	0.5	1.7×10^{-4}	$i \propto Dr^3\omega^4$; $H_z \propto D\omega^4a^4$
J	-15×10^{14}	-3×10^{12}	-850	$J \propto D\omega^2a^3$
J	-4×10^{16}	-3×10^{12}	1.0×10^{-4}	$J \propto D\omega a^3$

In the theory which will subsequently be developed, we shall consequently provide for the equivalent of a current density of the form ω^3r^2. This will have the advantage of making the current density depend on ω explicitly as well as on the linear velocity ωr, and will be of the type to give zero result for non-rotational motion. Of course it must be understood that, when the term giving rise to this apparent current density finally appears in the theory and equations to be developed, it will come in in a more general form as applicable to all kinds of motion, and in such a way as to make it part of a consistent theory of electrodynamics. All that we desire to state here is our intention to

*The objections to which have already been discussed.

adjust the general theory in such a way as to reduce to a condition of this type for the special case of uniform rotation.

It must be pointed out that all theories of the kind under consideration lead in their simplest aspects to a state of magnetization symmetrical about the axis of rotation. It is felt that, once a satisfactory theory providing for this condition is developed, the difficulties inherent in accounting for the non-coincidence of the magnetic and geographic axes, and the departure of the field from that of the types following from an effective current density which is a function only of r, might well be of secondary seriousness.

As regards the maintenance of the earth's charge, the facts are these. The earth is negatively charged to an extent such as to give rise to a potential-gradient of about 150 volts per meter. That charge is streaming out from the earth on account of the atmospheric conductivity at such a rate that 90 per cent. of it would be gone in 10 minutes if there were no means of replenishing the loss. The difficulties in explaining the replenishment in a manner such as to maintain consistency with all the facts are very great*; and, while theories have been proposed, all have their objections. So serious is the situation that Dr. G. C. Simpson† has been moved to raise the question of whether or not charge may be spontaneously generated in the earth's interior. Such a view must of course be moulded with the greatest care into any logical theory, as one is here striking at the very foundations of one's fundamental concepts of electric charge, and to tamper with the structure here may throw complete confusion into the most commonplace electrodynamic conclusions elsewhere. Nevertheless, in the theory about to be developed, the endeavour has been made to invoke the spirit of this idea of creation (or annihilation) of charge as a result of the earth's motion, but to do so in a manner in which its effect on the whole of the remainder of the electrodynamic scheme is perfectly precise and logically understandable. Speaking briefly, the effect becomes secured by making the modified electrodynamic equations of a type in which the equation of continuity no longer holds universally, at least for positive electricity. Under these conditions the rotation of the earth is made to give rise to a slow death of positive electricity, so that the earth charges negatively to such an extent that a steady state is finally attained when the conduction through the atmosphere into space balances the death of positive electricity. It turns out that it is only necessary to provide for a rate of death of positive electricity corresponding to a disappearance of half of one per cent. of the earth's mass in 10^{20} years, a death corresponding to the loss of one proton per c.c. per day.

Here we have no knowledge of the electric field of another planet to guide us in determining the way in which the field is to depend on the angular velocity and radius. We are, however, able to restrict our choice by the fact that the theory must not lead to the conclusion that the generation of charge by a small body rotating at high speed in the laboratory shall be easily measurable.

*See, for example, W. F. G. Swann, *Unsolved Problems of Cosmical Physics*, Jour. Franklin Inst., Vol. 195, pp. 433-474, 1923.

†G. C. Simpson, *Some Problems of Atmospheric Electricity:* Monthly Weather Review, Vol. 44, p. 121, 1916.

The modification introduced in the electromagnetic scheme to account for gravity is almost precisely that of the theory of Lorentz, in which the basic idea is the assumption of a difference between the force of attraction between unlike charges and the force of repulsion between corresponding like charges. The only reason for including this theory in the present paper is the desire to mould the gravitational theory into harmony with the others, in such a manner as to result in one complete scheme for the whole, a scheme, moreover, of a form invariant under the transformation of the restricted theory of relativity.

Generalization of the Electrodynamic Equations

Remarks on the Classical Scheme

The classical circuital relations of electrodynamics are, in the usual notation, and in Heavisidean units,

$$\frac{\rho \mathbf{u}}{c} = -\frac{1}{c}\frac{\partial \mathbf{E}}{\partial t} + \text{curl}\,\mathbf{H}, \tag{2}$$

$$\rho = \text{div}\,\mathbf{E}, \tag{3}$$

$$0 = \frac{1}{c}\frac{\partial \mathbf{H}}{\partial t} + \text{curl}\,\mathbf{E}, \tag{4}$$

$$0 = \text{div}\,\mathbf{H}. \tag{5}$$

These equations give the fields **E** and **H** when ρ and **u** are assigned, but they do not determine the motion of the electricity itself. This is provided for by a fifth relation not derivable from (2)-(5). This equation may be written

$$\left(\mathbf{E} + \frac{[\mathbf{wH}]}{c}\right) e = \frac{d}{dt}\frac{m_0 \mathbf{w}}{\left(1 - \frac{q^2}{c^2}\right)^{\frac{1}{2}}} + \mathbf{B}, \tag{6}$$

where **E**, **H**, is the field at some specified point inside the electron, neglecting the field of the electron itself, and where e is the electronic charge, **w** the electron's vector velocity, and q its resultant velocity, m_0 the rest mass, *viz.*, $e^2/6\pi ac^2$, with a the electronic radius, and where **B** is the radiation reaction.

Precision of statement demands that we say something more about this equation; but as the points involved have no direct bearing upon the generalization of the whole set of equations which we intend to discuss, and have to do primarily with the classical theory as such, we shall relegate remarks in this connection to the Appendix, note 2.

Now it is desirable that we note several points with regard to equations (2) to (6). These are:

(a) Equations (2)-(5) are capable of being solved, and give **E** and **H** in the form

$$\mathbf{E} = -\frac{1}{c}\frac{\partial \mathbf{U}}{\partial t} - \text{grad}\,\phi, \tag{7}$$

$$\mathbf{H} = \text{curl}\,\mathbf{U}, \tag{8}$$

where

(9) $$4\pi\phi = \iiint \frac{[\rho]}{R} d\tau,$$

(10) $$4\pi c\mathbf{U} = \iiint \frac{[\rho\mathbf{u}]}{R} d\tau,$$

ϕ and $\mathbf{U}$ being restricted by the relation

(11) $$\text{div } \mathbf{U} = -\frac{1}{c}\frac{\partial\phi}{\partial t},$$

the square brackets here being used to denote that the potentials are calculated in the usual retarded sense. It thus results that each element of electricity produces its own contribution to the field at a point, these contributions being additive to obtain the resultant field. Thus, if we should so wish, we could speak of the field $\mathbf{E}_+, \mathbf{H}_+$ produced by positive electricity and the field $\mathbf{E}_-, \mathbf{H}_-$, produced by negative electricity.

(*b*) Equations (2) and (3) contain, as a direct algebraical consequence, the equation

(12) $$-\frac{\partial\rho}{\partial t} = \text{div } \rho\mathbf{u}.$$

This is the equation of continuity, and provides for the conservation of electric charge. It may not be inappropriate to recall that the equation of continuity is not something which can be provided for other than through the equations. It is not true for everything. It is not true of the people in this country that the rate of decrease of population is equal to the difference between the rates at which people go out of and come into the country, because some are born here and some die here. Equation (12) denies the possibility of electricity being born or dying in the volume elements, and (11) which is implied by it (see Appendix, note 4), serves to restrict the values of ρ and $\rho\mathbf{u}$ assigned in our problems to those in harmony with it.

(*c*) An analysis of the steps by which (7)-(11) are derived from equations (2)-(5), will show that ρ and $\rho\mathbf{u}$ occur as they do in the solutions for $\mathbf{E}$ and $\mathbf{H}$ *not* because they happen to be density and density times velocity, but because they constitute the left hand sides of equations (2) and (3). In other words, consider a set of equations

(13) $$\frac{\mathbf{S}}{c} = -\frac{1}{c}\frac{\partial\mathbf{E}}{\partial t} + \text{curl } \mathbf{H},$$

(14) $$S_t = \text{div } \mathbf{E},$$

(15) $$0 = \frac{1}{c}\frac{\partial\mathbf{H}}{\partial t} + \text{curl } \mathbf{E},$$

(16) $$0 = \text{div } \mathbf{H},$$

where $\mathbf{S}$ is any vector with components S_x, S_y, S_z, and S_t is a scalar satisfying*

$$-\frac{\partial S_t}{\partial t}=\operatorname{div}\mathbf{S}. \tag{17}$$

The solution of these equations is

$$\mathbf{E}=-\frac{1}{c}\frac{\partial \mathbf{U}}{\partial t}-\operatorname{grad}\phi, \tag{18}$$

$$\mathbf{H}=\operatorname{curl}\mathbf{U}, \tag{19}$$

where

$$4\pi\phi=\iiint\frac{[S_t]}{R}d\tau,\quad 4\pi c\mathbf{U}=\iiint\frac{[\mathbf{S}]}{R}d\tau, \tag{20}$$

with ϕ and $\mathbf{U}$ subject to the condition

$$\operatorname{div}\mathbf{U}=-\frac{1}{c}\frac{\partial\phi}{\partial t}. \tag{21}$$

This fact is proved formally in Appendix, note 3.

Equation (17) is an analytical consequence of (13) and (14), and, whatever $\mathbf{E}$ and $\mathbf{H}$ may be, it is a condition automatically imposed on S_x, S_y, S_z, S_t by the mere fact of their being equated to the right hand sides of (13) and (14). It is the representative of equation (12) in the classical equations, and equation (21) has merely the function of preventing us from assigning, in any problem, values of $\mathbf{S}$ and S_t which would themselves violate (17), just as (11) prevents us from assigning values of ρ and $\mathbf{u}$ which would themselves violate the equation of continuity (12), an equation which forms part of the basis of solution of the problem (see Appendix, note 4). Of course (17) is no longer the equation of continuity except for the special case $\mathbf{S}=\rho\mathbf{u}$, $S_t=\rho$.

(*d*) With a proper understanding of the meanings of $\mathbf{E}$ and $\mathbf{H}$, an understanding provided, for example, in the definition of these quantities which follows from limiting cases of (6), the whole set of equations (2)-(6) is invariant (see note 5) under the transformation of the restricted theory of relativity, viz.,

$$x'=\epsilon(x-vt),\ y'=y,\ z'=z,\ t'=\epsilon\left(t-\frac{v}{c^2}x\right), \tag{22}$$

where

$$\epsilon=\left(1-\frac{v^2}{c^2}\right)^{-\frac{1}{2}}.$$

The invariance of (2)-(5) follows from the following facts. In the sense of their proper definitions H_x, H_y, H_z, $-iE_x$, $-iE_y$, $-iE_z$, constitute a six vector (i being $\sqrt{-1}$), so that the four quantities formed by the three components of the right hand side of (2) and the right hand side of (3), with an additional factor i, constitute a 4-vector. Moreover, the three components of $\rho\mathbf{u}/c$ together with the quantity ρ (with an i attached) form the four components of a 4-vector.

*The subscript t is used merely to call to mind the association of the fourth component with the t axis in the sense understood in the restricted theory of relativity.

Again, the three components of the right hand side of (4) and the right hand side of (5) with a factor i, constitute the four components of another 4-vector; while, as regards (6), after multiplying throughout by $(1-q^2/c^2)^{-\frac{1}{2}}$, the left hand side corresponds to three of the four components of a 4-vector, and the right hand side, if properly chosen as regards **B**, corresponds also to three of the four components of a 4-vector in the same direction.

Thus, in order to maintain invariance under the restricted theory of relativity in the case of such a set of equations as (13)-(16), all that is necessary is to provide, as corresponding to (6), a fifth equation which will lead to definitions of **E** and **H** such as to make **H**, $-i\mathbf{E}$ a six vector, and to choose the quantities S_x, S_y, S_z, icS_t so that they form the 4 components of a 4-vector satisfying the relation (17).

The Generalization

In proceeding to a generalization of the equations in line with the possibilities pointed out in (*c*) and (*d*), we first write down two sets of circuital relations of the classical type, one for the field due to positive electricity, and the other for the field due to negative electricity. It is the former set which we shall generalize, leaving the latter in its old form. We shall denote by subscripts + or − fields produced by or quantities associated with positive or negative electricity respectively.

For the positive electricity we shall generalize equations (2)-(5) by replacing $\rho\mathbf{u}$, ρ by $\mathbf{S}_+$, $(S_t)_+$, as in (*c*), where

$$\mathbf{S}_+ = \rho_+\mathbf{u}_+ + \alpha\mathbf{P} + \beta\mathbf{Q}, \tag{23}$$

$$(S_t)_+ = \rho_+ + \alpha P_t + \beta Q_t, \tag{24}$$

where P_x, P_y, P_z, icP_t constitute components of a 4-vector, Q_x, Q_y, Q_z, icQ_t, constitute the components of another 4-vector, $(\rho_+\mathbf{u}_+)_x$, $(\rho_+\mathbf{u}_+)_y$, $(\rho_+\mathbf{u}_+)_z$, $ic\rho_+$ constitute the components of another 4-vector, and α and β are invariants of the Lorentzian transformation.

Our complete set of circuital relations then becomes:

$$\frac{\mathbf{S}_+}{c} = -\frac{1}{c}\frac{\partial \mathbf{E}_+}{\partial t} + \text{curl } \mathbf{H}_+, \tag{25}$$

$$(S_t)_+ = \text{div } \mathbf{E}_+, \tag{26}$$

$$0 = \frac{1}{c}\frac{\partial \mathbf{H}_+}{\partial t} + \text{curl } \mathbf{E}_+, \tag{27}$$

$$0 = \text{div } \mathbf{H}_+, \tag{28}$$

$$\frac{(\rho_-\mathbf{u}_-)}{c} = -\frac{1}{c}\frac{\partial \mathbf{E}_-}{\partial t} + \text{curl } \mathbf{H}_-, \tag{29}$$

$$\rho_- = \text{div } \mathbf{E}_-, \tag{30}$$

$$0 = \frac{1}{c}\frac{\partial \mathbf{H}_-}{\partial t} + \text{curl } \mathbf{E}_-, \tag{31}$$

$$0 = \text{div } \mathbf{H}_-, \tag{32}$$

where $\mathbf{S}_+$, $(S_t)_+$ are given by (23) and (24).

It is, however, necessary to add to these a force equation, and an equation, moreover, which will provide for the six vector characteristics of the vectors $\mathbf{H}_+$, $-i\mathbf{E}_+$ and $\mathbf{H}_-$, $-i\mathbf{E}_-$. As a matter of fact, we shall write down two force equations, one for the motion of positive and one for the motion of negative electrons. These we shall write as follows:

$$\frac{m_+}{e_+}\frac{d}{dt}(k_+\mathbf{w}_+) = \mathbf{E}_+ + \frac{[\mathbf{w}_+\mathbf{H}_+]}{c} + \lambda\left(\mathbf{E}_- + \frac{[\mathbf{w}_+\mathbf{H}_-]}{c}\right), \tag{33}$$

$$\frac{m_-}{e_-}\frac{d}{dt}(k_-\mathbf{w}_-) = \mathbf{E}_- + \frac{[\mathbf{w}_-\mathbf{H}_-]}{c} + \lambda\left(\mathbf{E}_+ + \frac{[\mathbf{w}_-\mathbf{H}_+]}{c}\right), \tag{34}$$

where

$$k_+ \equiv \left(1 - \frac{q_+^2}{c^2}\right)^{-\frac{1}{2}}; \quad k_- \equiv \left(1 - \frac{q_-^2}{c^2}\right)^{-\frac{1}{2}}, \tag{35}$$

q_+ and q_- being the absolute velocities and $\mathbf{w}_+$ and $\mathbf{w}_-$ the vector velocities of the positive and negative electrons respectively. The factor λ is a constant very nearly equal to unity; and, in its difference from unity, it provides for a difference in the motional effect of unit field on an electron according as that field is produced by positive or negative electricity. It is to the factor λ that we shall look later for the explanation of gravitation.

Now, by taking the scalar product of $\mathbf{w}_+$ and each side of (33), we readily derive

$$\frac{m_+}{e_+}\left(\mathbf{w}_+ \cdot \frac{d}{dt}k_+\mathbf{w}_+\right) = (\mathbf{E}_+ \cdot \mathbf{w}_+) + \lambda(\mathbf{E}_- \cdot \mathbf{w}_+), \tag{36}$$

and in an analogous way from (34) we derive

$$\frac{m_-}{e_-}\left(\mathbf{w}_- \cdot k\frac{d}{dt}k_-\mathbf{w}_-\right) = (\mathbf{E}_- \cdot \mathbf{w}_-) + \lambda(\mathbf{E}_+ \cdot \mathbf{w}_-). \tag{37}$$

The expressions

$$k_+\frac{d}{dt}(k_+\mathbf{w}_+), \quad \frac{k_+ i}{c}\left(\mathbf{w}_+ \cdot \frac{d}{dt}k_+\mathbf{w}_+\right)$$

constitute a 4-vector, as do also the corresponding quantities with negative subscripts. It is shown in Appendix (note 5) that, with the definitions of $\mathbf{E}_+$, $\mathbf{E}_-$- $\mathbf{H}_+$, $\mathbf{H}_-$, appropriate to the solution, $\mathbf{H}_+$, $-i\mathbf{E}_+$, and $\mathbf{H}_-$, $-i\mathbf{E}_-$ constitute six vectors so that the whole set of equations (25)-(37) is invariant.

A further remark must be added to the effect that equations (33) and (34), although generalized from the classical force equation in one sense, appear to have lost generality in another sense, in that the kinetic reaction represented by the left hand side contains no terms involving higher time derivatives of the velocity than the first. This restriction is by no means essential, however, for, remembering that we may from any 4-vector obtain another by operating with

$k_+ \dfrac{d}{dt}$, it is obvious that we could generalize the left hand side of (33), for example, to

$$\frac{1}{k_+}\left\{ ak_+ \frac{d}{dt} k_+ \frac{d}{dt}(k_+\mathbf{w}_+) + bk_+ \frac{d}{dt} k_+ \frac{d}{dt} k_+ \frac{d}{dt}(k_+\mathbf{w}_+) + \dots \text{ etc.}\right\} \tag{38}$$

where a, b, ..., etc., are invariants. We could even insert invariants under the operators $\dfrac{d}{dt}$ without destroying the 4-vector character of the expression. Further, as far as relativity is concerned, the invariants are arbitrary except for the restriction that if $\mathbf{W}$ represents the first three of the components of the 4-vector the fourth must be equal to $k_+ i(\mathbf{W}.\mathbf{w})/c$, as pointed out in Appendix (note 6). Analogous remarks apply, of course, to equation (34).

This matter concerned with generalizations of the left-hand side of (33) and (34) plays, of course, no part in the applications to terrestrial magnetism, terrestrial electricity, and gravitation discussed in this paper; and it is introduced merely as part of the discussion of the whole scheme of generalized electrodynamics.

In a rather complex situation of the kind here involved, it is desirable that one pay proper attention to the definitions of the quantities he uses, such quantities as m, e, etc., We shall not introduce an interruption by discussing these matters here, but shall relegate them to the Appendix (note 7).

We must now return to the equations (23) and (24) and the circuital relations (25)-(32). We understand these relations as applying generally for any kind of motion, but it will be our purpose to adjust $\mathbf{P}$, P_t, $\mathbf{Q}$, Q_t, α and β, so that, for uniform rotation they give us the results we hope for in that case. Our equations for $\mathbf{E}$ and $\mathbf{H}$ will then be as follows:

$$\mathbf{E}_+ = -\text{grad}\,\phi_+ - \frac{1}{c}\frac{\partial \mathbf{U}_+}{\partial t}, \tag{39}$$

$$\mathbf{H}_+ = \text{curl}\,\mathbf{U}_+, \tag{40}$$

$$\mathbf{E}_- = -\text{grad}\,\phi_- - \frac{1}{c}\frac{\partial \mathbf{U}_-}{\partial t}, \tag{41}$$

$$\mathbf{H}_- = \text{curl}\,\mathbf{U}_-, \tag{42}$$

where

$$4\pi\phi_+ = \iiint \frac{[\rho_+ + \alpha P_t + \beta Q_t]}{R}\,d\tau, \tag{43}$$

$$4\pi c\mathbf{U}_+ = \iiint \frac{[\rho_+\mathbf{u}_+ + \alpha\mathbf{P} + \beta\mathbf{Q}]}{R}\,d\tau, \tag{44}$$

where

$$4\pi\phi_- = \iiint \frac{[\rho_-]}{R}\,d\tau, \tag{45}$$

$$4\pi c\mathbf{U}_- = \iiint \frac{[\rho_-\mathbf{u}_-]}{R}\,d\tau, \tag{46}$$

with the understanding that, on account of (17),

(47) $$-\frac{\partial\rho_+}{\partial t}-\frac{\partial}{\partial t}(\alpha P_t)-\frac{\partial}{\partial t}(\beta Q_t)=\text{div }\rho_+\mathbf{u}_+ +\text{div }(\alpha\mathbf{P})+\text{div }(\beta\mathbf{Q}),$$

just as, on account of the analogous equation for the unmodified scheme for negative charges,

(48) $$-\frac{\partial\rho_-}{\partial t}=\text{div }\rho_-\mathbf{u}_-.$$

As already remarked (see Appendix, note 4), (47) and (48) are the equivalent of the restricting relations:

(49) $$\text{div }\mathbf{U}_+ = -\frac{1}{c}\frac{\partial\phi_+}{\partial t},$$

(50) $$\text{div }\mathbf{U}_- = -\frac{1}{c}\frac{\partial\phi_-}{\partial t}.$$

Now (47) is no longer the equation of continuity. In other words, it is no longer true that $-\partial\rho_+/\partial t$ is represented by the net rate at which positive electricity goes *out* of the unit of volume. The equation provides for a true birth or death of positive electricity on account of the additional terms in it. The rate of death of charge density is $\frac{-\partial\rho_+}{\partial t}$ where, since div $\rho_+\mathbf{u}_+$ is zero for uniform rotation of a homogeneous sphere

(51) $$-\frac{\partial\rho_+}{\partial t}=\text{div }(\alpha\mathbf{P})+\text{div }(\beta\mathbf{Q})+\frac{\partial}{\partial t}(\alpha P_t)+\frac{\partial}{\partial t}(\beta Q_t),$$

an expression which we shall have occasion to use presently.

Again if for a piece of neutral matter, $\rho_+=-\rho_-$, and if $\mathbf{u}_+=\mathbf{u}_-$, we obtain for this case by addition from (43) and (45), and by addition from (44) and (46),

(52) $$4\pi(\phi_+ +\phi_-)=\iiint\frac{[\alpha P_t+\beta Q_t]}{R}\,d\tau,$$

(53) $$4\pi c(\mathbf{U}_+ +\mathbf{U}_-)=\iiint\frac{[\alpha\mathbf{P}+\beta\mathbf{Q}]}{R}\,d\tau,$$

and the values of $\mathbf{E}_+ +\mathbf{E}_-$ and $\mathbf{H}_+ +\mathbf{H}_-$ may be obtained directly from these by the aid of (39)-(42). In fact

(54) $$\mathbf{E}_+ +\mathbf{E}_- = -\text{grad }(\phi_+ +\phi_-)-\frac{1}{c}\frac{\partial}{\partial t}(\mathbf{U}_+ +\mathbf{U}_-);\quad \mathbf{H}_+ +\mathbf{H}_- =\text{curl }(\mathbf{U}_+ +\mathbf{U}_-).$$

It is of course true that these quantities are not strictly to be regarded as the electric and magnetic fields which would figure in a practical measurement for the cases under consideration. They differ from these on account of the difference between unity and the λ of equations (33)-(34); and, in view of the fact that the whole of the enormous amounts of positive and negative electricity in the earth participate separately in the determination if ϕ_+, $\mathbf{U}_+$ and ϕ_-, $\mathbf{U}_-$ it is not obvious that the small difference between λ and unity is of negligible account. It turns out that the magnitude of the gravitational constant is such as to justify the approximation, however, as is shown in Appendix (note 8), so that while it is not necessary to make any approximation at all, we shall do so in order to avoid complication of the expressions to the obscuration of the main ideas.

Choice of the quantities $\mathbf{P}$, P_t; $\mathbf{Q}$, Q_t. Returning to equations (23) and (24), where $\mathbf{P}$, P_t; $\mathbf{Q}$, Q_t first made their appearance as generalization on the terms $\rho_+\mathbf{u}_+$, ρ_+, and remembering that by operating successively with $k_+\dfrac{d}{dt}$ on the 4-vector $\rho_+\mathbf{u}_+$, ρ_+ic, we obtain a set of 4-vectors of successively increasing orders, it seems natural to take the first two of these as corresponding to the 4 vectors $\mathbf{P}$, icP_t; $\mathbf{Q}$, icQ_t respectively, so that, on this basis, we have

$$\mathbf{P}=k_+\frac{d}{dt}(\rho_+\mathbf{u}_+)=k_+\rho_+\dot{\mathbf{u}}+k_+\mathbf{u}_+\frac{d\rho_+}{dt}, \tag{55}$$

$$P_t=k_+\frac{d\rho_+}{dt}, \tag{56}$$

$$\mathbf{Q}=k_+\frac{d}{dt}k_+\frac{d}{dt}(\rho_+\mathbf{u}_+)=k_+^2\rho_+\ddot{\mathbf{u}}_++2k_+^2\dot{\mathbf{u}}_+\frac{d\rho_+}{dt}+k_+^2\mathbf{u}_+\frac{d^2\rho_+}{dt^2}, \tag{57}$$

$$Q_t=k_+\frac{d}{dt}k_+\frac{d\rho_+}{dt}. \tag{58}$$

By this device we generalize the $\rho\mathbf{u}$ of classical theory to an expression involving terms in $\dot{\mathbf{u}}$, $\ddot{\mathbf{u}}$, a generalization which fits in with one's intuitions as to the proper fitness of things, and yet we do this in a manner which retains harmony with the restricted theory of relativity. Indeed, it is possible to generalize the expressions for $\mathbf{S}_+$, $(S_t)_+$ of equations (23) and (24) still further by the addition of 4-vector terms of higher orders, and to do this without adding materially to the complexity of the theory. We shall, however, content ourselves with two terms.

For uniform rotation, and in view of the fact that $k_+=(1-q_+^2/c^2)^{-\frac{1}{2}}$ is constant with the time, equations (55)-(58) become, in terms of the distance r from the axis of rotation and the angular velocity ω,

$$\mathbf{P}=-k_+r\omega^2\rho_+, \tag{59}$$

$$P_t=k_+\frac{d\rho_+}{dt}, \tag{60}$$

$$\mathbf{Q}=k_+^2r\omega^3\rho_+, \tag{61}$$

$$Q_t=k_+^2\frac{d^2\rho_+}{dt^2}, \tag{62}$$

where we have neglected the terms involving $d\rho_+/dt$ and $d^2\rho_+/dt^2$ in equations (55) and (57) in comparison with the first terms on the right hand sides of these equations. This approximation involves neglecting quantities of the order $\dot{\rho}_+/\rho_+\omega$ and $\ddot{\rho}_+/\rho_+\omega^2$ compared with unity. Now although one of the purposes of the generalization discussed in this paper is to secure a finite value of $d\rho_+/dt$ to provide for the atmospheric electric current, it turns out that the value of $d\rho_+/dt$ necessary is such that $\dot{\rho}_+/\rho_+ = -1.7\times10^{-30}$, so that since $\omega = 7.3\times10^{-5}$, for the earth, the quantity $\dot{\rho}_+/\rho_+\omega$ is quite negligible compared with unity. Again, even though $\ddot{\rho}$ were so large that $\dot{\rho}$ changed by 100 per cent. in one second $\ddot{\rho}_+/\rho_+\omega^2$ would only amount to a quantity of the order 3×10^{-22} and would be negligible compared with unity. As a matter of fact our theory will provide for a constant value of $\dot{\rho}_+$.

In applying these expressions in (52) and (53) we should, strictly speaking, observe that the potentials are retarded, as indicated. However, in view of the uniformity of the rotation the retarded potentials only differ from the unretarded ones in virtue of such change of the numerators with time as results from the slow death of positive electricity provided for by (51), and the effect of neglecting the retarded features is insignificant (see Appendix, notes 9 and 10). The extreme closeness of approximation to a steady state also renders insignificant the contribution $-\frac{1}{c}\frac{\partial}{\partial t}(\mathbf{U}_+ + \mathbf{U}_-)$ to $\mathbf{E}_+ + \mathbf{E}_-$ as given by (54) (see Appendix, note 9).

$\mathbf{P}$ represents a vector in the direction opposite to the radius vector, *i.e.*, in the direction of the acceleration, while $\mathbf{Q}$ represents a vector perpendicular thereto. $\beta\mathbf{Q}$ thus acts in (53) exactly like a current density in giving rise to a magnetic field. Moreover, it turns out that the value which has to be provided for $\frac{d\rho_+}{dt}$ in order to give rise to the necessary atmospheric electric current is such that both $d\rho_+/dt$ and $d^2\rho_+/dt^2$ are negligible as regards their power to give rise to values of P_t and Q_t which would contribute appreciably to $\phi_+ + \phi_-$ through (52) and so to $\mathbf{E}_+ + \mathbf{E}_-$, through (54), (see Appendix, note 9).

The smallness of $\phi_+ + \phi_-$ is very important. Had we attempted to choose for $\mathbf{S}_+$, $ic(S_t)_+$ a 4-vector such as would cause a piece of moving matter to be the equivalent of a current in the general field equations in such a way that the apparent current density was proportional to the velocity of the matter, we could not have avoided the presence of a fourth component of undesirable magnitude. In fact, the four components would have been in the ratio ku_x, ku_y, ku_z, kic, so that if in the general equations (39)-(53), $A\mathbf{u}$ had survived as a numerator for $\mathbf{U}_+ + \mathbf{U}_-$ after the cancellation of the parts $\rho_+\mathbf{u}_+$ and $\rho_-\mathbf{u}_-$ corresponding to the positive and negative electricity on the sphere, there would also have survived in the numerator of the integral for $\phi_+ + \phi_-$ the quantity A. The result of this would have been that we should have been compelled to admit that a rotating sphere which gave rise to a given magnetic field as a result of its rotating neutral matter would also give rise to an electric field, and this electric field would be as large as would be that electric field which we should have had if we had obtained the magnetic field in question by the rotation of a solid sphere of electricity. There is thus a very fundamental objection to any theory which supposes that a piece of neutral matter may act to produce magnetic field entirely in virtue of its velocity.

It is not therefore to the expressions αP_t and βQ_t that we shall look for an explanation of the earth's electric field, but rather to the accumulation of unbalanced negative charge resulting from the continual death of the positive electricity. This accumulation of negative charge piles up until the field which it produces causes an atmospheric electric current into space of such amount as to maintain equilibrium with the rate of death of positive electricity.

It is by equating the atmospheric electric current to the volume integral of $\partial\rho_+/\partial t$ throughout the sphere that we shall determine the magnitude of the numerical factor in $\alpha\mathbf{P}$; for, as will appear, div $(\beta\mathbf{Q})$ is zero for the case of uniform rotation.

If Γ_r and Γ_ψ are respectively the components of a vector in the direction of the radius vector and tangential to a circle about the origin, we have, in cylindrical coordinates for the case where there is no variation of the quantities along the z axis,

$$\text{Div }\mathbf{\Gamma} = \frac{1}{r}\frac{\partial}{\partial r}(r\Gamma_r) + \frac{1}{r}\frac{\partial \Gamma_\psi}{\partial \psi}. \tag{63}$$

From this, and from (61), we see that, if β is independent of ψ as it is from symmetry*,

$$\text{Div }\beta\mathbf{Q} = 0. \tag{64}$$

Now it turns out that while the main terms of (51) are of the first order of magnitude in α, those involving the time derivatives of αP_t and βQ_t are of the third order in α and are quite negligible (see Appendix, note 9). Indeed, they would be absolutely zero except for the slight departure from a steady state resulting from the slow death of the positive electricity. Omitting, then, the terms in question and using (59) and (64), (51) becomes

$$\frac{-\partial\rho_+}{\partial t} = -\text{div }\omega^2\rho_+\alpha k_+ r = -\frac{\omega^2\rho_+}{r}\frac{\partial}{\partial r}(\alpha k_+ r^2). \tag{65}$$

*It might be of interest to call attention to a further fact. We have seen that $\alpha\mathbf{P}$ represents a vector in a radial direction, and $\beta\mathbf{Q}$ represents a vector in a direction tangential to a circle about the origin. The 3-vector formed from the 4-vector of next higher order is radial and opposite in direction to $\alpha\mathbf{P}$. In fact the 3-vectors formed from the 4-vectors of successively higher orders are at right angles to each other, each one being advanced one right angle in relation to its predecessor. It follows from this that, if we call the ordinary vector $\rho\mathbf{u}$ of electromagnetic theory the first, $\alpha\mathbf{P}$, the second, $\beta\mathbf{Q}$, the third, and so on then, for uniform rotation, all vectors of odd orders are such as to have no divergence and so contribute nothing to the violation of the equation of continuity, while those of even order do give rise to such violation. On the other hand, the vectors of odd order, being all parallel to $\rho\mathbf{u}$, act in the vector potential of (53) as the equivalents of currents flowing in circles about the axis of rotation. They are thus of a nature such as to provide in a general way for a magnetic field like that of the earth. On the other hand, the vectors of even order, being radial, are totally unsuitable for such purposes. Happily the magnitudes which it becomes necessary to assign to them are such that their influence on the magnetic field is totally negligible.

The form of $-\partial\rho_+/\partial t$ as a function of r and ω will depend upon the form of the invariant α. If α is taken of such form that, for uniform rotation as a rigid body $k_+\alpha$ is a constant independent of r, we have, for this case

$$\frac{-\partial\rho_+}{\partial t} = -2\omega^2\rho_+\alpha k_+. \tag{66}$$

In the steady state, the atmospheric electric current from the whole earth will consequently be

$$J = \frac{8\pi}{3}\,\omega^2\rho_+\alpha k_+ a^3, \tag{67}$$

and if the material density D is proportional to the number of protons per c.c., as is approximately true, and if α is a constant, we should have J proportional to $D\omega^2a^3$ approximately. In Table I the values of J are given for this case, the values here being in electrostatic units, and based on the known value -3×10^{12} e.s.u. for the earth. It will be seen that the small sphere would give an easily measurable current density amounting to 10^3 e.s.u. If, however, we can secure a value of J proportional to the first power of ω, so that J is proportional to $D\omega a^3$, the current for the small sphere would be practically immeasurable, and even for the sun the value suggested would not be unreasonable in comparison with the known value for the earth. As we shall see it is possible to choose, for α, an invariant which for uniform rotation reduces to

$$\alpha = \alpha_0 k^{-1}\omega^{-1} \tag{68}$$

where α_0 is independent of r and ω so that, anticipating to the extent of assuming this value for α, we have, from (67),

$$J = \frac{8\pi}{3}\,\alpha_0\rho_+\omega a^3$$

or, if m is the combined mass of a proton and electron, e_+ the charge of a proton, and D the density of the earth,

$$J = \frac{8\pi}{3}\,\alpha_0\frac{e_+}{m}D\omega a^3. \tag{69}$$

The atmospheric electric current from the whole earth amounts to about -3×10^{12} e.s.u. so that, putting $a=6.4\times10^8$ cm., $\omega=7.3\times10^{-5}$, $D=5.5$, $e_+=4.8\times10^{-10}$ e.s.u., $m=1.64\times10^{-24}$ we find,

$$\alpha_0 = -1.2\times10^{-26}. \tag{70}$$

From (66) and (68)

$$-\frac{1}{\rho_+}\frac{d\rho_+}{dt} = -2\omega\alpha_0,$$

so that

$$-\frac{1}{\rho_+}\frac{\partial\rho_+}{\partial t} = 1.75\times10^{-30},$$

which corresponds to a loss of charge amounting to less than 1 proton per cc.

per day. The required rate of death would correspond to a loss of 0.5 per cent. of the earth's mass in 10^{20} years.

Using in (68) the value of α_0 given in (70) and referring to (53), we readily find that, quite apart from the fact that the vector $\alpha\mathbf{P}$ is in a direction perpendicular to that necessary to give the earth's magnetic field, its magnitude is quite negligible as regards any measurable influence it may have on $\mathbf{U}_+ + \mathbf{U}_-$ (see Appendix, note 10). It is thus to the term $\beta\mathbf{Q}$ in (53) that we must look for the origin of the magnetic field. As we have already remarked, this term is a vector in the proper direction, and it is, in fact, the analytical equivalent of a tangential current density of amount $\beta\mathbf{Q}$, since it occurs in the vector potential exactly in the way that such a current density would occur. Putting $\mathbf{i}$ for this apparent current density we thus have

$$\mathbf{i} = \beta\mathbf{Q}$$

or, from (61),

(71) $$i = -\beta\rho_+ k_+^2 r\omega^3,$$

the minus sign indicating that β must be negative if i is to correspond in sign to ω.

Now we have seen that, in order to conform to the relative requirements of the sun, the earth and the small sphere of high angular velocity it is necessary to have i proportional to $Dr^3\omega^4$. Hence the invariant β must be such that, for uniform rotation, it reduces to

$$\beta_0 r^2\omega / k_+^2$$

where β_0 is independent of r and ω. However, the factor k_+ is so nearly equal to unity that its presence or absence does not appreciably affect the conclusions.

Choice of the invariants α and β.

As we have already seen, it is necessary to choose for α and β invariants which, for a case of uniform rotation, reduce to $\alpha_0 k_+^{-1}\omega^{-1}$ and $\beta_0 r^2\omega/k_+^2$ respectively, where α_0 and β_0 are independent of r and ω.

The simplest way to form invariants which are not mere constants is to take the scalar product of two 4-vectors. The 4-vectors naturally at our disposal for this purpose are the velocity 4-vector, and all the 4 vectors of higher order obtained by operating successively with $k_+ \dfrac{d}{dt}$.

Table II gives the components of these 4-vectors, T_1, T_2, T_3, etc., referring symbolically to the whole 4-vector and $\dfrac{d}{ds}$ being written for the operator $k_+ \dfrac{d}{dt}$.

The last two columns give the form to which the 4-vector degenerates for a case of uniform rotation. The first of these two columns gives the resultant 3-vector formed by the first three components of the corresponding 4-vector and the direction of the 3-vector is indicated by the arrow attached to the circle. It will be observed that the 3 vectors corresponding to T_1, T_2, T_3, ... etc., are advanced successively in direction by one right angle.

TABLE II

4-Vector	Components				Degenerate Forms	
	x	y	z	4th	Resultant 3-vector	4th component
T_1	k_+u_x	k_+u_y	k_+u_z	k_+ic	$k_+r\omega$ ↻	kic
T_2	$\frac{d}{ds}(k_+u_x)$	$\frac{d}{ds}(k_+u_y)$	$\frac{d}{ds}(k_+u_z)$	$\frac{d}{ds}(k_+ic)$	$k^2{}_+r\omega^2$ ↙	0
T_3	$\frac{d^2}{ds^2}(k_+u_x)$	$\frac{d^2}{ds^2}(k_+u_y)$	$\frac{d^2}{ds^2}(k_+u_z)$	$\frac{d^2}{ds^2}(k_+ic)$	$k^3{}_+r\omega^3$ ↺	0
T_4	$\frac{d^3}{ds^3}(k_+u_x)$	$\frac{d^3}{ds^3}(k_+u_y)$	$\frac{d^3}{ds^3}(k_+u_z)$	$\frac{d^3}{ds^3}(k_+ic)$	$k^4{}_+r\omega^4$ ↗	0
T_5	etc.	etc.	etc.	etc.	$k^5{}_+r\omega^5$ ↻	0

These degenerate forms are obtained immediately if we remember that for uniform rotation $\frac{dk_+}{dt}$ is zero. We can readily see from the directions of the 3-vectors and from the fact that all the fourth components are zero except the first, that any scalar product of a 4-vector of odd order with one of even order is zero. Only scalar products of odd or of even 4-vectors survive, and these are obtained immediately in degenerate form by multiplying together the corresponding quantities in the column next to the last.

There are several ways in which we may build up invariants which degenerate to the desired forms $\alpha_0 k_+^{-1}\omega^{-1}$ and $\beta_0 r^2\omega/k^2{}_+$. Thus for example

$$\frac{(T_1 . T_3)^{1/4}}{(T_3 . T_5)^{1/4}} \text{ degenerates to } k_+^{-1}\omega^{-1}$$

and

$$\frac{(T_1 . T_3)^{7/4}}{(T_3 . T_5)^{3/4}} \text{ degenerates to } k_+ r^2\omega,$$

where the parenthesis with the dot denotes the scalar product of the two 4-vectors designated by the corresponding T's. The factor k_+ in the last expression in place of $k_+{}^{-2}$ causes no trouble since it is very nearly equal to unity and therefore does not affect the functional form of β appreciably. Thus, two of many forms which serve our purpose for α and β are

$$\alpha = \alpha_0 \frac{(T_1 . T_3)^{1/4}}{(T_3 . T_5)^{1/4}}, \tag{72}$$

$$\beta = \beta_0 \frac{(T_1 . T_3)^{7/4}}{(T_3 . T_5)^{3/4}}. \tag{73}$$

With this form of β our apparent current density is

$$i = -\beta_0 \rho_+ k_+{}^3 \omega^4 r^3,$$

or, if m is the combined mass of a proton and an electron, e_+ the charge on a proton, and D the density of the substance of the earth

$$i = \beta_0 \frac{e_+ D}{m} k_+^3 \omega^4 r^3.$$

From (87), in Appendix, note 1, we consequently have, for the magnetic field at the pole

$$-\frac{64\pi}{315} \beta_0 \frac{e_+ D}{cm} k_+^3 \omega^4 a^4,$$

where the negative sign indicates for a positive β_0 a field directed radially inwards from west to east; k_+ is sensibly unity so that putting $H = 0.5$ c.g.s. unit at the pole,

$$e_+/m = 2.9 \times 10^{14}, \; D = 5.5, \; \omega = 7.3 \times 10^{-5}, \text{ and } a = 6.4 \times 10^8 \text{ cm.}$$

we find

(74) $$\beta_0 = -3.1 \times 10^{-24}.$$

We have thus determined forms of **P** and **Q** and of α and β which will make equations (25)-(34) give rise to a magnetic field comparable with that of the earth and to an atmospheric electric current, and so to a potential gradient comparable with that of the earth. Moreover, the forms of our expressions are such as to be consistent with the requirements imposed by the sun and by a small sphere rotating at high speed, and they are consistent with restricted relativity.

Gravitation

The incorporation of gravitation in the scheme comes about essentially along the lines of the theory of Lorentz.

In the first place we shall regard the right hand side of (33) and (34) when multiplied by e, as representing respectively the force on a proton and electron to the extent of believing that if we evaluate the total force due to one piece of neutral matter on another by adding up the forces (as above understood) on the individual protons and electrons, the result will represent the gravitational force between the pieces of matter. If we discuss the case for matter at rest we can neglect the terms involving velocity and magnetic field. As a matter of fact, it is easy to see that these terms are negligible even when the matter is in motion, as are also those correcting terms on the *fields* brought about by the modifications we have introduced into the electrodynamic laws (see Appendix, note 11). Thus, confining ourselves to the values of $\mathbf{E}_+$ and $\mathbf{E}_-$ calculated for stationary matter, and using ordinary electrostatic units, we have, for the forces δF_+ and δF_- on a proton and electron respectively and due to an element of matter of volume $d\tau_1$ containing N_1 protons and N_1 electrons,

$$\delta F_+ = \delta F_- = \frac{N_1 e^2 (1-\lambda)}{R^2} d\tau_1,$$

so that the total force on an element of matter of volume $d\tau_2$ containing N_2 protons and electrons per c.c. is

$$\delta F = \frac{2N_1N_2(1-\lambda)e^2}{R^2} d\tau_1 d\tau_2.$$

If m refers to the sum of the masses of a proton and electron, and if D_1 and D_2 are the densities of the substances

$$\delta F = 2D_1D_2\left(\frac{e}{m}\right)^2(1-\lambda)d\tau_1 d\tau_2.$$

This must be equated to

$$-\frac{GD_1D_2}{R^2} d\tau_1 d\tau_2$$

where G is the gravitational constant

Hence

$$\lambda - 1 = \frac{G}{2}\left(\frac{m}{e}\right)^2. \tag{75}$$

Putting $G = 6.6\times 10^{-8}$, $\dfrac{e}{m} = 2.9\times 10^{14}$

we have

$$\lambda = 1 + 3.9\times 10^{-37}. \tag{76}$$

SUMMARY

Collecting our results from equations (23)-(32) for the field equations, from (33) and (34) for the force equations, from (55)-(58) for the forms of **P**, P_t, **Q**, Q_t, and from (72) and (73) and Table II for α and β, we have as follows:

$$\frac{1}{c}\left\{\rho_+\mathbf{u}_+ + \alpha_0\frac{(T_1 . T_3)^{\frac{1}{4}}}{(T_3 . T_5)^{\frac{1}{4}}}k_+\frac{d}{dt}(\rho_+\mathbf{u}_+) + \beta_0\frac{(T_1 . T_3)^{\frac{7}{4}}}{(T_3 . T_5)^{\frac{3}{4}}}k_+\frac{d}{dt}k_+\frac{d}{dt}(\rho_+\mathbf{u}_+) + \frac{\partial\mathbf{E}_+}{\partial t}\right\} = \text{curl }\mathbf{H}_+ \tag{77}$$

$$\rho_+ + \alpha_0\frac{(T_1 . T_3)^{\frac{1}{4}}}{(T_3 . T_5)^{\frac{1}{4}}}k_+\frac{d\rho_+}{dt} + \beta_0\frac{(T_1 . T_3)^{\frac{7}{4}}}{(T_3 . T_5)^{\frac{3}{4}}}k_+\frac{d}{dt}k_+\frac{d\rho_+}{dt} = \text{div }\mathbf{E}_+, \tag{78}$$

$$-\frac{1}{c}\,\frac{\partial\mathbf{H}_+}{\partial t} = \text{curl }\mathbf{E}_+, \tag{79}$$

$$0 = \text{div }\mathbf{H}_+, \tag{80}$$

$$\frac{1}{c}\left(\rho_-\mathbf{u}_- + \frac{\partial\mathbf{E}_-}{\partial t}\right) = \text{curl }\mathbf{H}_-, \tag{81}$$

$$\rho_- = \text{div }\mathbf{E}_-, \tag{82}$$

(83) $$-\frac{1}{c}\frac{\partial \mathbf{H}_-}{\partial t}=\text{curl } \mathbf{E}_-,$$

(84) $$0=\text{div } \mathbf{H}_-,$$

(85) $$\frac{m_+}{e_+}\frac{d}{dt}(k_+\mathbf{w}_+)=\mathbf{E}_+ +\frac{[\mathbf{w}_+\mathbf{H}_+]}{c}+\lambda\left(\mathbf{E}_- +\frac{[\mathbf{w}_+\mathbf{H}_-]}{c}\right),$$

(86) $$\frac{m_-}{e_-}\frac{d}{dt}(k_-\mathbf{w}_-)=\mathbf{E}_- +\frac{[\mathbf{w}_-\mathbf{H}_-]}{c}+\lambda\left(\mathbf{E}_+ +\frac{[\mathbf{w}_-\mathbf{H}_+]}{c}\right),$$

where the meanings of the T's are given in Table II, and the value of λ by (75).

These equations are invariant under the transformation of the restricted theory of relativity, and they can readily be generalized to a form invariant in the sense of the general theory of relativity. With the values $\alpha_0=-1.2\times10^{-26}$ $\beta_0=-3.1\times10^{-24}$ $\lambda=1+3.9\times10^{-37}$ as given by equations (70), (74) and (76) they predict, for a sphere of the earth's size rotating with the earth's angular velocity, a magnetic field comparable with the earth's field but symmetrical about the axis of rotation. They give a death of positive electricity amounting to a disappearance of 0.5 per cent. of the earth's mass in 10^{20} years, leaving a surplus of negative electricity which passes off as an atmospheric electric current equal in magnitude to the earth's atmospheric electric current and corresponding, therefore, to a potential gradient equal to the earth's potential gradient. Moreover, the value which they predict for the magnetic field of the sun is correct within the limits of our knowledge of that magnitude, and the value which they predict for the sun's atmospheric electric current is not unreasonable, although we have no data to test it. They predict for a small sphere 10 cm. in radius and of the earth's density, rotating 100 times per second, a magnetic field, and a rate of accumulation of surplus negative electricity practically immeasurable under the possible conditions of experimentation. They predict an attraction between neutral matter to the proper extent determined by the gravitational constant.

The generalization which we have made in the electrodynamic equations is only one of many which may be made for the above purposes. Thus, as we have seen, the two additional terms in (77) and in (78) might be extended into a series involving higher time derivatives of the velocity, and the invariants involving the T's might be chosen in various ways; again, the left hand sides of (85) and (86) might be extended into series involving higher time derivatives of the velocity.

It is important to observe that, although the additional terms have been chosen by an appeal to the forms to which they degenerate for uniform rotation, the equations in the form presented in (77)-(86) or their further generalizations along lines indicated above are to be regarded as applicable to any kind of motion. It may be added that our reason for throwing the onus of the generalization on the positive electricity rather than the negative, or both, is that we have wished to avoid the discussion of such difficult questions as might arise in the matter of the conservation of charge in the case of rapid changes of velocity of β rays of very high speed. Even the question of conservation of charge to the extent required by experience in the case of such changes of speed as are suffered by alpha particles is one which invokes further discussion, but is too large a subject to enter into here. However, it will readily be seen that we have at our disposal plenty of latitude for avoiding such difficulties. Thus, to take a hypothetical case involving, for simplicity, only uniform rotation, suppose that we should encounter in the atom a situation of this kind involving rotation with very high angular velocity. It is possible that expression (66), with α as given by (68), and with the value of α_0 given by (70) would lead to a rate of death of charge inconsistent with the facts. However, it is quite possible to replace the invariant α as given by (72) and which degenerates to (68) by, for example,

$$\alpha = \alpha_0 \frac{(T_1 . T_3)^{1/4}}{(T_3 . T_5)^{1/4}} \left\{ 1 + \gamma \frac{(T_3 . T_5)^{s/4}}{(T_1 . T_3)^{s/4}} \right\}^{-1},$$

where γ is a very small constant and s is a positive integer. For uniform rotation α will now degenerate to

$$\alpha = \alpha_0 k_+^{-1} \omega^{-1} \{ 1 + \gamma k_+^s \omega^s \}^{-1}.$$

By choosing a suitably large value of s and a suitably small value of γ we can make $\gamma k_+^s \omega^s$ insignificant for angular velocities comparable with that of the earth's rotation, or even with that of a sphere rotating with any practicable speed attainable with bodies of laboratory dimensions; and yet we can arrange that the term is as large as we choose for the high angular velocities concerned in the atomic problem. For very high angular velocities α will thus become

$$\alpha = \alpha_0 k_+^{-1} \omega^{-1} / \gamma k_+^s \omega^s$$

and (66) will become

$$-\frac{1}{\rho_+} \frac{\partial \rho_+}{\partial t} = \frac{-2\omega\alpha_0}{\gamma k_+^s \omega^s},$$

which on account of the large value of the denominator for very large values of ω will result in very small values of $\frac{1}{\rho_+} \frac{\partial \rho_+}{\partial t}$ for such cases.

Another question related to this matter arises even in the slow rates of death of charge involved in the earth's rotation. Are we to assume that the death takes place uniformly with time in each proton? A more reasonable view would be to the effect that only occasionally is a proton affected and then the whole proton dies. On such a view it would be necessary to appeal to the same sort of arguments as are made in attempts to reconcile quantum phenomena with classical electrodynamics, and regard equations (77)-(86) as holding only in a statistical sense.

The Reasonableness of the Equations

From a purely mathematical standpoint such a generalization as we have made is satisfying in that it replaces the term $\rho\mathbf{u}$ of classical electrodynamics by an infinite series containing all orders of time derivatives of the velocities. If one asks what the physical significance of these extra terms is, we can well reply that they have as much right to physical significance as the term $\rho\mathbf{u}$, and only lack of familiarity with them suggests otherwise. If, however, one is pressed for a statement which will be more agreeable to the non-mathematical mind, he may speak thus:—

Suppose we have a circle of wire carrying a current. Then, according to ordinary physical ideas, we imagine the magnetic field as produced entirely by the velocity and density of the electricity. However, the electricity has acceleration $\omega^2 r$, rate of change of acceleration, and so on. Why should it not be the case that, even for a steady state* where $\mathbf{E}$ and $\mathbf{H}$ are independent of the time, the acceleration, rate of change of acceleration, etc., give rise to magnetic fields? If such were true, then, in the case of a current flowing in a circle, for example, we should get from that current when produced by a density ρ and a velocity $\mathbf{u}$, a different magnetic field from that produced by a density 0.5ρ and a velocity $2\mathbf{u}$. It is such a state of affairs as this that the extra terms in our equations provide for; and, if the phenomena which they represent were of sufficiently large magnitude to impress themselves upon us in every day experiments, we should doubtless learn to attach as strong a physical significance to these terms as we now attach to the main term $\rho\mathbf{u}$.

Of course the actual values of the terms in our equations are extremely small when applied to such motions as are concerned in the earth's rotation. Thus, for this case, the first two additional terms in (77) amount at most to only one part in 10^{26} and 5 parts in 10^{19} respectively in comparison with the term $\rho\mathbf{u}_+$; and, as we have seen, the quantity λ in (85) and (86) differs from unity by only 3.9×10^{-37}.

The question of whether the foregoing ideas do constitute the germ of an explanation of terrestrial magnetism and atmospheric electric phenomena is one which is perhaps apart from their merits as forming a consistent scheme in themselves. It is not without interest to observe, however, that if anyone objects to them as constituting a theory of the phenomena in question he may at least take the facts that the earth's magnetic field and atmospheric-electric current are as small as they are as evidence of an extraordinary completeness in the symmetry of the equations for the fields due to positive and negative electricity, particularly as regards any generalization of those equations along the lines we have indicated.

Appendix

Note 1. Let there be a circular ring carrying a current I e.s.u., the ring being of such size that its diameter subtends an angle 2β at the centre of a sphere

*We here limit the discussion to that for a steady state as, in other cases, even the classical electrodynamics cause the magnetic field to depend upon the accelerations and higher derivatives.

of radius σ. Then the magnetic potential Ω, at a point R, θ, outside the sphere is*, in e.m.u.,

$$\Omega=\frac{2\pi I}{c}\sin^2\beta\sum_{s=1}^{s=\infty}\left(\frac{1}{s+1}\right)\frac{\sigma^{s+1}}{R^{s+1}}\frac{dP_s(\beta)}{d\mu}P_s(\theta),$$

where $\mu=\cos\beta$, and the P's are the Legendre coefficients.

Apply this to a ring of cross section $\sigma d\beta d\sigma$ carrying a current of density $i=AD\omega^m r^n=AD\omega^m\sigma^n\sin^n\beta$, as specified in the text. Integrate from $\beta=$zero to π and from $\sigma=$zero to a. Then form minus the differential coefficient with respect to R and put $R=a$, to obtain the vertical component of the field at the surface.

In this way we obtain

$$H_z=\frac{2\pi}{c}AD\omega^m a^{n+1}(n+1)\sum_{s=1}^{s=\infty}\frac{P_s(\theta)}{n+s+3}\int_{-1}^{+1}(1-\mu^2)^{\frac{n-1}{2}}\mu P_s(\beta)d\mu.$$

For $\theta=0$, the quantity under the sum becomes a function of n only, so that

$$H_z=AD\omega^m a^{n+1}F(n).$$

Working out, in greater detail, the result for $n=3$, we readily find

$$H_z=\frac{32}{35c}A\pi D\omega^m a^4\left[\frac{P_1(\theta)}{3}-\frac{P_3(\theta)}{9}\right],$$

and, for $\theta=0$, *i.e.*, at the pole,

$$H_z=\frac{64}{315c}A\pi D\omega^m a^4. \tag{87}$$

Note 2. This equation makes its appearance in electrodynamics through the assumption that the electron moves so that the integral of $\left(\mathbf{E}+\frac{[\mathbf{uH}]}{c}\right)\rho d\tau$ taken all over the electron is zero, $\mathbf{E}$ and $\mathbf{H}$ here referring to the resulting field including the field of the electron itself. It is by separating this integral into two parts, one due to the field of the electron and the other due to the external field, and then evaluating the former in terms of the motion of the electron that the terms on the right hand side of (6) find their origin, and with them the significance of electronic mass in relation to the charge, and electronic radius. An assumption has to be made as to the variation of the shape of the electron with its motion. It seems to have been customary to assume that, at each instant the electron suffers the Lorentzian contraction appropriate to a steady

*Maxwell's *Treatise on Electricity and Magnetism*, 3rd edition, Vol. 2, p. 333.

velocity equal to the velocity at that instant. Under these conditions the law of motion derived is not invariant under the Lorentzian transformation, as was shown by S. Jacobsohn*.

Moreover, such an assumption as to change of shape with velocity would, in certain cases, lead to the conclusion that parts of the electron would have to travel with a velocity greater than that of light. For example, when the acceleration is so great that, starting from rest, the centre of the electron attains the velocity c in a time a/c, where a is the radius of the spherical electron at rest, the assumption would cause the rear parts of the electron to attain a velocity greater than c, in order that the postulated contraction could be assumed.

The equation of motion of an electron has been expressed in a modified form by Leigh Page†, who assumes that, at each instant, the electron is a sphere with all of its parts moving with equal though not constant velocity when it is referred to a system of axes with respect to which it is momentarily at rest, and that in this system of axes its acceleration and higher derivatives are such as to give zero for the integral of $\left(\mathbf{E}+\frac{[\mathbf{uH}]}{c}\right)\rho d\tau$ when evaluated over it. When this law is expressed in terms of an arbitrary system of axes, it assumes the form

$$\iiint\left(\mathbf{E}+\frac{[\mathbf{uH}]}{c}\right)\rho d\tau=\frac{e^2}{6\pi ac^2}\,\frac{d}{dt}\,\mathbf{w}\left(1-\frac{q^2}{c^2}\right)^{-\frac{1}{2}}+\mathbf{B}$$

where $\mathbf{B}$ involves all the time derivatives of the motion, and where $\mathbf{E}$ and $\mathbf{H}$ refer to the external field, and $\mathbf{w}$ refers to the velocity of the centre of the electron. If one assumes that $\mathbf{E}$ and $\mathbf{H}$ in this equation are sensibly constant over the electron, the left hand side of the equation assumes the form $\left(\mathbf{E}+\frac{[\mathbf{wH}]}{c}\right)e$; and the equation so obtained is invariant‡. However, it would seem that the invariance does not hold to a degree of approximation which takes account of the variation of the external field over the electron, a condition which is not out of the realms of practical interest when one considers the sharp changes of field which an electron may encounter when coming into the vicinity of another which is moving with a velocity comparable with c, and whose tubes of force are therefore concentrated greatly towards its equatorial plane.

It would thus seem that the assumption of zero force on the electron does not lead, in any of its forms, to expressions which are well adapted formally to the requirements of relativity. However, it is possible to write down an equation which has all the generality of those suggested by electrodynamics and more by equating $k\left(\mathbf{E}+\frac{[\mathbf{wH}]}{c}\right)e$ to a series composed of vectors formed from the first three components of the acceleration 4-vector and the first three components of all

*In a paper entitled *Note on the Force Equation of Electrodynamics* read before the International Mathematical Congress, Toronto, August, 1924.

†Phys. Rev., Vol. 11, p. 398, 1918.

‡L. Page: Phys. Rev., Vol. 24, pp. 627-630, 1924.

the 4-vectors of higher order, each vector being provided with an invariant coefficient, the understanding being that **E** and **H** refer to the external field at a specified point in the electron, and $k=\left(1-\frac{q^2}{c^2}\right)^{-\frac{1}{2}}$. The generalization may be further enhanced by the insertion of invariants under any of the time derivatives, a procedure which does not destroy the 4-vector character. It is known that the three components of $k\left(\mathbf{E}+\frac{[\mathbf{wH}]}{c}\right)$ together with $ki(\mathbf{E}.\mathbf{w})/c$ constitute a 4-vector, i being $\sqrt{-1}$. Thus the equation obtained will be invariant provided that the invariant coefficients are such that the 4-vector composed of the series of terms referred to possesses the property that its fourth component is i/c times the scalar product of **w** with its other three components (see Appendix, note 6). The series of 4-vectors contains derivatives of all orders so that it has all the desired characteristics for exhibiting radiation reactions and the like, and its first and most important term is, for the case where the invariants associated with it are constants, identical in form with the corresponding term in the ordinary electrodynamic equation. Its generality is, however, enhanced in comparison with the electrodynamic expression by the greater arbitrariness of the invariants associated with the various 4-vectors. It is in such an extended sense as this that we shall understand equation (6) and equations (33) and (34) for the most general case.

Note 3. Equation (16) is automatically satisfied by (19). Substituting from (19) in (15) we find that the latter equation is satisfied if

$$\operatorname{curl}\left(\mathbf{E}+\frac{1}{c}\frac{\partial \mathbf{U}}{\partial t}\right)=0,$$

and this is satisfied by (18).

Substituting from (18) and (19) in (13), we find that (13) is satisfied if

$$\frac{\mathbf{S}}{c}=\frac{1}{c^2}\frac{\partial^2\mathbf{U}}{\partial t^2}+\frac{1}{c}\operatorname{grad}\frac{\partial\phi}{\partial t}+\operatorname{curl}\operatorname{curl}\mathbf{U},$$

i.e., if

$$-\frac{\mathbf{S}}{c}=\nabla^2\mathbf{U}-\frac{1}{c^2}\frac{\partial^2\mathbf{U}}{\partial t^2}-\operatorname{grad}\left\{\operatorname{div}\mathbf{U}+\frac{1}{c}\frac{\partial\phi}{\partial t}\right\},$$

and this is satisfied if (21) is true.

Again, substituting from (18) in (14) we find the latter satisfied if

$$S_t=-\frac{1}{c}\frac{\partial}{\partial t}\operatorname{div}\mathbf{U}-\nabla^2\phi,$$

i.e., if

$$-\frac{S_t}{c}=\nabla^2\phi-\frac{1}{c^2}\frac{\partial^2\phi}{\partial t^2}+\frac{1}{c}\frac{\partial}{\partial t}\left\{\operatorname{div}\mathbf{U}+\frac{1}{c}\frac{\partial\phi}{\partial t}\right\}$$

which, again, is true if (21) holds. Hence (13)-(16) which also imply (17) are satisfied by the solutions (18)-(21).

Note 4. Consider the second of equations (20). Divide all space into volume elements $d\tau$. The value of $\mathbf{U}_\lambda$ of $\mathbf{U}$ corresponding to the direction of an element of length $d\lambda$ at any point P at time t is determined by the values of $\mathbf{S}_\lambda$ in the various volume elements estimated at times $t-R/c$ where R is the radius vector from P to the corresponding volume element. Now consider a point P' displaced an infinitesimal amount $d\lambda$ from P, and again divide all space into volume elements $d\tau'$ such that each of them may be regarded as obtained from a corresponding volume element $d\tau$ by displacement equal and parallel to $d\lambda$. The value of $\mathbf{U}_\lambda$ at P' at time t is determined by the values of $\mathbf{S}_\lambda$ in the volume elements $d\tau'$ at the times R/c earlier. Thus, corresponding to the same time t at P and P', the *effective* times in *corresponding* volume elements are the same and the distances of corresponding volume elements from P and P' respectively are equal, so that

$$4\pi c(\mathbf{U}_{\lambda_{p'}}-\mathbf{U}_{\lambda_p})=d\lambda\iiint\frac{\left[\frac{\partial\mathbf{S}_\lambda}{\partial\lambda}\right]}{R}d\tau$$

that is

$$4\pi c\,\frac{\partial\mathbf{U}_\lambda}{\partial\lambda}=\iiint\frac{\left[\frac{\partial\mathbf{S}_\lambda}{\partial\lambda}\right]}{R}d\tau$$

where the square brackets have the same significance as before. By considering displacements $d\mu$, $d\nu$, perpendicular to $d\lambda$, we thus immediately derive

$$4\pi c\ \mathrm{div}\ \mathbf{U}=\iiint\frac{[\mathrm{div}\ \mathbf{S}]}{R}d\tau. \tag{88}$$

Turning now to the first of equations (20), it is obvious that

$$4\pi\,\frac{\partial\phi}{\partial t}=\iiint\frac{\left[\frac{\partial S_t}{\partial t}\right]}{R}d\tau \tag{89}$$

where again the square brackets have the same significance as before; for the effective volume elements are the same at t as at $t+dt$ for a fixed point P, and the values of S_t which are effective at times t and $t+dt$ at P are those to be found at $t-R/c$ and $t+dt-R/c$ in the volume element corresponding to R.

Since in every volume element at *all instants*

$$\mathrm{div}\ \mathbf{S}+\frac{\partial S_t}{\partial t}=0,$$

we have, from (88) and (89),

$$\mathrm{div}\ \mathbf{U}=-\frac{1}{c}\,\frac{\partial\phi}{\partial t},$$

showing that this equation is the analytical equivalent of (17).

Note 5. On transforming equations (2)-(5) to the system whose coordinates x', y', z', t', are specified by (22) they revert to the same form in dashed letters, with the understanding that

$$(90)\qquad E_x' = E_x, \qquad E_y' = \epsilon\left(E_y - \frac{v}{c}H_z\right), \qquad E_z' = \epsilon\left(E_z + \frac{v}{c}H_y\right),$$

$$(91)\qquad H_x' = H_x, \qquad H_y' = \epsilon\left(H_y + \frac{v}{c}E_z\right), \qquad H_z' = \epsilon\left(H_z - \frac{v}{c}E_y\right),$$

$$(92)\qquad \rho' = \epsilon\rho\left(1 - \frac{u_x v}{c^2}\right).$$

Moreover if $d\tau$ and $d\tau'$ are corresponding elements of volume

$$\rho' d\tau' = \rho d\tau,$$

so that the densities ρ and ρ' have corresponding significances to observers who agree on the equality of total charges in corresponding volume elements. It remains to show, however, that $\mathbf{E}'$, $\mathbf{H}'$, and $\mathbf{E}$, $\mathbf{H}$, have corresponding significances for corresponding observers. To test this point it is necessary to appeal to the definitions of $\mathbf{E}$ and $\mathbf{H}$. It would seem that the customary treatments of the matter leave much to be desired. This will immediately be apparent to anyone who takes as his definition for $\mathbf{H}$ the force on unit pole, and tries to convince himself that $\mathbf{H}'$ is the corresponding quantity which the observer in the dashed system would measure for the force on a unit pole. As a matter of fact, the whole significance of the invariance is best exhibited by a line of reasoning which does not invoke any definitions of $\mathbf{E}$ and $\mathbf{H}$ in terms of the force on anything; for at best such a procedure becomes meaningless when applied to the field at a *point inside* an electron, for example, in which case it is impossible to fall back upon the approximation that the field is sensibly constant over the entity which is used to measure it. We shall first outline the procedure in terms of the more logical attitude.

Significance of the invariance apart from definition of the fields in terms of forces on assigned entities.

In the cases of the classical equations (2)-(6) the starting point is the assumption of an invariant (the charge) associated with any element of volume. The density ρ is defined as the value of the invariant divided by the volume element. This leads to the law of transformation (92). Equations (7)-(11) or their analytical equivalents (2)-(5) now constitute the *definitions* of $\mathbf{E}$ and $\mathbf{H}$. Since equations (2)-(5) revert to the same form on transformation with $\mathbf{E}'$, $\mathbf{H}'$ as defined by (90) and (91) replacing $\mathbf{E}$ and $\mathbf{H}$, with ρ' given by (92), and with $\mathbf{u}'$ as given by the Lorentzian transformation replacing $\mathbf{u}$, their solutions (7)-(11) also revert to corresponding forms with ρ' as defined by (92). Since, however, with ρ' defined by (92) the equations insure the invariance of $\rho d\tau$, it follows that the ρ' defined by (92) is the quantity obtained by dividing the primary invariant (the charge in a volume element) by the volume element. Thus in terms of the original invariant, ρ', $\mathbf{E}'$ and $\mathbf{H}'$ and ρ, $\mathbf{E}$, $\mathbf{H}$, have corresponding insignificances to corresponding observers. So far there has been no statement of law—all has been the erection of scaffolding. The law of physics now enters when we state such a relation as (6) as holding between the motion of an electron and the quantities $\mathbf{E}$ and $\mathbf{H}$ as above defined. The definitions of $\mathbf{E}$ and $\mathbf{H}$ have provided for $\mathbf{H}$, $-i\mathbf{E}$, transforming as a six vector so that, writing $k\equiv(1-q^2/c^2)^{-\frac{1}{2}}$, we have, as is well known, that $k\left(\mathbf{E}+\frac{[\mathbf{wH}]}{c}\right)$, $ki(\mathbf{E}.\mathbf{w})/c$ constitute a 4-vector. Consequently if k times the right hand side of (6) together with the scalar product of this right hand side with $i\mathbf{w}/c$ forms a 4-vector the relation (6) will be invariant (see Appendix, note 6).

The whole law of physics implied in the scheme of equations is then to the effect that it is possible to assign to the volume elements of space numbers (the charges) which are the same for both observers and other numbers $\mathbf{u}$ which transform between the observers as a velocity transforms, and such that with the definition of ρ, $\mathbf{E}$, and $\mathbf{H}$ we have made in terms of these two sets of numbers, equation (6) holds for the motion of an electron.

The argument proceeds in similar manner in the case of the general equations (23)-(35). Here again we have for the positive electricity an invariant, the positive charge in the volume element. In terms of it we define ρ_+ by dividing by the volume element. $\mathbf{P}$, icP_t and $\mathbf{Q}$, icQ_t then become defined in such manner as we have adopted in equations (55)-(58), for example; and, the invariants α and β become constructed out of such quantities as have already been defined. If α, β, $\mathbf{P}$, P_t, $\mathbf{Q}$, Q_t are to have in their expressions anything other than the quantities already defined or quantities which can be constructed from them, it will be necessary to make explicit definitions of these quantities in senses which would have meanings for the two observers. Presumably this would be done by defining them in terms of certain invariants as ρ_+ is defined in terms of the invariant charge in the volume element*. $\mathbf{E}_+$, $\mathbf{H}_+$, now become defined by (25)-(28), or by their analytical equivalents of the form (18)-(20) as applied to this case. Equations (29)-(32) being of the classical form, $\mathbf{E}_-$ and $\mathbf{H}_-$ become defined as in the classical case. The foregoing definitions ensure that $\mathbf{H}_+, -i\mathbf{E}_+$ and $\mathbf{H}_-$, $-i\mathbf{E}_-$ transform as six vectors, and so the invariance of (33) and (34) is ensured even, if necessary, in a more general form provided always that, if $\mathbf{W}$ is the left hand side of one of these equations, the quantities $k\mathbf{W}$, $ik(\mathbf{W}.\mathbf{w})/c$ constitute a 4-vector. This condition is satisfied by equations (33) and (34) as they stand. (See Appendix, note 6.)

Significance of the invariance when the fields are defined in terms of the forces on assigned entities. For one who wishes to retain the concept of the field as the force on an assigned entity, the argument regarding the invariance may assume the following form.

Suppose that $k\mathbf{w}$, $ki(\mathbf{W}.\mathbf{w})/c$ constitute 4 components of a 4-vector involving time derivatives of the motion of the electron. Suppose that, leaving the meanings of $\mathbf{E}$ and $\mathbf{H}$ unspecified for the moment, we write

(93)
$$\mathbf{E}+\frac{[\mathbf{wH}]}{c}=\mathbf{W},$$

so that

(94)
$$(\mathbf{E}.\mathbf{w})=(\mathbf{W}.\mathbf{w}).$$

It then follows as a direct algebraical consequence that, since $k\mathbf{W}$, $ki(\mathbf{W}.\mathbf{w})/c$ constitute a 4-vector

(95)
$$\mathbf{E}'+\frac{[\mathbf{w}'\mathbf{H}']}{c}=\mathbf{W}',$$

(96)
$$(\mathbf{E}'\mathbf{w}')=(\mathbf{W}'.\mathbf{w}'),$$

where $\mathbf{E}'$, $\mathbf{H}'$ are related to $\mathbf{E}$ and $\mathbf{H}$ by (90) and (91).

Now $\mathbf{W}$ involves constants and the time derivatives of the motion in the undashed system, and $\mathbf{W}'$ involves these in the same form in the dashed system. However, it is not to be concluded that (93) and (94) in transforming to (95) and (96) have remained invariant under the transformation in the practical sense unless we can show that $\mathbf{E}'$, $\mathbf{H}'$ and $\mathbf{E}$, $\mathbf{H}$ have corresponding meanings to the two observers in the two systems. In order to provide for this it is necessary to say what we intend to mean by $\mathbf{E}$ and $\mathbf{H}$. Thus, suppose we define $\mathbf{E}$ as the value of $\mathbf{W}$ when $\mathbf{w}$ is zero, and $\mathbf{H}$ as the limiting coefficient of $\mathbf{w}/c$ (taken in the sense of a vector product, with $\mathbf{w}$ directed so as to make the vector product a maximum) in the expression for $\mathbf{W}-\mathbf{E}$, when $\mathbf{w}$ is made infinitesimal†.

Then, obviously from equation (95) the corresponding definitions in the dashed system will make $\mathbf{E}'$ and $\mathbf{H}'$ the electric and magnetic fields in that system. Thus our definitions of $\mathbf{E}$ and $\mathbf{H}$ combined with the truth of such a relation as (93) have provided for a relation between the vectors in the two coordinate systems such as is given by (90) and (91). In other words we have provided for a condition of affairs in which $\mathbf{H}$, $-i\mathbf{E}$ transforms as a 6-vector. It then follows that the right hand side of (2) and the right hand side of (3) with a factor i attached form a 4-vector, and the right hand side of (4) and the right hand side of (5) with a factor i attached form a 4-vector. Hence, since $\frac{\rho\mathbf{u}}{c}$, ρi form a 4-vector the whole set of equations (2)-(6) is invariant provided that the right hand side of (6) satisfies the requirements that $\mathbf{W}$ must have‡.

*These remarks considered with the definitions of the quantities $\mathbf{S}_+$, $(S_t)_+$ apply in similar fashion to the quantities $\mathbf{S}$, S_t of equations (13)-(16). It is important to notice that the fact that the charge is not conserved does not destroy its *invariance* in the sense of relativity theory. The form of the equations provides for $\rho' d\tau' = \rho d\tau$ at corresponding instants even though both of these quantities may vary with the time.

†Such definitions are closely in line with those which we should implicitly use in measuring $\mathbf{E}$ and $\mathbf{H}$ in terms of the motions of streams of electrons.

‡It is important to observe that the procedure calls for the use on the right hand side of (93) of such a quantity $\mathbf{W}$ that $k\mathbf{W}$ forms three components of a 4-vector whose fourth component is $ki(\mathbf{W}.\mathbf{w})/c$. In the language of classical relativity the 4-vector must be perpendicular to the velocity 4-vector. The 4-vector $k\frac{d}{dt}(k\mathbf{w})$, $k\frac{d}{dt}(kic)$ whose first three components represent the generalization of the acceleration in the customary expression for $\mathbf{W}$ satisfies this condition, as shown in Appendix, note 6.

Some sort of an explanation may appear necessary for the apparently complicated definition of **E** given above. This definition is a generalization of the elementary definition of **E** in terms of the force on unit charge, a definition understood presumably in a vague sort of way as the mass acceleration of the charge. Strictly speaking, it is not possible to make a definition of **E** which involves only the acceleration of the electron when $\mathbf{w}=0$ regardless of what values the other derivatives may have. At least, any such definition would be useless for the purpose of building up such a relation as (93) since that relation certainly will involve, in general, all the derivatives even when $\mathbf{w}=0$. It is useless to consider a case where all the other derivatives of $\mathbf{w}$ are zero except $\dot{\mathbf{w}}$, for such a condition of motion would be, in general, inconsistent with the equation of motion proposed, as regards its application at two consecutive instants.

Having defined **E**, we might take equation (93) as the definition of **H** without restricting ourselves to the limiting case where $\mathbf{w}$ approaches zero. In this case, the equation would be correct by definition. However, this procedure would not provide necessarily for the truth of (2)-(5) in terms of the **E** and **H** defined. The truth of these relations as holding between **E** and **H** defined in the manner stated would now constitute the fact of physics which the whole scheme expressed.

A knowledge of the form of the right hand side of (93) and of the constants occurring in it is necessary for the purpose of making the definitions of **E** and **H** cited above. In the absence of any attempt to derive the whole scheme of equations from some unifying principle, the form in question, and the numerical constants in it must be regarded as chosen in such a way that the **E** and **H** which become defined in terms of them satisfy (2)-(5). It is by no means certain nor likely that the choice of the form could always be made so that these equations are satisfied for any assigned distribution of ρ and $\rho\mathbf{u}$ and the assumption that it can constitutes the assumption implied in the whole scheme of equations.

Practically all that we have written with regard to the invariance of (2)-(5) may be said with regard to the invariance of (13)-(16) when taken in conjunction with a force equation such as (93). Any definitions of S_x, S_y, S_z, icS_t which provide for the 4-vector character of these quantities will insure the invariance of $S_t d\tau$. In general, the definitions would be founded upon the invariants peculiar to the form of $\mathbf{S}$, S_t.

The idea becomes clear by considering the more general equations (25)-(28), with $\mathbf{S}$, S_t defined as in (23) and (24). Here, since $\rho\mathbf{u}$, ρic is a 4-vector, $\rho d\tau$ is invariant, as in the classical equations. Thus, ρ may be defined by the two observers as the ratio of the invariant charge in the volume element to the magnitude of the volume element. In such a case as that implied in equations (55)-(58), where $\mathbf{P}$, P_t, $\mathbf{Q}$, Q_t are expressed in terms of ρ and $\mathbf{u}$, no further definitions are necessary to complete their meaning. In other cases, where they might involve in their expression concepts other than electric density, as for example density of magnetism of one sign, the meaning of the new concept in question would have to be sought in terms of a definition based on whatever invariants were appropriate to the forms of the expression.

Definitions of $\mathbf{E}_+$, $\mathbf{E}_-$, $\mathbf{H}_+$, $\mathbf{H}_-$, necessary to provide for the invariance of equations (25)-(35) may be made in the manner outlined above with slight complications depending upon the more elaborate forms of equations (33) and (34). Thus, let

$$k_+\mathbf{W}_+,\ k_+i\ (\mathbf{W}_+.\mathbf{w}_+)/c \text{ and } k_-\mathbf{W}_-,\ k_-\ i(\mathbf{W}_-\mathbf{w}_-)/c$$

be 4-vectors specified in terms of the motion of the positive and negative electrons, respectively. Then, holding in abeyance for a moment the definitions of $\mathbf{E}_+$ $\mathbf{H}_+$, $\mathbf{E}_-$, $\mathbf{H}_-$, let us write

$$\mathbf{W}_+=\mathbf{E}_++\frac{[\mathbf{w}_+\mathbf{H}_+]}{c}+\lambda\left(\mathbf{E}_-+\frac{[\mathbf{w}_+\mathbf{H}_-]}{c}\right) \tag{97}$$

and

$$\mathbf{W}_-=\mathbf{E}_-+\frac{[\mathbf{w}_-\mathbf{H}_-]}{c}+\lambda\left(\mathbf{E}_++\frac{[\mathbf{w}_-\mathbf{H}_+]}{c}\right). \tag{98}$$

Then, as a direct result of the 4-vector character associated with the $\mathbf{W}$'s it follows that

$$\mathbf{W}_+'=\mathbf{E}_+'+\frac{[\mathbf{w}_+'\mathbf{H}_+']}{c}+\lambda\left(\mathbf{E}'_-+\frac{[\mathbf{w}'_+\mathbf{H}'_-]}{c}\right), \tag{99}$$

$$\mathbf{W}'_-=\mathbf{E}'_-+\frac{[\mathbf{w}'_-\mathbf{H}'_-]}{c}+\lambda\left(\mathbf{E}'_++\frac{[\mathbf{w}'_-\mathbf{H}'_+]}{c}\right), \tag{100}$$

where $\mathbf{E}_+'$, $\mathbf{H}_+'$, $\mathbf{E}_-'$, $\mathbf{H}_-'$ are defined in terms of the corresponding undashed vectors by relation of the forms (90) and (91). By putting $\mathbf{w}_+=0$ in (97), the equation serves to define $\mathbf{E}_++\lambda\mathbf{E}_-$. By putting $\mathbf{w}_-=0$ in (98) the equation serves to define $\mathbf{E}_-+\lambda\mathbf{E}_+$. Hence both $\mathbf{E}_-$ and $\mathbf{E}_+$ are defined.

Then, by proceeding towards the limit when $\mathbf{w}_+=0$, in the manner adopted for the classical equations, (97) serves to define $\mathbf{H}_++\lambda\mathbf{H}_-$, since $\mathbf{E}_+$ and $\mathbf{E}_-$ are now defined; and in a similar way by proceeding towards the limit when $\mathbf{w}_-=0$, (98) serves to define $\mathbf{H}_-+\lambda\mathbf{H}_+$. Hence $\mathbf{H}_-$ and $\mathbf{H}_+$ become defined.

From the similarity in form between (97) and (99) and between (98) and (100), it now follows that if the observer in the dashed system defines the fields in the way corresponding to that adopted by the observer in the undashed system he will find the quantities $\mathbf{E}'_+$, $\mathbf{E}'_-$, $\mathbf{H}'_+$, $\mathbf{H}'_-$, and, as we have observed, these are related to the corresponding undashed quantities by the transformation of a six vector. The invariance of equations (25)-(32) then follows as before, since $\frac{\rho\mathbf{u}}{c}$, ρi constitute a 4-vector, and $\mathbf{S}_+$, $(S_t)_+$ constitute a 4-vector.

Note 6. Since, to take one case which will represent all of its kind,

$$k_+\left(\mathbf{E}_++\frac{[\mathbf{w}_+\mathbf{H}_+]}{c}\right)+\lambda k_+\left(\mathbf{E}_-+\frac{[\mathbf{w}_+\mathbf{H}_-]}{c}\right)$$

constitutes 3 components of a 4-vector of which the fourth is

$$i\{k_+(\mathbf{E}_+.\mathbf{w}_+)+\lambda k_+(\mathbf{E}_-.\mathbf{w}_+)\}/c,$$

it is obvious that this particular 4-vector has the property that its fourth component is i/c times the scalar product of $\mathbf{w}_+$ and its other three components. Hence, any other 4-vector $k_+\mathbf{W}$, k_+icW_t which is to be equated to it must have this property. In other words, we must have $k_+\mathbf{W}_t = k_+(\mathbf{W}\cdot\mathbf{w}_+)/c^2$. The acceleration 4-vector $k\dfrac{dk\mathbf{w}}{dt}$, $kic\dfrac{dk}{dt}$ has this property. For

$$\frac{i}{c}\left(\mathbf{w}\cdot k\frac{dk\mathbf{w}}{dt}\right) = \frac{i}{c}\left(kw_x\frac{dkw_x}{dt} + kw_y\frac{dkw_y}{dt} + kw_z d\frac{kw_z}{dt}\right)$$

$$= \frac{i}{2c}\frac{d}{dt}k^2(w_x^2 + w_y^2 + w_z^2)$$

$$= \frac{1}{2c}\frac{d}{dt}\left[\frac{q^2}{1-q^2/c^2}\right] = \frac{ic}{2}\frac{d}{dt}(k^2) = ick\frac{dk}{dt}.$$

The property is not shared in general by the 4-vectors corresponding to the higher time derivatives so that, in the use of such an expression as (38) it is necessary that the invariant coefficients a, b, etc., shall be restricted by the requirement that the desired condition shall be satisfied by the 4-vector as a whole.

Note 7. From a purely formal standpoint we may say that the content of the equations (25)-(34) with the relations (55)-(58) and with the understanding that the scheme of equations is to be invariant under the Lorentzian transformation and that the invariants α and β have been decided upon, is contained in the following statement. It is possible to assign to the volume elements of space positive and negative numbers (the charges) which are the same for observers moving with constant relative velocity, and by dividing these numbers by the corresponding volume elements to define other numbers ρ_+ and ρ_-. Further, it is possible to assign to the volume elements vectors, $\mathbf{u}_+$ and $\mathbf{u}_-$ (each transforming as a velocity) and other numbers and vectors associated with α and β in a way specified by the forms of these expressions, the numbers and vectors satisfying the conditions

$$\text{div}\,(\rho_-\mathbf{u}_-) + \frac{\partial\rho_-}{\partial t} = 0, \quad \text{and div}\,\mathbf{S}_+ + \frac{\partial(S_t)_+}{\partial t} = 0.$$

Further, it is possible to assign a certain number λ and to assign to certain entities (protons and electrons) whose presence we believe we can detect, numbers e_+, e_-, m_+, m_-, such that if $\mathbf{E}_+$, $\mathbf{E}_-$, $\mathbf{H}_+$, $\mathbf{H}_-$, are calculated by (25)-(32), and the values when so determined are substituted in (33) and (34), these equations will give the motions of the observable entities.

It will be, of course, obvious from the above that, apart from the terms involving α and β, if we should, for example, multiply e_+, e_-, and all the ρ's by a constant factor η and the m's by η^2, the equations would serve to determine for the entities the same motions as before. Similar remarks apply in general when the terms in α and β are retained, the exact extent of the ambiguity in the numbers depending, however, upon the forms of the terms whose factors are α and β.

The numerical values of the constants m_+, m_-, e_+, e_-, etc., become fixed ultimately by a certain association with charge magnitudes defined in terms of mechanical force. Thus, consider two parallel plates charged equally and negatively. The ordinary elementary theory of electrostatics leads to the expression $F=\sigma^2/2$ for the force per unit area between the plates. In the spirit of the general standpoint outlined above, which left an ambiguity in the numbers e_+, e_-, etc., the expression $F=\sigma^2/2$ is really a definition of the numerical value of σ in terms of F as measured mechanically. If we now form a picture of Millikan's oil drop experiment as carried out in such a way that the field used is that due to one of the above plates then, if f is the gain in the force on the drop, measured mechanically, resulting from the acquisition of one electron we have $f=X_-e_-$ where X_- is the field due to the charge density σ on the plate. Thus

$$f=\frac{\sigma}{2}e_-=\frac{\sqrt{2F}}{2}e_-, \tag{101}$$

which serves to determine the magnitude of e_-.

The number e_- having been determined, m_- becomes determined as X_-e_- divided by the initial acceleration which the electron suffers in the field X_-, this field being measured as above. If one desires a degree of logical precision which requires specification of the other time derivatives of the motion besides the acceleration he will be driven to the consideration of matters from the standpoint discussed in Appendix, note 5.

For the determination of e_+ and m_+ the procedure may be imagined as one similar to that outlined above, with the use, however, of a proton and a positively charged plate* to produce the field. Such an experiment is, at the present time, an ideal one only. The incorporation of extra theoretical considerations such as belief in the equality of e_+ and e_-, and of the recognized structure of the hydrogen molecule as a means to obtain m_+ relieves us of the necessity of this ideal experiment.

Finally, λ becomes determined in terms of the gravitational constant G by equation (75). An ideal experiment more closely in line with the significance of λ as it occurs in equations (33) and (34), is one in which the experiment on the determination of e_- is repeated with the field produced by a positively charged plate. In this case equation (101) is replaced by $f=\frac{\sqrt{2F}}{2}\lambda e_-$, which serves to determine λ in terms of e_- as already determined. It need hardly be pointed out that the determination of λ by this procedure would be impracticable in view of the limits of experimental accuracy.

*The use of the positively charged plate with the proton is only necessary formally in order to avoid the appearance of λ in the discussion at this point.

Absolute precision in the logical reasoning demands that, in the above, we pay attention to the gravitational attraction between the plates when measuring σ and to the gravitational pull of the plate on the oil drop and even on the electron whose charge is being measured. These corrections are, of course, absolutely insignificant from the practical standpoint; but, in order to satisfy one who would feel confused as to the meaning of things until some ideal method of eliminating them had been *stated*, it will suffice to suppose that the experiment on the determination of e_- is carried out with successively larger charge densities on the plates and with successively smaller oil drops. The value of e_- desired is then the limit corresponding to an infinitely large value of σ as above measured, and an infinitely small oil drop.

Note 8. Taking, for example, the motional effect on an initially stationary electron as the apparent measure of the electric field, that is, as the measure appropriate for one who regards the laws of classical electrodynamics as exact, we see from (34) that the apparent field will be represented by

$$\mathbf{E}_- + \lambda \mathbf{E}_+ = \lambda(\mathbf{E}_+ + \mathbf{E}_-) + (1-\lambda)\mathbf{E}_-.$$

Here $\mathbf{E}_+$ and $\mathbf{E}_-$ are the fields due to all the positive and negative electricity in the neutral body. For the case where the body is at rest, $\mathbf{E}_-$ is the field calculated from all the negative electricity in the body on the basis of the law of inverse squares; and when the body is in motion with a velocity appreciably less than that of light $\mathbf{E}_-$ has sensibly the same value as for the statical case. For the statical case, the term $(1-\lambda)\mathbf{E}_-$ represents merely the *gravitational* pull of the body on the electron as will be evident on referring to the section on Gravitation. It is therefore negligible. Moreover, since λ is very nearly equal to unity, $\lambda(\mathbf{E}_+ + \mathbf{E}_-)$ may be replaced by $\mathbf{E}_+ + \mathbf{E}_-$. As a matter of fact, in the case which interests us, viz., that of uniform rotation, it turns out that $\mathbf{E}_+ + \mathbf{E}_-$ is sensibly zero for a neutral body so that $(1-\lambda)\mathbf{E}_-$ is not necessarily negligible *compared* with it. However, in this case, as is shown in the text, we seek the electric field as resulting from an excess of negative over positive, that excess resulting from the death of positive electricity provided for by the equations. When compared with the electric field, for which we shall provide in this way, $(1-\lambda)\mathbf{E}_-$ is negligible.

Again, the force per unit charge on a negative electron moving in the absence of electric field is

$$\mathbf{F} = \frac{[\mathbf{w}_-\mathbf{H}_-]}{c} + \lambda \frac{[\mathbf{w}_-\mathbf{H}_+]}{c} = \lambda \frac{[\mathbf{w}_- . \overline{\mathbf{H}_+ + \mathbf{H}_-}]}{c} + (1-\lambda)\frac{[\mathbf{w}_-\mathbf{H}_-]}{c}.$$

Now $\mathbf{H}_-$ is the magnetic field produced on the basis of the classical theory by the rotation of all the negative electricity in the body. For the case of the earth, it amounts at the pole to 2×10^{18} gauss. However, as shown in the section on Gravitation, $\lambda - 1 = 3.9 \times 10^{-37}$. Hence, the term $(1-\lambda)\frac{[\mathbf{w}_-\mathbf{H}_-]}{c}$ is negligible in any situation which provides for a value of $\mathbf{H}_+ + \mathbf{H}_-$ comparable with the earth's magnetic field. Further, since λ is nearly equal to unity, we have to a very high degree of approximation

$$\mathbf{F} = \frac{[\mathbf{w} \, . \, \overline{\mathbf{H}_+ + \mathbf{H}_-}]}{c}.$$

If then we measure the apparent magnetic field $\mathbf{H}_0$ as the mechanical force divided by $\mathbf{w}_-$ when $\mathbf{w}_-$ is directed so as to make the force a maximum, the quantity we shall obtain for $\mathbf{H}_0$ will be equal to $\overline{\mathbf{H}_+ + \mathbf{H}_-}$.

Note 9. For uniform rotation of a homogeneous body $\partial/\partial t = d/dt$. Moreover, as subsequently chosen in the section "Choice of the Invariants", the values of α and β given by (72) and (73) reduce for uniform rotation to

(102) $$\alpha = \alpha_0 k_+^{-1} \omega^{-1}, \ \beta = \beta_0 k_+ r^2 \omega$$

where α_0 and β_0 are constants. Hence, using (60), (62), (63) and (64), we have, from (51),

(103) $$\frac{-\partial \rho_+}{\partial t} = -2\omega\alpha_0\rho_+ + \alpha_0\omega^{-1} \frac{\partial^2 \rho_+}{\partial t^2} + \beta_0 k^3{}_+ r^2 \omega \frac{\partial^3 \rho_+}{\partial t^3}.$$

Now while not necessary, it constitutes a simplification if we can neglect the last two terms on the right hand side of (103). The procedure is justifiable if the solution for ρ_+ obtained from

(104) $$\frac{-\partial \rho_+}{dt} = -2\omega\alpha_0\rho_+,$$

when substituted in the two terms in question, makes them negligible in comparison with $2\omega\alpha_0\rho_+$. From (104) we obtain

(105) $$\frac{\partial^2 \rho_+}{\partial t^2} = 2\omega\alpha_0 \frac{\partial \rho_+}{\partial t} = 4\omega^2 a_0^2 \rho_+ ; \ \frac{\partial^3 \rho_+}{\partial t^3} = 8\omega^3 a_0^3 \rho_+ .$$

Substituting from (104) and (105) in the terms of (103) under discussion, we find that the ratio of their sum to $2\omega\alpha_0\rho_+$ is $2a_0^2 + 4\beta_0 a_0^2 k_+^3 \omega^3 r^2$. Now the values of α_0 and β_0 given by (70) and (74) respectively make the first term of this expression equal to 2.9×10^{-52} and the second term equal to $1.8 \times 10^{-75} \omega^3 r^2$ so that, even for such speeds as would make ωr equal to the velocity of light the expression only amounts to $16 \times 10^{-55} \omega$, and even for such angular velocities as are associated with atomic phenomena, it is quite negligible. Thus, for all practical purposes equation (103) may be replaced by (104), to which (65) is equivalent for the form of α given by (68).

Now, as regards the contributions of the terms αP_t and βQ_t to the electric field of the rotating body: using (60), (62), (104), (105), and (103), we have

$$\alpha P_t+\beta Q_t=\alpha k_+\frac{d\rho_+}{dt}+\beta k_+\frac{d^2\rho_+}{dt^2}=-2\alpha_0^2\rho_++4\alpha_0^2\beta_0k_+^3\omega^3r^2\rho_+.$$

It is evident from (52) that, at any rate if we calculate the potential in the unretarded sense, $\alpha P_t+\beta Q_t$ acts exactly like a density of electricity ρ_0 given by

$$\rho_0=\alpha P_t+\beta Q_t=-2\alpha_0^2\rho_++4\alpha_0^2\beta_0k_+^3\omega^3r^2\rho_+$$

or, putting $r=R_0\sin\theta$,

$$\rho_0=-2\alpha_0^2\rho_++4\alpha_0^2\beta_0k_+^3\omega^3R_0^2\rho_+\sin^2\theta.$$

It is, of course, easily possible to determine the field due to this distribution throughout the sphere. It is expressible as a series of spherical harmonics. However, we are concerned only with showing that the order of magnitude of the field is small. The order of magnitude of the field will be the same as that calculated by putting $\sin\theta=1$ everywhere; and, of course putting $k_+=1$. In this way we obtain for the field at the surface, a quantity of the order

$$E_0=\frac{8\pi\alpha_0^2Dea}{m}\left(\frac{2\omega^2a^2\beta_0}{5}-\frac{1}{3}\right)$$

where we have replaced ρ_+ by De/m, D being the earth's density, e the charge on the proton, and m the sum of the masses of a proton and an electron. E_0 as expressed above is in electrostatic units if e is in electrostatic units. Using he value of α_0 given by (70) the value of β_0 given by (74) and the data for ω, D,and a given just prior to equation (74) we find, for the earth,

$$E_0=-1.2\times10^{-27}\text{ e.s.u.},$$

a quantity insignificant compared with the earth's potential gradient of about 150 volts per meter for which the theory provides otherwise.

If we had used in (52) the retarded values in the numerator, the values of ρ_0 assigned to any volume element would have differed from the value used above by a quantity $\Delta\rho_0$ which is such that

$$\left|\Delta\rho_0\right|<\left|\frac{2a}{c}\left(\frac{\partial\rho_0}{dt}\right)\right|$$

where $2a/c$ represents the time for light to travel a distance equal to the diameter of the sphere. Thus,

$$\left|\frac{\Delta\rho_0}{\rho_0}\right|<\left|\frac{2a}{c\rho_0}\left(\frac{\partial\rho_0}{dt}\right)\right|$$

or, using (104),

$$\left|\frac{\Delta\rho_0}{\rho_0}\right|<-\frac{4a\omega\alpha_0}{c}.$$

Even for a value of $a\omega$ equal to the velocity of light, $|\Delta\rho_0/\rho_0|$ is only $-4\alpha_0$, *i.e.*, 4.8×10^{-26}. Thus, the effect of using the unretarded in place of the retarded potential is negligible in its influence on the field E_0 calculated above, which field E_0 is itself negligible as we have seen.

Turning now to the contribution of the term $-\frac{1}{c}\frac{\partial}{\partial t}(\mathbf{U}_+ + \mathbf{U}_-)$ in equation (54), it is only because of the slight departure from a steady state caused by the slow death of the positive electricity that this term makes any contribution at all. The magnitude $\delta\mathbf{E}$ of the contribution is certainly less than that which we should calculate if we should assign to each volume element in (53) vectors equal in magnitude to the absolute values of $\alpha\mathbf{P}$ and $\beta\mathbf{Q}$ but in the same direction, and moreover, in the same direction for each volume element. The absolute value which we should calculate for the integral in (53) by using such a process would be 4π times the potential V due to a distribution ρ_1 given by*

$$\rho_1 = |\alpha P| + |\beta Q|$$

and we should thus have

(106) $$\left|\delta E\right| < \left|\frac{1}{c^2}\frac{\partial V}{\partial t}\right|.$$

The use of unretarded values in calculating V does not, of course, affect the order of magnitude.

Using (59), (61), and (102), we have

$$\rho_1 = (\alpha_0 r\omega + \beta_0 k_+^3 r^3\omega^4)\rho_+.$$

The potential at a point due to this distribution is of the same order of magnitude as that obtained by putting $k_+ = 1$, replacing r by $R_0 \sin\theta$, R_0 being the radius vector from the centre of the sphere, and then discarding everywhere the difference between $\sin\theta$ and unity. With these approximations, we find for V at a point on the surface of the sphere of radius a,

$$V = \frac{1}{a}\int_0^a (\alpha_0\omega R_0 + \beta_0\omega^4 R_0^3)\rho_+ R_0^2 dR_0 = \tfrac{1}{4}(a^3\alpha_0\omega + \tfrac{2}{3}a^5\omega^4\beta_0)\rho_+.$$

Thus, using (104) and (106), and putting $\rho_+ = De/m$, we have

$$\left|\delta E\right| < \left|\frac{2\omega^2\alpha_0^2 a^3\pi}{c^2}\left(1 + \tfrac{2}{3}a^2\frac{\beta_0}{\alpha_0}\omega^3\right)\frac{De}{m}\right|,$$

where e is in e.s.u., and we have introduced the factor 4π, so that δE is in e.s.u. Putting in the values appropriate for the earth, we find

$$|\delta E| < 6.2\times10^{-32} \text{ e.s.u.},$$

which is insignificant compared with the earth's potential gradient.

*We shall omit Clarendon type in absolute values.

Note 10. If the whole of the positive electricity in the earth were rotating alone as a solid mass with the earth's angular velocity, the ordinary current density, which would be tangential, would amount at any point to $r\omega\rho_+$. A fraction 5×10^{-19} of this current density would give rise to a field comparable with that of the earth. Now, using (59) and (68),

$$\alpha P = \alpha_0 r\omega\rho_+$$

and, from (70)

$$\alpha P = 1.2\times10^{-26}(r\omega\rho_+) = 2.4\times10^{-8}(r\omega\rho_+\times5\times10^{-19}).$$

Hence the apparent radial current density at any point is only 2.4×10^{-8} of the tangential current density which would apply to that point in a current distribution of density proportional to the distance from the axis of rotation and of amount such as to produce a magnetic field equal to that of the earth.

Again, as regards the retarded feature in (53), following a line of argument similar to that used in Note 9, and using (59), (61), and (104), we have, that the errors in **P** and **Q** resulting from failure to take account of the retarded feature are $\Delta\mathbf{P}$ and $\Delta\mathbf{Q}$, such that

$$\left|\frac{\Delta P}{P}\right| = \left|\frac{\Delta Q}{Q}\right| < 2\frac{a}{c}\,\omega\alpha_0,$$

so that since $a\omega$ must be less than c, each of these quantities is less than $2\alpha_0$ and is therefore exceedingly small. The effect of $\Delta\mathbf{P}$ on the contributions of $\alpha\mathbf{P}$ from each volume element is consequently exceedingly small compared with that contribution itself, which we have already found negligible. The only effect of $\Delta\mathbf{Q}$ is to change our calculation of $\mathbf{U}_+ + \mathbf{U}_-$, and consequently of $\mathbf{H}_+ + \mathbf{H}_-$ to an exceedingly small extent as a result of a very slight change in the apparent current density to be assigned to each volume element of the sphere.

Note 11. In virtue of (33) and (34) the combined force on a unit of positive and a unit of negative charge moving with equal velocity **w** is

$$\delta\mathbf{F} = (1-\lambda)\left(\mathbf{E}_+ - \mathbf{E}_- + \frac{[\mathbf{w}\,.\,\overline{\mathbf{H}_+ - \mathbf{H}_-}]}{c}\right).$$

Here $\mathbf{E}_+ - \mathbf{E}_-$ and $\mathbf{H}_+ - \mathbf{H}_-$ really represent additions of quantities of like signs since $\mathbf{E}_+$, and $\mathbf{H}_+$ represent the field due to all the positive electricity in the piece of matter and $\mathbf{E}_-$ and $\mathbf{H}_-$ represent the field due to the corresponding negative electricity.

Now $\mathbf{E}_+$ and $\mathbf{H}_+$ are very nearly equal to the values calculable on strict classical theory, so that, to a high degree of approximation, for a neutral body,

$$\delta\mathbf{F} = -2\,(1-\lambda)\left\{\mathbf{E}_- + \frac{[\mathbf{w}\mathbf{H}_-]}{c}\right\}.$$

$\mathbf{E}_-$ and $\mathbf{H}_-$ being of course calculable on classical theory. Now it is easy to see, and is well known, that $\mathbf{H}_-$ is of the order $\mathbf{w}/c$, so that, up to and including the order $\mathbf{w}/c$

$$\delta\mathbf{F} = -2(1-\lambda)\mathbf{E}_-.$$

Further, as is well known, $\mathbf{E}_-$ differs from the field calculable on the basis of pure electrostatics only by quantities of the order $\mathbf{w}^2/c^2$. Hence, to a high degree of approximation, the force on an element of matter $d\tau_2$ containing N_2 protons and N_2 electrons per c.c. and due to an element of matter $d\tau_1$ containing N_1 protons and N_1 electrons per c.c. is

$$\delta F = \frac{2N_1N_2e^2(1-\lambda)}{R^2}\, d\tau_1\, d\tau_2$$

where e is in electrostatic units.

THE EFFECT OF SURFACE DRAG ON SURFACE WINDS

By Dr. W. J. Humphreys,
Meteorological Physicist, Weather Bureau, Washington, D.C.

In the application of mathematics to concrete problems we take our equations wherever we find them and turn them to our immediate needs. If, fortunately, we know of some other problem, however different the subject matter, already so solved that by merely changing the meanings of the symbols and the values of the constants we thereby resolve our own, then all the better. The usefulness of this transfer *en masse* of a system of equations from one field to another is well illustrated in the following, in which Ekman's* classical paper on ocean drift, by proper change of interpretation, becomes an equally good resolution of difficult questions, directly solved by Taylor† and others, concerning the effect of turbulence and surface drag on the winds of the lower atmosphere.

Ekman assumed a wind of constant strength and direction over an initially quiet body of water of great extent and considerable depth, and found what the resulting movements of the water would be on the attainment of a steady state. On the other hand, we shall assume a broad, straightaway current of water, or drift of prairie sod, say, under an initially quiet atmosphere, and try to find the resulting steady-state movements of the lower air. Clearly, the two problems, the effect of wind on water drift, and of a water current on air drift, are identical, except as to the numerical values of certain constants. Hence the equations used by Ekman in the solution of the ocean drift problem may be taken over without change for the solution of the similar wind problem. Clearly, too, the final results will be directly applicable to all steady winds, whether over land or over water, for a wind is only motion as between earth and air.

Let the origin of the rectangular coordinates x, y, and z be on the surface (land or water) of the earth, z being vertical and positive upwards, and the positive direction of y 90° counterclockwise from the positive direction of x, as seen from above. Let

u, v = the velocity components of the air drift in the directions of x, y;

X, Y = the x, y components of extraneous forces per unit mass of the air, assumed constant with height;

ρ = the density of the air, regarded constant, which it is, roughly, through at least the first half kilometer;

On the influence of the earth's rotation on ocean currents, Arkiv. för Mat., Astr. o. Fys., 1905.

†Phil. Trans., A215, 1-26, 1914.

μ = the coefficient of drag, or eddy viscosity, of the air, considered constant through the eddy layer—an allowable first approximation;

t = the time.

Then, in the northern hemisphere (or southern, with proper change of signs),

$$X = 2v\omega \sin \phi; \quad Y = -2u\omega \sin \phi;$$

in which ω is the angular velocity of the earth's rotation, and ϕ the latitude.

Hence, from the equations of motion of a viscous fluid,* remembering that, by assumption, u and v can vary only with time and height,

$$\frac{\partial u}{\partial t} = 2v\omega \sin \phi + \frac{\mu}{\rho} \frac{\partial^2 u}{\partial z^2}$$

$$\frac{\partial v}{\partial t} = -2u\omega \sin \phi + \frac{\mu}{\rho} \frac{\partial^2 v}{\partial z^2}.$$

For steady motion, and on writing

$$\rho\omega \sin \phi/\mu = a^2,$$

these equations reduce to

$$\frac{\partial^2 u}{\partial z^2} + 2a^2 v = 0,$$

$$\frac{\partial^2 v}{\partial z^2} - 2a^2 u = 0,$$

the general solution of which is

$$u = C_1 e^{az} \cos (az + k_1) + C_2 e^{-az} \cos (az + k_2)$$

$$v = C_1 e^{az} \sin (az + k_1) - C_2 e^{-az} \sin (az + k_2)$$

in which C_1, C_2, k_1, k_2, are arbitrary constants.

Clearly, the drag is zero, hence u and v are zero, when z is infinite. Indeed the drag is approximately zero at half a kilometer, or thereabouts, above the surface, since at that level the speed and direction of the wind are practically what they would be under the existing horizontal pressure gradient if the atmosphere were a perfect fluid.

If, then, u and v are each to be zero when z is infinite, obviously C_1 is zero. Hence

$$u = C_2 e^{-az} \cos (az + k_2)$$

$$v = -C_2 e^{-az} \sin (az + k_2)$$

*See, *e.g.*, Lamb: *Hydrodynamics*, 5 ed., pp. 546-547.

and
$$\frac{du}{dz} = -aC_2e^{-az}\{\sin(az+k_2)+\cos(az+k_2)\}$$
$$= -a\sqrt{2}\,C_2e^{-az}\sin(az+k_2+45°)$$
$$\frac{dv}{dz} = aC_2e^{-az}\{\sin(az+k_2)-\cos(az+k_2)\}$$
$$= -a\sqrt{2}\,C_2e^{-az}\cos(az+k_2+45°).$$

If the tangential drag of the surface on the air is T, and in the positive direction of y, then
$$\mu\left(\frac{du}{dz}\right)_{z=0}=0, \quad -\mu\left(\frac{dv}{dz}\right)_{z=0}=T.$$

Hence
$$k_2=-45°$$
$$C_2=\sqrt{u_0^2+v_0^2}=V_0$$

(1)
$$\left.\begin{array}{l} u=V_0e^{-az}\cos(45°-az) \\ v=V_0e^{-az}\sin(45°-az) \end{array}\right\}$$
$$V_0=T/\mu a\sqrt{2}=T/\sqrt{2\mu\rho\omega\sin\phi}\,.$$

Since the tangential drag T of the water (or land) on the air is, by assumption, in the positive direction of y, and of necessity in the direction of the surface movement with reference to the air, it follows from the equations (1) that, in the northern hemisphere, the drift of the air just above the surface ($z=0$, components u and v of V_0 equal to each other and positive) is 45° to the right of the direction of the surface current relative to the air. In the southern hemisphere the air drift is 45° to the left of the surface current.

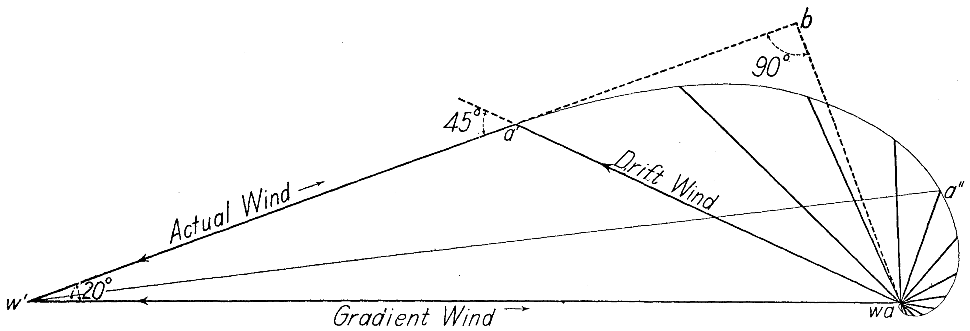


Fig. 1

Since the value of a is independent of height, z may be counted from any level, provided T is the drag between the superjacent layers of air at that level, and V_0 the corresponding drift velocity. Hence the projection of the drift envelop on to the ground or water surface is an equiangular, 45° or 135°, spiral about the initial contact point of air and surface.

The relations between the air and surface movements are shown in Fig. 1 (as over water—over land the angle i is around 30°), in which wa is the contact position of certain adjacent air and surface particles at a given instant, w' the position of the same surface particle at a given subsequent instant, and a' the position, at the same subsequent instant, of the initial air particle. That is, with reference to the earth beneath, ww' is the direction and speed of the surface and aa' the consequent direction and speed of the air. Hence $a'w'$, inclined 135° to aa', is the direction and speed of the surface with reference to the air, and, of course, $w'a'$ that of the air with reference to the surface. That is, $w'a'$ is the direction and speed of the surface wind.

It should be noted in this connection that since the angle $w'a'w$ between the directions of the surface wind and the "drift wind" has to be 135°, therefore the speed of the surface wind decreases as its inclination to the isobar increases, becoming zero as this angle reaches its limit, 45°.

From the equations (1) it is further obvious that the angle between the directions of the driving surface and the resulting air drift increases uniformly with height, two right angles for each π/a gain therein; and also that the velocity of the drift continuously decreases, as the height grows greater, falling to the $e^{-\pi}$th, or, approximately, the 1/23 part of its initial value for any π/a ascent.

The speeds and directions of the drift currents at the heights 0, $\pi/10a$, $2\pi/10a$, and so on, of which wa'' is an example, are shown in Fig. 1, projected onto the surface, or xy plane. The direction and velocity of this actual wind at the height corresponding to the drift wa'' are given by $w'a''$ and similarly for other heights.

If we assume the turbulence drag to be substantially zero at that level at which the wind, as we go up, attains the direction of the isobars, that is, where the angle of inclination i, Fig. 1, becomes zero, and call this height H, then, clearly

$$H = 3\pi/4a = \tfrac{3}{4}\pi\sqrt{\frac{\mu}{\rho\omega \sin \phi}}. \tag{2}$$

To find the total momentum given to the air per unit of time by the drag T, let F_x be the total flow in the positive direction of x across a strip of infinite height and unit width at right angles to x, and F_y the similar total flow in the direction of y. Then

$$F_x = \int_0^\infty u dz = V_0 \int_0^\infty e^{-az} \cos(45° - az) dz = \frac{V_0}{a\sqrt{2}} = \frac{T}{2\rho\omega \sin \phi},$$

$$F_y = \int_0^\infty v dz = V_0 \int_0^\infty e^{-az} \sin (45° - az) dz = 0.$$

Hence the integrated momentum of the drift wind is 90° to the right of the direction of the surface flow with reference to the air, and its value per second per strip of unit width at right angles to this direction and infinite height (practically, to the level of gradient direction),

$$\rho F_x = \frac{T}{2\omega \sin \phi}.$$

Since the above discussion applies to any velocity, between air and earth in any direction, let, as we may, $w'a'$, Fig. 1, be the direction and velocity of the surface wind, and let i be the inclination angle, or angle between the direction of the "surface wind" (wind at bottom of region of constant eddy viscosity, say 10 meters above the surface) and the direction of the isobar, or direction of the gradient wind. Then, on the assumption that the viscosity is constant up to the gradient-direction level, and from there on zero, the gradient velocity is given by $w'a$, determined by drawing $a'a$ so as to make the angle $w'a'a$ 135°.

From Fig. 1, it also is evident that the surface wind W_s is given, in terms of the gradient wind W_g by the equation

$$W_s = W_g(\cos i - \sin i),$$

and the drift wind, W_d, by the equation

$$W_d = W_g \sin i \cdot \sqrt{2}.$$

The figure also shows that, theoretically, the wind attains gradient velocity both below and above the level of gradient direction, and that at this level the velocity is distinctly in excess of the gradient value. All these surprising deductions have actually been observed.

Since H, the height of the gradient-direction wind, is given by observation, we have, from (2), in known terms

$$\mu = 16H^2\rho\omega \sin \phi / 9\pi^2.$$

Clearly, then, the direction and velocity of the wind at different heights up to the level of gradient direction furnishes, as indicated by Fig. 1, a means of determining whether or not μ is essentially constant. This has been done, and the assumption that normally μ is substantially constant with height fully sustained,* however much it may vary with the roughness of the surface and velocity of the wind.

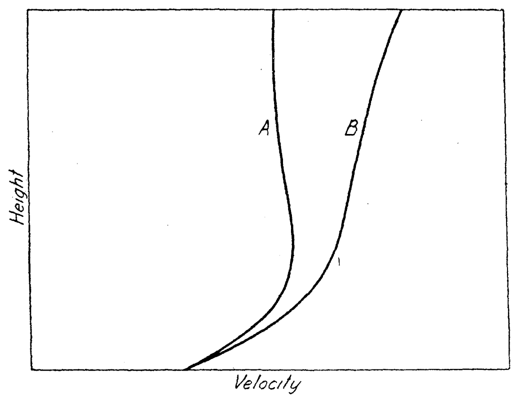


Fig. 2

If the driving forces X, Y, per unit mass, hence the gradient wind, are the same at all levels, as above assumed, then the speed of the actual wind varies with height substantially as indicated by curve A, summer type, Fig. 2. If,

*Taylor: Phil. Trans., A215, p. 1, 1915; Whipple: Quart. Jour. Roy. Met. Soc., 46, 39, 1920.

however, these forces increase linearly with height, this speed will vary as indicated by curve B, winter type, Fig. 2.

The averages of the wind velocities at various heights in the lower atmosphere, from observations made daily, or nearly so, over a period of several months, give a curve of the same general type as the theoretical ones shown in Fig. 2. This is well illustrated by Fig. 3, a typical example from many based

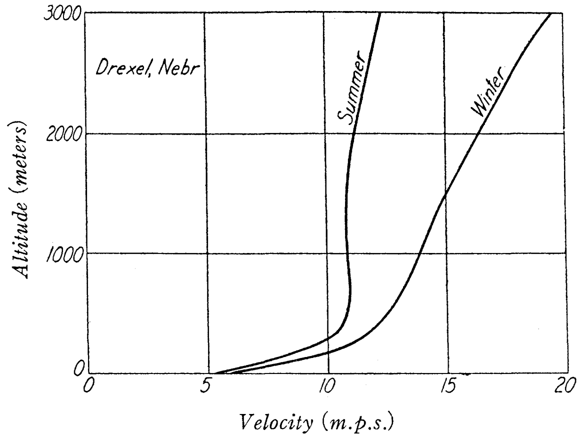


Fig. 3

on observations covering several years at Drexel, Nebraska, and kindly furnished by W. R. Gregg, U.S. Weather Bureau. Clearly, then, so far as change of velocity is concerned, the assumptions on which the theoretical analysis is based seem to be substantially correct. And the same is true of the wind direction, since, as long known, the angle between this direction and that of the current isobar decreases with increase of elevation, to approximately zero at half a kilometer, or thereabouts, above the surface.

Addendum, August 1925

After mathematics has shown us what the correct conclusion is, it often happens that we then can arrive at the same result in a less formal way—by verbal reasoning instead of symbolic logic. Such is the case in the present instance, as we shall show by considering the action of a steady wind on initially still water. This air-on-water action is chosen instead of the precisely similar converse, because it is easier, perhaps, to visualize.

Let, then, aa', Fig. 4, represent tbe steady direction and velocity o the surface air. The water, being free to flow, and on a rotating sphere, will deflect from the wind direction— to the right in the northern hemisphere. Let ww' be the direction and velocity of the surface water. Numerically, at latitude ϕ

$$ww' = 7.9 \times 10^{-3} aa' / \sqrt{\sin \phi}$$

nearly.* Clearly, then, $w'a'$ is the direction and velocity of the wind with

*Durst: Quart. Jour. Roy. Met. Soc., 50, p. 115, 1924.

reference to the moving surface of the water—but little different from aa', its geographic direction and velocity. Hence the drag of the air on the water, T per unit surface area, is in the direction $w'a'$. But the water, being free to move, automatically adjusts its motion so as to balance this force by the corresponding force incident to earth rotation. That is, in each vertical column of water of unit cross-section there is a momentum M at right angles to $w'a'$ of such magnitude that

$$T = 2M\omega \sin \phi$$

Now, the connecting medium between the wind above and the moving water beneath the surface layer is that surface layer, however thin. Hence the reaction on this surface by the momentum M at right angles to $w'a'$ is the same, this being the resultant of all the subsurface momenta, as would obtain if the neutralizing components parallel to $w'a'$ did not exist. Hence the drag of the

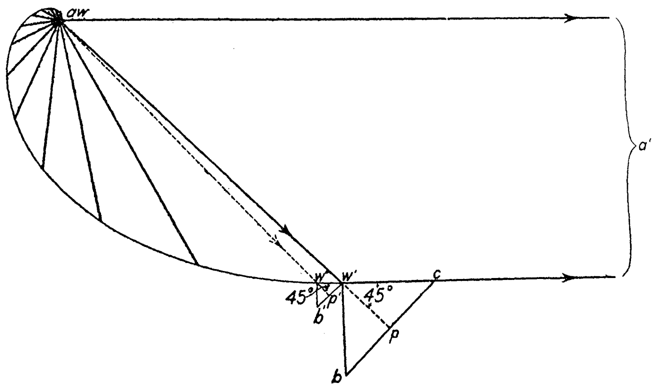


Fig. 4

under water on the surface is in the direction $w'b$ normal to $w'a'$. Furthermore, the component of this water drag normal to the direction of surface flow is equal and opposite to that of the wind drag, there being no other horizontal forces operating on it. If, therefore, wp is the direction of surface flow and if pc, normal thereto, represents the component of T in that direction, then pb, equal and opposite to pc, similarly represents the component of the water drag on the surface at right angles to the direction of flow.

From the similarity of triangles and the fact that $bw'a'$ is a right angle, it follows that $pw'a'$ is half a right angle. Hence, under the conditions assumed, as to steady winds and initially still deep water, the final direction of surface drift makes an angle of 45° to the direction of the wind with reference to the water surface—to the right in the northern hemisphere, to the left in the southern.

Consider next the water as made up of separate layers, each moving as a solid and connected to its neighbours by interfacial viscosity, and let ww'' be the direction and velocity of the second layer. Clearly, then $w''w'$ is the direction and velocity of the first or surface layer with respect to the second. Also, exactly as before, we find that the drift of this second layer is 45° to the right (northern hemisphere) of the direction of the first layer with respect to the second.

In this manner any depth may be reached, however thin the separate layers. Hence the projection of the envelope of the radii vectors of the water drifts onto the surface is a 45° (or 135°) equiangular spiral about any common starting point of wind and water.

Since the above directional relations are independent of velocity, it follows that the amount of spiralling is directly proportional to depth, provided the viscosity, here owing to turbulence convection, is constant. Hence, from the direction and velocity of the surface drift and of the drift at any known depth the movement of the water at any other depth may be determined graphically. Also from these observed drifts the viscosity readily can be determined.

Exactly similar conclusions, arrived at in precisely the same way, apply to the atmosphere, as already explained.

THE CONVECTIVE ENERGY OF SATURATED AIR IN A NATURAL ENVIRONMENT

By Sir Napier Shaw,
Professor of Meteorology, Imperial College of Science and Technology, London, England.

Natural environment in respect of the convectional ascent of air at any epoch is specified by the graph of temperature in relation to pressure obtained from a balloon-sounding at that epoch. It can be set out as a θ, ϕ diagram, *i.e.*, a diagram with temperature as abscissa and entropy as ordinate, by transforming the graph of pressure and temperature into one with temperature, θ, on a linear scale and potential temperature, θ_0, on a logarithmic scale; because the relation between entropy ϕ and potential temperature θ_0 is

$$\phi = c_p \times \log \theta_0 + \text{const.}$$

Fig. 1 shows some examples of such entropy-temperature curves for South East England; the normals for summer and for winter and the conditions on July 5 and 6, 1923. The dimensions are adjusted for 1 kilogramme of air. Note that the curve of environment gives no information about humidity. A separate graph is required for that. On such a diagram area represents energy in thermal units or equivalent dynamical units. One square centimetre on Fig. 1 represents approximately 800 joules per kilogramme. Note also that the isothermal condition is represented by a vertical line, and the dry adiabatic condition by a horizontal line.

Figs. 2 and 3 represent the initial and intermediate stages, Fig. 4 the final stage of the transformation of a graph from the pressure, temperature curve to the θ, ϕ curve. The *process* is, of course, the same for any curves whatever, drawn on the p, θ paper; the curves selected for illustration of the transformation are the family of saturation adiabatics for air which passes from high pressure to low pressure without communication of heat, and from which the condensed water falls away. They have been called *pseudo-adiabatics* but *irreversible adiabatics* is a better name.

Fig. 2 represents the adiabatics for saturated air on a p, θ, diagram with logarithmic scales. On this diagram an adiabatic for dry air is a straight line parallel to that drawn from the left-hand top corner to 330 on the bottom line. Fig. 3 shows the family of curves transformed into a $p_1\theta_0$ diagram (pressure-potential temperature). The transformation can be effected by setting out the dry adiabatic line from the point at standard pressure (1000 mb) and from the vertical line of potential temperature passing through that point, setting off at any pressure a point as much to the right-hand side as the point at the same pressure on the p, θ curve is to the right of the dry adiabatic. From the two diagrams Fig. 4 is constructed by setting off the temperatures corresponding to successive pressures as abscissae on the horizontal linear scale of the final diagram, and the corresponding potential-temperatures as ordinates on the logarithmic scale. In all these diagrams the figures are adjusted for a mass which contains one kilogramme of dry air.

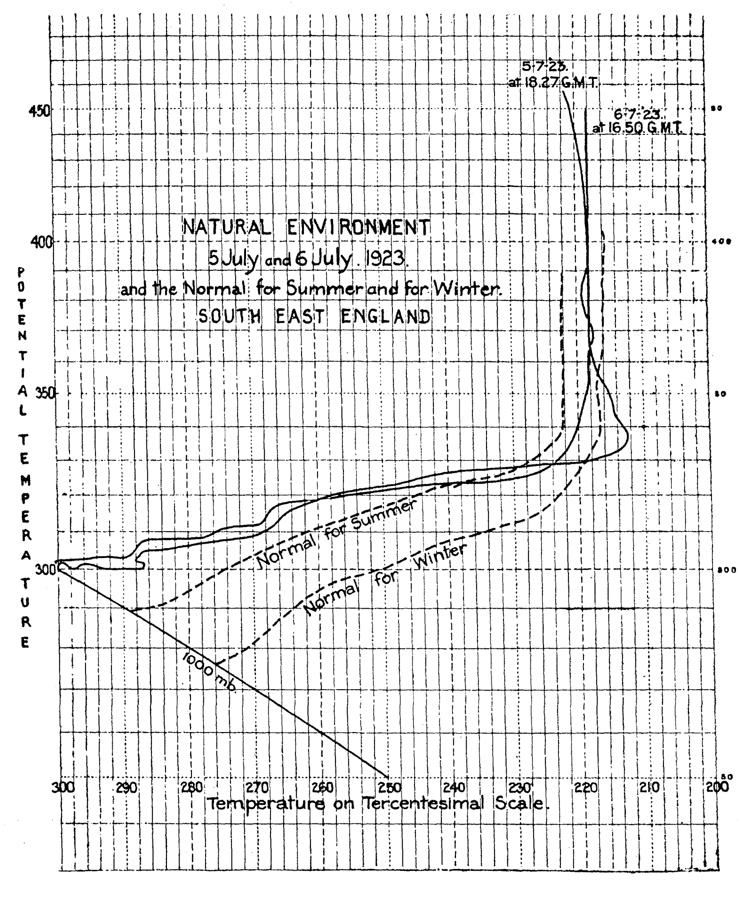


FIG. 1—Curves of natural environment referred to temperature and entropy (logarithm of potential temperature) as coordinates (1) normal curve for summer in South East England, (2) normal curve for winter, (3) and (4) curves for Benson (Oxford) July 5 and July 6, preceding the great London thunderstorm of July 9, 1923.

Fig. 5 represents the state of the environment (with two saturation adiabatics as guides) on two occasions in England, July 22 and November 25, 1922, and we note the following:

Dry air at any point of a curve of natural environment in equilibrium has no convective energy. Even in the extreme case of an environment which is

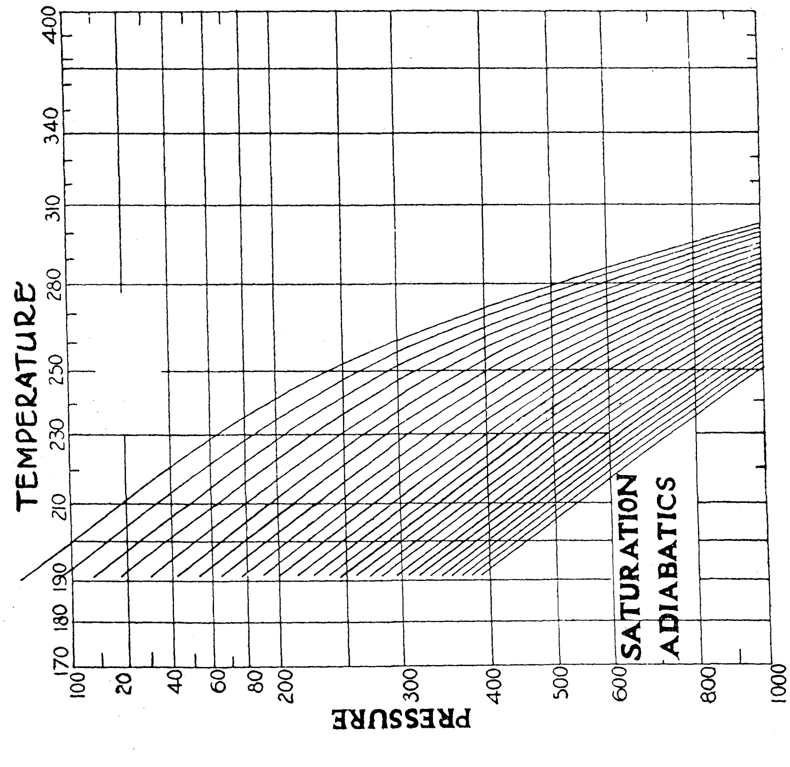


Fig. 2—Irreversible adiabatics for saturated air referred to pressure and temperature, both on logarithmic scales. The adiabatics for dry air are straight lines parallel to that drawn from the left hand top corner to 330 on the bottom line.

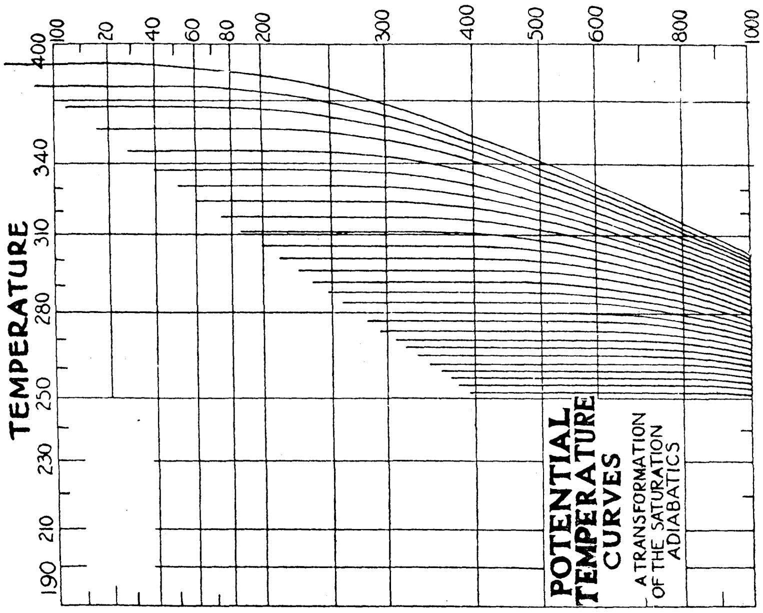


Fig. 3—Irreversible adiabatics for saturated air referred to pressure and potential temperature (with 1000 mb as standard pressure) both on logarithmic scales.

itself in convective equilibrium, dry air at any point is only part of a labile layer. It develops no energy in rising an infinitesimal distance if its temperature is infinitesimally increased above that of its environment.

On the other hand air which is saturated at any point of a curve of environment may or may not have convectional energy. It has convectional energy if the saturation adiabatic passing to lower pressure lies on the warmer side, that is to say on the left-hand side, of the curve of environment. In that case it must sooner or later cross the graph of environment at some second point and the area enclosed between the two graphs expresses the convectional energy of a saturated kilogramme of air at the first crossing point. The second point

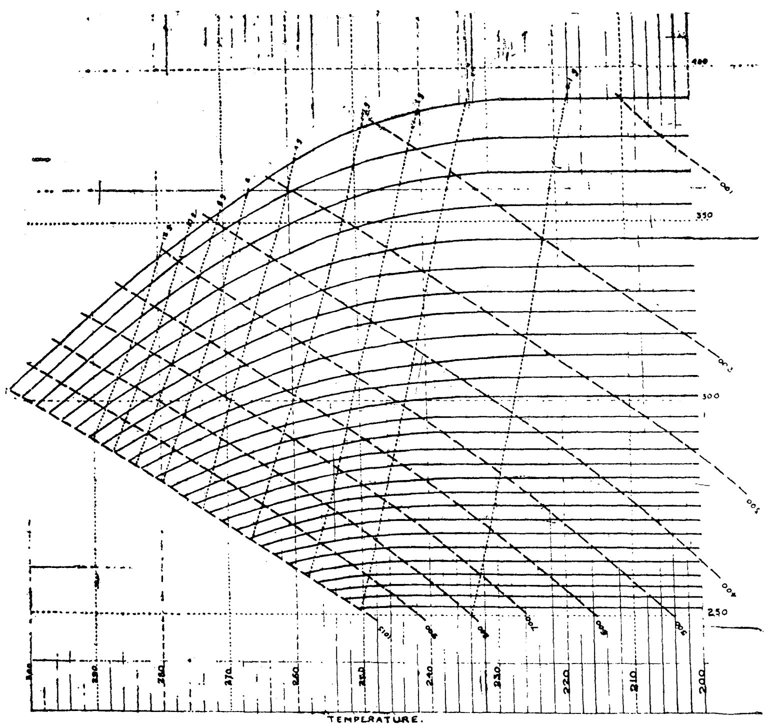


FIG. 4—Indicator diagram for dry and saturated air composed of (1) Isothermal lines (vertical), (2) Adiabatic lines for dry air (horizontal), (3) Adiabatic lines for saturated air referred to temperature, linear scale, and entropy, linear scale, or potential temperature, logarithmic scale.

indicates the position at which the saturated air, after passing through the pressure-changes, would find itself in equilibrium, provided of course that the energy which it possessed at the start had been disposed of somehow or other *en route.* If it carried its energy with it as kinetic energy it would be in vigorous motion when it arrived in the position of equilibrium.

The energy thus represented in the cases examined is found to range between zero and about 10 kilowatt-seconds, 13 horse-power seconds, or more, for every kilogramme of saturated air, which goes up. In Fig. 5 there is no convectional energy in the case of the lower curve, November 25, 1922. For July 22, 1922, the energy amounts to about .7 kilowatt second, 1 h.p. second. The auxiliary lines on the diagram show the change of pressure and the precipitation automatically associated with the ascent.

Fig. 6 shows two corresponding curves for Canadian ascents. The upper curve of environment is for July 5, 1911, taken during the hottest spell of weather to be found in the records of Toronto. I have not actually measured the area

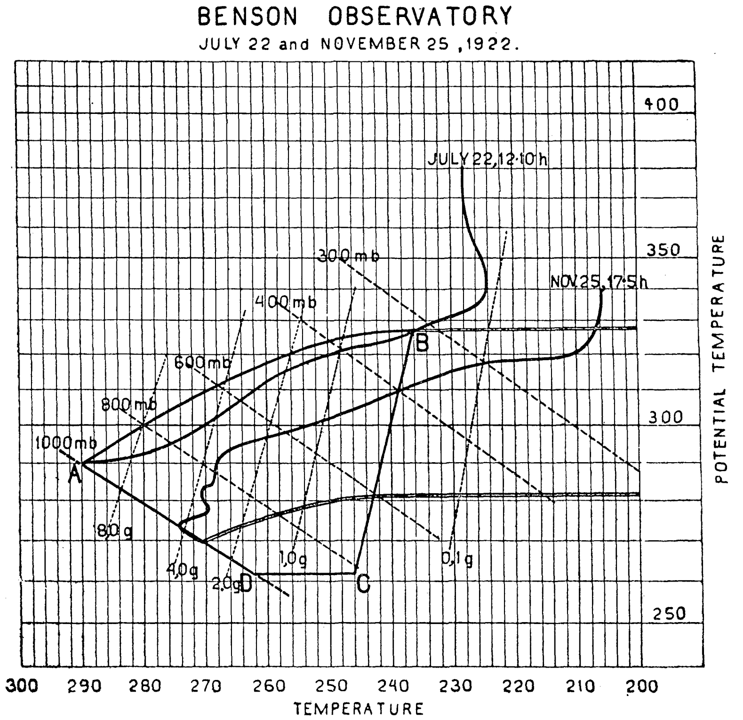


FIG. 5—Soundings by balloon at Benson (Oxford) on July 22 and November 25, 1922, set out in the form of the indicator diagram of Fig. 4 with adiabatics for saturated air, drawn from the points at ground level as auxiliary curves.

which would represent the energy of a kilogramme of air at the base if it were saturated, but guess it at about 12 kilowatt seconds. The second curve, that for February 4, 1914, is most interesting; there is a little pocket of convectional energy for dry air in the lowest layers up to about 1500 metres. I think it must have acquired that energy from the warming effect of the lakes acting upon a cold surface layer at that time of year; above the limit of 1500 metres the air is obviously very stable. Probably the condition, before the lakes changed it, would be a continuation of the curve from the 1500 metre-level to the ground.

Reverting again to the curves of Fig. 1, those for July 5 and 6, 1923, show convective energy of about 10 kilowatt seconds. They preceded the thunderstorm of July 9, 1923, perhaps the worst ever experienced in London. It is a fair conclusion that the commotion of the thunderstorm was the expression of the energy which was represented on the diagram. It was approximately the same as that which the air over Toronto would have had if saturated on July 5, 1911.

We now remember that the 10 kilowatt seconds, or 13 horse-power seconds, is derived from the saturation of 1 kilogramme of dry air. That is a very small unit for meteorological purposes. Let us think of a cubic kilometre of saturated air, a much more suitable meteorological unit; that would be 15×10^9 horse-power seconds, or, say, four million horse-power hours. Or, if we suppose saturated air a kilometre thick to be drawn from an area 100 kilometres square, we get 15×10^{13} horse-power seconds, or 40,000,000,000 horse-power-hours.

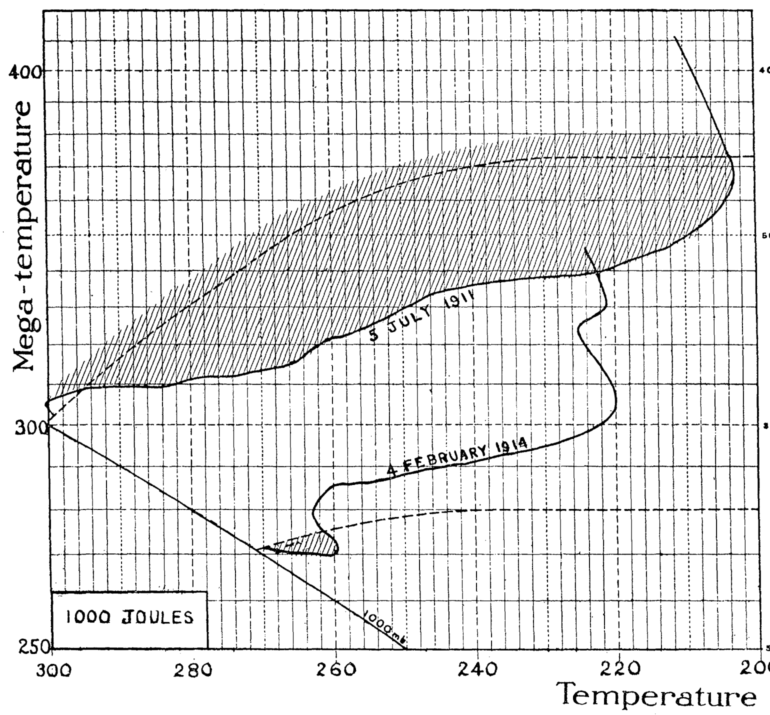


FIG. 6—Soundings by balloon at Woodstock, Ont., on July 5, 1911, and February 4, 1914, referred to temperature and entropy as coordinates.

The area of the rectangle on the diagram marked 1000 Joules represents the development of 1000 Joules (10^{10} ergs) of energy by a mass containing one kilogramme of dry air which passes clockwise through the changes indicated by the perimeter of the area.

The stippled area represents the surplus energy of a kilogramme of air saturated at the lowest point in the environment represented by the graph of the soundings.

The name *mega-temperature* is used in the diagram for potential temperature referred to a standard pressure of 1000 millibars.

The potential temperature is expressed on the "tercentesimal scale", *i.e.*, in centigrade degrees from a zero 273° below the freezing point of water.

The energy of a vortex with a ring of maximum velocity of 20 metres per second, 100 kilometres in radius, a height of one kilometre, a core revolving like a solid and velocity falling off from the maximum in accordance with the law of equal moment of momentum or vr constant, works out, I believe, at the same order of magnitude as that of 2000 cubic kilometres of saturated air in the conditions of July 22, 1922. Whether the two could ever be related as cause and effect I cannot now stop to inquire.

There are many other interesting features of these curves of which other specimens are available, notably the existence of stretches of air in convective equilibrium from the ground upwards for observations from soundings in the early afternoon, and inversions at the surface in the early morning; patches of air in conditions approximating to convective equilibrium in the upper levels and inversions at various heights—all these are shown more clearly upon the θ, ϕ diagram than in other forms of representation.

A transformation into a p, v diagram is easily carried out; but it involves some assumption about the humidity of the air at each point, because the state of the air as regards humidity affects the value of v at any pressure. In the θ, ϕ diagram the assumption appears in the selection of a value for R in the characteristic equation $p = R\rho\theta$ but the changes in R with changes of humidity are so small, that the assumption of no change may be allowed as making no important difference in the results of computation. Moreover a p, v diagram which ranges from a pressure of 1000 millibars to 100 millibars is very elongated and very inconvenient in shape. The crossing of the adiabatic and isothermal lines at right angles in the θ, ϕ diagram is a great advantage in practice.

The θ, ϕ diagram constructed as a diagram of temperature, potential temperature, the latter being on a logarithmic scale, is accordingly recommended for the study of the physical processes of the atmosphere.

THE INFLUENCE OF ELECTROMAGNETIC INDUCTION WITHIN THE EARTH UPON TERRESTRIAL MAGNETIC STORMS

BY PROFESSOR S. CHAPMAN,
Imperial College of Science and Technology, London, England,

AND

MR. T. T. WHITEHEAD,
Mathematical Master, Burnley Grammar School, Burnley, England.

INTRODUCTION

1. A tentative theory of magnetic storms was outlined in a former paper[1] by one of us, on the basis of an investigation of the principal average features of magnetic storms in the region between latitudes 60° north and south. Streams of electric corpuscles from the sun, having charges mainly or entirely of like sign, were regarded as the primary cause. Professor F. A. Lindemann[2] pointed out certain difficulties encountered in subsequent numerical developments[3] of the theory, and suggested that they might be avoided, without otherwise modifying the theory, by supposing the streams of corpuscles from the sun to be ionized but electrically neutral. This proposal must itself be revised, however, since it has been shown[4] by one of us that a neutral ionized stream would not impinge on the dark hemisphere of the earth in the way that the theory requires, and as actually occurs, as shown by auroral displays. The stream must have a slight excess charge of one sign, and the requirements of auroral theory[5] would be met[6] if the unbalanced excess amounts to one unit charge per hundred or more equal and opposite charge; the stream must be ionized in order that the magnetic force on the excess charge may be communicated to the whole mass of the stream. Professor E. A. Milne has recently shown how the sun can and must emit ionized corpuscles[7].

2. A subsequent study[8] of magnetic disturbance phenomena, in high as well as in middle and low latitudes, indicates that the above theory of magnetic storms requires modification in further respects. In particular, it is concluded that magnetic disturbance is directly produced mainly, if not entirely, in polar regions, where also aurorae occur. The polar magnetic variations clearly indicate the presence, during magnetic disturbance, of intense electric currents along a narrow zone in the upper atmosphere, which ordinarily coincides nearly with the zone of maximum auroral frequency as determined by Fritz. This is an addition to the many evidences of connection between aurorae and magnetic disturbance. It may also be regarded as giving general confirmation to the conclusion arrived at theoretically by Professor C. Størmer[5] that charged

NOTE—The references are to the bibliography at the end of the paper.

streams from the sun reach the earth only along narrow zones centred at the north and south poles of the axis of the earth's magnetization.

On the other hand, not only are magnetic storms not confined to the regions of visible aurorae, but some of their most striking effects in middle and low latitudes are greater at the equator than in latitudes nearer the auroral zones, *e.g.*, 40° or 50°. This is specially the case with regard to the diminution of horizontal magnetic force, which in the middle belt of the earth is so prominent a feature of magnetic disturbance. The facts may, however, be consistent with the view that the currents whose magnetic disturbing effect is observed during storms originate directly only in polar regions; one possibility is that some of the polar electric currents find return paths in conducting regions of the atmosphere in middle and low latitudes Another possibility is that the varying field of the high-latitude currents *induces* secondary currents in lower latitudes, both in the atmosphere and in the earth. The object of the present paper is to indicate how such questions of induction can be investigated analytically and numerically. A particularly simple case is treated numerically in detail, and the results confirm the view that primary electric currents in polar regions, whose direct magnetic field in low latitudes is inappreciable, can induce secondary currents which have important magnetic effects in low latitudes. They therefore render it unnecessary to suppose, as one of us formerly did,[1] that the solar corpuscles impinge on the earth in low latitudes (where aurorae are seldom or never observed) as well as in high latitudes.

The problems of induction which we consider are very limited in scope. They relate only to the part of the disturbing field which is symmetrical about the earth's axis, and therefore corresponds to the "storm-time" variations of the field[1]. This is the more important part of the field in middle and low latitudes, though not in high latitudes. The corresponding currents, whether primary or secondary, will flow entirely along parallels of latitude. We treat the earth as symmetrical about a common geographical and magnetic axis, and regard the primary E.M.F. (electromotive force) as existing only in two narrow zones, one round each pole. In this case it is clear, from considerations of symmetry, that no mere spreading of the currents out of these zones into higher or lower latitudes is possible. The investigation is therefore necessarily confined to the second possibility mentioned above. It is hoped to discuss later the more general non-axially-symmetrical case (including the local-time or diurnal variations during magnetic storms) in which both possibilities exist.

The numerical calculations of this paper, though so limited, render possible a simplified theory of magnetic storms, in which the primary actions are confined to the polar regions, where alone we have direct evidence of the entry of corpuscles from the sun. During intense storms, of course, the region of auroral occurrence is much more widespread than at ordinary times, and the region in which magnetic disturbance has its primary seat will be correspondingly greater. The numerical calculations of this paper refer to relatively weak magnetic disturbance.

Although this paper bears closely on the theory of magnetic storms, no revised general theory is here propounded. The existence of E.M.F.'s in the auroral zones is assumed without explanation: since the currents exist so also must the E.M.F.'s. Our object is merely to solve certain definite problems of electromagnetic induction which are considered to bear on terrestrial magnetic phenomena, and whose solution may afford precise knowledge on which a revised theory can be partly based.

Statement of the Mathematical Problems Considered

3. We consider a uniform sphere of radius a which, within a concentric "core" of radius $qa (q<1)$ is of uniform electrical conductivity k, while the layer around this core is non-conducting. The magnetic permeability throughout the whole sphere is taken to be unity. Outside, and concentric with, this sphere there is a thin uniformly conducting shell which, for mathematical convenience, is treated as of infinitesimal thickness, though the product (K') of the thickness into the specific electrical conductivity is finite. The radius of this shell will be denoted by a' or $q'a$ $(q'>1)$. Outside the shell there are two conducting circular rings, with their planes parallel and their centres on that diameter of the sphere (hereafter called the axis of the sphere) which is normal to their planes. For definiteness these rings are treated as anchor rings, of small circular section of radius b_0. The distance of the central circles of these rings from the centre of the sphere will be denoted by a_0 or q_0a $(q_0>1)$. These central circles are north and south of the equatorial plane of the sphere in equal colatitudes (θ_0); their radius r_0 in their own planes will therefore be $a_0 \sin \theta_0$.

The specific conductivity of the material of the two rings will be denoted by k_0, so that the total conductivity (K_0) across a section will equal $\pi b_0^2 k_0$, and the total "linear" resistance (ρ_0) of either ring will be $2\pi r_0/K_0$. The self induction L_0 of either ring is approximately given by

$$L_0 = \pi a_0 \sin \theta_0 \{4 \log_e (8a_0 \sin \theta_0/b_0) - 7\},$$

and the mutual induction M_0 of the two rings (for this purpose treated as linear) is approximately given by

$$M_0 = 4\pi a_0 \sin^2 \theta_0 \int_0^{\frac{1}{2}\pi} \frac{2 \sin^2 \phi - 1}{(1 - \sin^2 \theta_0 \sin^2 \phi)^{\frac{1}{2}}} d\phi$$

$$= 4\pi a_0 \sin^2 \theta_0 \,.\, \tfrac{1}{2}\pi \left\{ \frac{1}{8} \sin^2 \theta_0 + \frac{3}{32} \sin^4 \theta_0 + \frac{75}{1024} \sin^6 \theta_0 + \dots \right\}.$$

It is supposed that up to time $t=0$ no currents flow in any of the conducting regions. At that instant an E.M.F. equal to E_0, a constant, is applied, in the same direction round each of the two zones. Equal currents will be set up in each, rising finally to a steady value E_0/ρ_0. During this process secondary currents will be induced in the shell and conducting core; these currents will be transient, finally dying away. Our object is to determine the distribution and relative magnitudes of the magnetic fields of the primary currents (in the zones) and of the secondary currents, for points on the surface of the sphere.

All the currents will flow along circles of latitude, and the magnetic force will everywhere lie in meridian planes through the axis of the model. One pole of the axis will be called north and the other south. The tangential component of the magnetic force at the surface of the sphere, reckoned positive toward the north pole, will be called the horizontal magnetic force, and denoted by H_n. The radial component, reckoned positive inward along the radius, will be called the vertical magnetic force, and denoted by H_v. The magnetic fields will be discussed in terms of these two components.

In the actual numerical application of the analysis, the conducting shell is supposed absent, *i.e.*, $K'=0$. It is hoped later to discuss the complete problem numerically, taking more than one illustrative set of initial numerical data.

The model just described presents, in an idealized form, the leading features of the terrestrial case which are relevant to the magnetic problems discussed in the Introduction. The study of the diurnal magnetic variations[9] has shown that the interior of the earth, below a depth of about 160 miles or 250 km., is much more conducting than the surrounding layer, and has indicated a probable value of k. Thus, for the earth, we take

$$a=6.37.10^8 \text{ cm.} \quad q=0.96 \quad k=3.65.10^{-13} \text{ e.m.u.}$$

Above the earth, at a certain height in the atmosphere, there is a conducting layer, for which the average value of K' is roughly known, from the study of the diurnal magnetic variations. The conductivity is, however, by no means uniform, being greater by day than by night, and in low latitudes than in high. In a first essay on a difficult problem it is convenient to regard the shell as of uniform conductivity, equal to the average value of K'. Also assuming a mean height of the layer of about 40 miles or 65 km., the appropriate data for the shell may be taken as:

$$K'=1.25.10^{-5} \text{ e.m.u.} \qquad q'=1.01.$$

The two conducting zones or anchor rings are an idealized representation of the auroral regions of high conductivity, existing above a height of 80 or 100 km. Any numerical values for the size and conductivity of these zones must be tentative, particularly since the actual quantities vary greatly with the auroral intensity. For the calculations of this paper the adopted values are

$$b_0 = 20 \text{ miles} = 3.2.10^6 \text{ cm.},$$
$$k_0 = 10^{-10} \text{ e.m.u.},$$

so that

$$K_0 = 3.2.10^3 \text{ e.m.u.}$$

This is about 40 times as great as the conductivity of a strip of the conducting shell, equal in width to the adopted diameter ($2b_0$) of the zones. Also we assume

$$\theta_0 = 20°, \quad q_0 = 1.02$$

corresponding to an angular radius of 20° for the auroral zone, and a mean height of 80 miles or about 130 km. From these we deduce

$$\rho_0 = 4.34.10^5 \text{ e.m.u.},$$
$$L_0 = 1.28.10^{10},$$
$$M_0 = 2.40.10^7.$$

Apart from the shell and the two zones, the atmosphere is regarded as non-conducting.

The above numerical data for the two zones would probably represent the facts better if b_0 were increased, but for a given value of K_0 or ρ_0 (which in the actual auroral regions must vary widely according to the degree of auroral intensity) a change in b_0 practically only affects L_0; if b_0 were increased tenfold L_0 would be about halved. Such a change in L_0 would appreciably modify the results of this paper, but without affecting the main conclusions to be derived from them. L_0 is perhaps not likely to be much less than half the above value, since it depends logarithmically on b_0, the chief unknown variable.

With regard to the above problem, it may be remarked that instead of supposing the E.M.F. (E_0) to arise suddenly at time $t=0$ (corresponding to the sudden commencement of a magnetic storm), it would perhaps be more appropriate physically to suppose E_0 always present in the zones, but that these are non-conducting until at time $t=0$ they suddenly acquire a conductivity k_0, which thereafter remains constant; the two assumptions are, however, mathematically equivalent.

In the case of actual magnetic disturbances the conductivity of the zones, and the currents in them, do not remain constant, as is supposed in the above problem, but themselves also finally decay. The analysis would not be rendered much more complicated if the zonal currents were assumed to start from nothing and finally to disappear again, but this has not been done in the present paper, because it would have involved assuming a rate of decay of k_0 or E_0; it seemed to us preferable not to make any such assumption, but rather, at least in the first illustrative computations, to determine what rates of current change are involved by the actual constants of the terrestrial model, so far as we could estimate them.

The numerical calculations have been made partly by Mr. T. T. Whitehead, but mainly by Mrs. N. Roberts, to whom grateful acknowledgment is made, and also to Messrs. F. Lord and A. W. King, who have assisted in the checking. Grateful acknowledgment is due to the Government Grant Committee of the Royal Society for financial assistance towards the cost of part of the computations.

The Direct Field of the Zonal Currents

4. We first consider the direct field of a single zonal current i, treated as a linear current of zero cross section, and reckoned positive (like the E.M.F. also) when directed from east to west. Let C and O (Fig. 1) represent the respective centres of the earth and the zone, and let P be any point on the

earth's surface, in colatitude θ, its polar coordinates relative to O being $r(=OP)$ and $\phi(=COP)$. Let $\Omega_0(r, \phi)$ denote the magnetic potential at P due to the current in this one zone. Then

$$\Omega_0(r, \phi)=2\pi i\sum_{n=1}^{\infty}(-1)^{n+1}\frac{1.3\ldots(2n-1)}{2.4\ldots 2n}\left(\frac{r_0}{r}\right)^{2n}P_{2n-1}(\cos\phi),$$

if $r>r_0$ (the radius OA of the zone), or, if $r<r_0$,

$$\Omega_0(r, \phi)=2\pi i\sum_{n=0}^{\infty}(-1)^{n+1}\frac{1.3\ldots(2n-1)}{2.4\ldots 2n}\left(\frac{r}{r_0}\right)^{2n+1}P_{2n+1}(\cos\phi).$$

The components of magnetic force at P, in the directions of increasing r and ϕ respectively, are given by

$$H_r=-\frac{\partial\Omega_0}{\partial r},\quad H_\phi=-\frac{\partial\Omega_0}{r\partial\phi}.$$

The north and vertical components of magnetic force at P are given by

$$H_n=-H_r\sin\psi+H_\phi\cos\psi,$$
$$H_v=-H_r\cos\psi-H_\phi\sin\psi,$$

where $\psi=OPC$. When numerical values of r_0, θ_0, θ, i and a are assigned, H_n and H_v can be readily determined, by finding r, ϕ, ψ graphically or numerically, and substituting in the above equations.

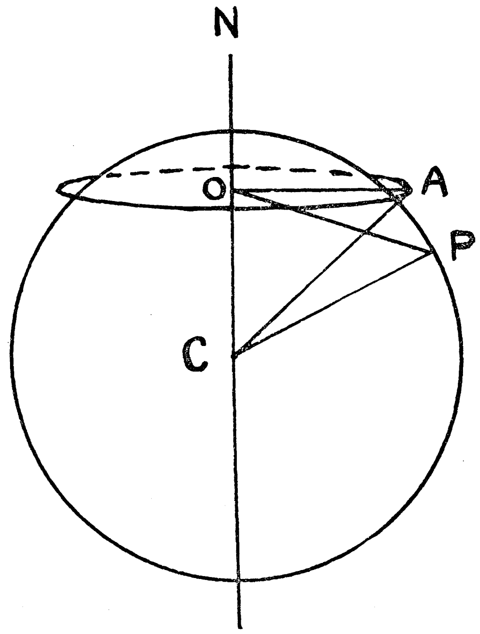


Fig. 1

The values of H_n, H_v at a point P in colatitude θ due to currents i in the *two* zones, northern and southern, are given by $H(\theta)+H(\pi-\theta)$ calculated for the northern zone alone (with appropriate suffix n or v).

Table I gives a rough indication of the distribution of H_n and H_v for the two zones, on the basis of the values of θ_0 and q_0 given in §3; for the purpose of this illustration the current i is taken as 100,000 amperes, or 10,000 e.m.u.; currents of this and greater intensity certainly occur in auroral regions. The values of H_n and H_v are expressed in the units customary in terrestrial magnetic work ($1\gamma = 10^{-5}$ e.m.u.). The current is supposed to flow from east to west.

TABLE I

VALUES OF NORTH FORCE, H_n, AND VERTICAL FORCE, H_v, DUE TO TWO EQUAL WESTERLY CURRENTS OF 100,000 AMPERES EACH AT 80 MILES HEIGHT IN LATITUDES ±70°

Latitude (N)	0°	10°	20°	30°	40°	50°	60°	69°	70°	71°	80°	90°
H_n (in γ)	−0.2	−0.2	−0.3	−0.4	−0.7	−1.2	−2.5	−80.	−155.	−80.	− 4.	0
H_v (in γ)	0	−0.2	−0.4	−0.6	−1.3	−3.	−7.	−86.	0	+86.	+35.	+28.

The vertical force is zero under the zone, and rises to numerical maxima (of opposite sign) in adjacent latitudes on either side; from these its value decreases to zero at the equator, and to a numerical minimum at the pole. The values of H_n and H_v close to the zones are calculated by treating the zone as an infinite rectilinear current, since the above formulae converge very slowly near the zones. Except in this region, H_n and H_v are given in the Table at intervals of 10° latitude. The main object of the Table is to show how small between latitudes ±50° is the direct field of such zonal currents, though near the zones the field is considerable.

5. The values of H_n and H_v could also be got by expanding the magnetic potential of the zonal currents in a series of spherical harmonic terms referred to the earth's centre as origin, instead of O, the centre of a zone. This is less convenient than the semigraphical method actually used, because of the slow convergence of the series obtained. But such an expansion is necessary if spherical harmonic functions are to be used, as in this paper, in calculating the secondary currents induced by variable currents flowing along the zones.

The solid angle subtended by the northern zone at a point on the axis distant r from C (where $r < a_0$) is given by

$$2\pi\left\{1-\frac{a_0\cos\theta_0-r}{(a_0{}^2+r^2-2a_0r\cos\theta_0)^{\frac{1}{2}}}\right\}$$

$$=2\pi\left[1-\cos\theta_0+\sum_{n=1}^{\infty}(r/a_0)^n\{P_{n-1}(\cos\theta_0)-\cos\theta_0P_n(\cos\theta_0)\}\right].$$

When multiplied by i this gives the magnetic potential at P due to a current i along the zone. The corresponding potential at the same point, due to an equal current in the southern zone, is got by substituting $\pi-\theta_0$ for θ_0. Hence the combined field of the two zonal currents, at such a point on the axis, has the potential

$$4\pi i\left[1+\sum_{n=1}^{\infty}(r/a_0)^{2n-1}\{P_{2n-2}(\cos\theta_0)-\cos\theta_0P_{2n-1}(\cos\theta_0)\}\right]\equiv i\sum_{n=1}^{\infty}z_nr^n/a^{n-1},$$

apart from a constant term which can be ignored: where

$$z_n = 0, \quad (n \text{ even}), \tag{5.1}$$
$$z_n = (4\pi/q_0{}^n a) \{P_{n-1}(\cos\theta_0) - \cos\theta_0 P_n(\cos\theta_0)\},$$
$$= (4\pi \sin^2\theta_0/nq_0{}^n a)P_n'(\cos\theta_0), \quad (n \text{ odd}).$$

At a point P in colatitude θ, at a radial distance $r(<a_0)$, we therefore have

$$\Omega = i \sum_{n=1}^{\infty} z_n(r^n/a^{n-1})P_n(\cos\theta). \tag{5.2}$$

Relations Between Primary and Induced Magnetic Fields

6. When the zonal currents i vary, each spherical harmonic term in the expansion (5.2) of their magnetic potential gives rise to induced currents (in the core and atmospheric shell) whose potential involves that spherical harmonic factor, and no other. The relation between the parts of the primary and induced fields associated with any one spherical harmonic $P_n(\cos\theta)$ will naturally depend on the rate of variation of the primary field, that is, on the time-factor involved in i. The relation is expressible conveniently only when the time-factor is periodic or exponential, or can be expanded as a series of periodic or exponential terms. In the present case there is no imposed periodicity, and it is necessary to consider exponential terms[10]. We shall, in fact, suppose that i can be expressed as follows:

$$i = i_0 + \sum_{m=1}^{\infty} i_m \exp(-l_m t), \tag{6.1}$$

where the numbers i_0, i_m, l_m have to be determined so as to satisfy the conditions of the particular problem.

Consider a typical term in the expansion (for $r<a_0$) of the magnetic potential of the field of the zonal currents, *i.e.*,

$$i_m \exp(-l_m t) z_n (r^n/a^{n-1}) P_n(\cos\theta). \tag{6.2}$$

Actually only odd values of n occur (cf. 5.2), but the following relations are valid whether n is odd or even.

Corresponding to the term (6.2) there will be an additional term, in the total magnetic potential for the region outside the atmospheric shell $(a'<r<a_0)$, given by

$$i_m \exp(-l_m t) S_{n,m} (a^{n+2}/r^{n+1}) P_n(\cos\theta), \tag{6.3}$$

arising from the induced currents in the shell and core.

The total magnetic potential for the region between the shell and core will contain two corresponding terms given by

$$i_m \exp(-l_m t)\{C_{n,m}(a^{n+2}/r^{n+1}) + s_{n,m}(r^n/a^{n-1})\}P_n(\cos\theta), \tag{6.4}$$

where the term containing $C_{n,m}$ as factor arises from the induced currents in the core, and the other term is due to the external currents, in the shell and zones.

Inside the core the field does not possess a scalar potential, but may be specified by a vector potential A, the r, θ components of which are zero, while the azimuthal component A_ϕ contains a corresponding term

$$i_m \exp(-l_m t)\, c_{n,m} r^n\, R_n(4\pi k l_m r^2) dP_n(\cos\theta)/d\theta \tag{6.5}$$

where

$$R_n(x^2) = 1 - \frac{x^2}{2(2n+3)} + \frac{x^4}{2.4.(2n+3)\,(2n+5)} - \ldots, \tag{6.6}$$

so that $R_n(x^2)$ is a numerical multiple of $x^{-n-\frac{1}{2}} J_{n+\frac{1}{2}}(x)$.

The above terms in the magnetic potential all depend directly on the zonal currents, and correspond to "forced vibrations" in the analogous theory of an elastic system. The magnetic potential may also include terms corresponding to the "free vibrations" in the elastic problem, these additional terms being required in order to satisfy prescribed initial conditions.

Corresponding to each spherical harmonic $P_n(\cos\theta)$, there will be an infinite series of additional terms analogous to (6.5) in the vector potential A_ϕ for the region within the core, namely

$$c'_{n,k} \exp(-l'_{n,k} t) r^n R_n(4\pi k l'_{n,k} r^2) dP_n(\cos\theta)/d\theta \tag{6.7}$$

where the suffix k takes all positive integral values. These terms correspond to current systems in the core which have no magnetic field outside it, so that the currents must flow in different directions at different depths. There are an infinity of such current systems associated with each spherical harmonic function P_n, and each freely decays at its own characteristic rate, specified by $l'_{n,k}$; the coefficients $l'_{n,k}$ depend only on the properties of the core (cf. §7).

No additional terms are necessary for the magnetic potential in the regions outside the core.

As regards the coefficients z, s, c, S, C, the large letters refer to fields in regions further from the earth's centre than are the currents to which the fields are due, while the small letters refer to fields in regions nearer than the currents are to the earth's centre; c, C refer to the regions inside and outside the core (up to the shell), and s, S refer to the regions inside the shell (to the boundary of the core) and outside it.

7. The following equations connect the above coefficients z, s, c, S, C; they are derived from the equations of current-induction in spheres and spherical shells, and are taken, with slight changes of notation, from our previous memoir[10] (cf. §§7.9, 8.16, 8.17, taking $i\beta^2$ as real). The equations are true whatever the values of the coefficients l_m or $l_{n,k}$:

(7.1) $$\frac{C_{n,m}}{s_{n,m}}=\frac{n}{n+1}q^{2n+1}\left\{1-\frac{R_n(4\pi kl_mq^2a^2)}{R_{n-1}(4\pi kl_mq^2a^2)}\right\},$$

(7.2) $$\frac{C_{n,m}}{c_{n,m}}=nq^{2n+1}\{R_{n-1}(4\pi kl_mq^2a^2)-R_n(4\pi kl_mq^2a^2)\},$$

(7.3) $$s_{n,m}=z_n-\frac{n+1}{n}\,\frac{4\pi K'l_mq'a'}{2n+1}\left(S_{n,m}-\frac{n}{n+1}z_n\right),$$

(7.4) $$C_{n,m}=S_{n,m}-\frac{4\pi K'l_mq'a'}{2n+1}\left(S_{n,m}-\frac{n}{n+1}z_n\right).$$

These equations suffice to determine s, c, S, C when z_n and l_m are known. It may be noted that if the shell is absent, corresponding to $K'=0$, $S_{n,m}$ becomes identical with $C_{n,m}$, and $s_{n,m}$ with z_n (cf. 7.3, 7.4). If the shell is present but the core absent, $C_{n,m}$ is zero and the coefficients $c_{n,m}$, $c'_{n,k}$ do not arise

The time-factors $l'_{n,k}$ associated with the freely decaying current-systems in the core which have no magnetic field outside the core are determined by the condition that the corresponding $C'_{n,k}$, given by an equation analogous to (7.2), is zero while $c'_{n,k}$ is not zero; consequently

(7.5) $$R_{n-1}\,(4\pi kl'_{n,k}\,q^2a^2)=R_n(4\pi kl'_{n,k}\,q^2a^2).$$

It is convenient to suppose the roots $l'_{n,k}$ of 7.5 to be numbered by the suffix k in order of increasing magnitude. The same will be supposed in regard to the coefficients l_m (cf. 8.5). It may be noted that $l_0=0$, $s_{n,0}=z_n$, while $S_{n,0}$, $C_{n,0}$, and $c_{n,0}$ are zero.

The Equation of Induction for the Zones

8. The equation of current flow for either zone has the form

(8.1) $$E_0-dN/dt=\rho_0 i,$$

where E_0 is the westerly applied E.M.F., ρ_0 is the total resistance of the zone, i is the current and N is the number of lines of magnetic force which thread the zone from the north side to the south. N is equal to $(L_0+M_0)i$, due to the self induction and mutual induction of the zones, together with a term due to the induced currents flowing in core and shell. This last part can be found by integrating the radial component of magnetic force corresponding to all the terms $S_{n,m}$ (cf. 6.3) in the potential for the region outside the shell, over the spherical cap of radius a_0 bounded by the zone; for the northern zone the inward radial component must be used, and it is obtained by differentiating 6.3 with respect to r, afterwards putting $r=a_0$. Thus

$$N=(L_0+M_0)i-\sum_{m=1}^{\infty}i_m\exp(-l_mt)\sum_{n=1}^{\infty}(n+1)S_{n,m}(a/a_0)^{n+2}\int_0^{\theta_0}\int_0^{2\pi}P_n(\cos\theta)q_0{}^2a^2\sin\theta d\theta d\phi.$$

This expression, which is valid for both zones, reduces to

$$(L_0+M_0)i-2\pi a^2\sum_{m=1}^{\infty} i_m \exp(-l_m t)\sum_{n=1}^{\infty} S_{n,m}(\sin^2\theta/nq_0{}^n)P_n'(\cos\theta),$$

so that, by (5.1),

$$N=(L_0+M_0)i-\tfrac{1}{2}a^3\sum_{m=1}^{\infty} i_m \exp(-l_m t)\sum_{n=1}^{\infty} S_{n,m}\, z_n. \tag{8.2}$$

On substituting from (6.1) and (8.2) in (8.1), and collecting together the terms which involve the same exponential factor, we get

$$(E_0-\rho_0 i_0)+\sum_{m=1}^{\infty} i_m \exp(-l_m t)[l_m\{L_0+M_0-\tfrac{1}{2}a^3\sum_{n=1}^{\infty} S_{n,m}z_n\}-\rho_0]=0. \tag{8.3}$$

Hence (8.1) will be satisfied if

$$i_0=E_0/\rho_0, \tag{8.4}$$

which gives the steady final value of the zonal currents, and if the numbers l_m satisfy the equations

$$L_0+M_0-\tfrac{1}{2}a^3\sum_{n=1}^{\infty} S_{n,m}\, z_n=\rho_0/l_m. \tag{8.5}$$

It must be remembered that the factors $S_{n,m}$ are themselves complicated functions of l_m (cf. §7).

The values of i_m for $m>0$ are not determined by (8.1) or the equations derived from it; they are available to satisfy any prescribed initial conditions of the problem (§3). The simplest of these conditions is that the initial current in the zones is zero, *i.e.*, putting $i=0$, $t=0$ in (6.1),

$$i_0+\sum_{m=1}^{\infty} i_m=0. \tag{8.6}$$

Also the initial magnetic field and current intensity must everywhere be initially zero. As regards non-conducting regions, this implies that the initial value of the complete coefficient of every separate spherical harmonic term in the expression for the magnetic potential must vanish. Thus, considering the region outside the shell, the condition (8.6) ensures that the direct magnetic field of the zonal currents shall initially vanish for every spherical harmonic. It is therefore sufficient to deal only with the field of the induced currents. Thus we have

$$\sum_{m=1}^{\infty} i_m S_{n,m}=0, \tag{8.7}$$

which must be satisfied for every (odd) value of n; moreover, considering the region between the shell and the core, we must have

$$\sum_{m=1}^{\infty} i_m C_{n,m}=0, \tag{8.8}$$

$$\sum_{m=0}^{\infty} i_m s_{n,m}=0. \tag{8.9}$$

But the last three equations (for any value of n) are not independent; it is obvious from (7.3), (7.4) that

$$s_{n,m}-\frac{n+1}{n}C_{n,m}=z_n-\frac{n+1}{n}S_{n,m};$$

multiplying this equation by i_m and summing with respect to m, from 0 to ∞, we have

$$\sum_{m=0}^{\infty}i_m s_{n,m}-\frac{n+1}{n}\sum_{m=0}^{\infty}i_m C_{n,m}=z_n\sum_{m=0}^{\infty}i_m-\frac{n+1}{n}\sum_{m=0}^{\infty}i_m S_{n,m}.$$

Thus (8.6) together with any two of (8.7), (8.8) and (8.9) imply the truth of the third of the three latter equations. We adopt the equations (8.6), (8.8), (8.9) as those from which the values of $i_m (m>0)$ are to be determined; the coefficients $C_{n,m}$ and $s_{n,m}$ are connected by (7.1), but this does not seem to suffice to make (8.9) a consequence of (8.6) and (8.8) (together with (8.5)).

Finally, each spherical harmonic term in the vector potential for the region within the core must initially vanish at all depths. Though (cf. (6.5)) the function $dP_n(\cos\theta)/d\theta$ is not itself a spherical harmonic function, when multiplied by $\cos\phi$ it becomes a tesseral harmonic of degree n and order 1; every such term in the vector potential multiplied by $\cos\phi$ must separately vanish when $t=0$. We therefore have the equations (one for every odd value of n):

$$\sum_{k=1}^{\infty}c'_{n,k}R_n(4\pi k l'_{n,k}r^2)+\sum_{m=1}^{\infty}i_m c_{n,m}R_n(4\pi k l_m r^2)=0. \tag{8.91}$$

In this equation the i_m's and $c_{n,m}$'s, and also the l_m's, $l'_{n,k}$'s are supposed known and the values of $c'_{n,k}$ are to be determined; since the equation must be satisfied for all values of r within the core, it affords an infinite number of linear equations in the $c'_{n,k}$'s, for example, by expanding the functions R_n in powers of r, and equating the coefficient of each power of r separately to zero. The solution of these equations is, however, unnecessary if all that is required is the magnetic field at the surface $r=a$, outside the core.

It may be added that no freely decaying system of currents in the core *and shell* can exist for which the magnetic force outside the shell is zero; this would correspond to $S_{n,m}=0$, $z_n=0$ in (7.3), (7.4), whence it follows that $s_{n,m}$, $C_{n,m}$ are also zero, and also $c_{n,m}$ by 7.2; this means that such a system has no magnetic field between the shell and core, and that the currents flow wholly in the core—the case already considered.

If instead of assuming E constant, as above, it is desired to calculate the effect of a sudden increase of conductivity at the commencement of a storm, and its subsequent decay, this can be effected approximately by taking E to be of the form

$$E=E'_0(e^{-\alpha t}-e^{-\beta t});$$

α and β are to be chosen to fit the assumed rise and decay of conductivity. Each such term in E gives rise to a particular integral i_α or i_β of (8.1), with a corresponding exponential factor; and there are corresponding currents in the core and shell. Also

(8.92) $$i_\alpha = E_0'/\left\{\rho_0 - \alpha\left(L_0 + M_0 - \tfrac{1}{2}a^3 \sum_{n=1}^{\infty} S_{n,\alpha} z_n\right)\right\},$$

and similarly for i_β, where $S_{n,\alpha}$ is obtained by substituting α for l_m in $S_{n,m}$. When α is assigned, $S_{n,\alpha}$ is determined by (7.1), (7.3), (7.4).

This solution fails if α is a root (l_m) of (8.5); in this case $i_\alpha \exp(-\alpha t)$ must be replaced by $i_\alpha' t \exp(-\alpha t)$ where

(8.93) $$i_\alpha' = E_0'/\rho_0.$$

Similarly for β if this is a root of (8.5).

The coefficients l_m are independent of E, and are determined by (8.5) as before, but (8.6) is to be replaced by the condition

(8.94) $$i_\alpha + i_\beta + \Sigma i_m = 0,$$

unless α or β or both are roots of (8.5), in which case the corresponding terms i_α or i_β or both are absent from (8.94). Similarly (subject to the same exception) two additional terms, corresponding to those having m as suffix, but with m replaced by α or β, must be added to each of the equations (8.7) – (8.91). The coefficient i_0 now vanishes.

The Zones and Core alone

9. It is convenient to write down the special forms which the equations take when the shell is absent, *i.e.*, when $K'=0$. In this case

(9.1) $$s_{n,m} = z_n,\ C_{n,m} = S_{n,m} = \frac{n}{n+1} q^{2n+1}\left\{1 - \frac{R_n(4\pi k l_m q^2 a^2)}{R_{n-1}(4\pi k l_m q^2 a^2)}\right\} z_n,$$

and the time-coefficients l_m are the roots of the equation

(9.2) $$L_0 + M_0 - \tfrac{1}{2}a^3 \sum_{n=1}^{\infty} \frac{n}{n+1} q^{2n+1}\left\{1 - \frac{R_n}{R_{n-1}}\right\} z^2{}_n = \rho_0/l_m,$$

where the argument of the functions R is the same as in (9.1). The equations determining the i_m's are

(9.3) $$\sum_{m=0}^{\infty} i_m = 0,$$

and either (8.7) or (8.8), which are now identical, by (9.1); using (9.1), and cancelling out the common factors which are independent of m, they become

$$\sum_{m=1}^{\infty} i_m \left\{1 - \frac{R_n(4\pi k l_m q^2 a^2)}{R_{n-1}(4\pi k l_m q^2 a^2)}\right\} = 0.$$

The equations determining the currents inside the core (if a knowledge of these currents is desired) are the same as (8.91).

The Zones and Shell alone

10. The special forms of the equations in the case when the core is absent are obtained by writing $k=0$. This leads to

$$c_{n,m}=0,\ C_{n,m}=0,\ c'_{n,k}=0, \tag{10.1}$$

$$s_{n,m}=z_n/\left\{1-\frac{4\pi K' l_m q' a'}{2n+1}\right\}, \tag{10.2}$$

$$S_{n,m}=-\frac{n}{n+1}\,\frac{4\pi K' l_m q' a'}{2n+1}\,s_{n,m}. \tag{10.3}$$

The time-coefficients l_m are the roots of the equation (8.5) with the above value of $S_{n,m}$ inserted. The equations determining the coefficients i_m are (8.6), (8.9); the equations (8.8) are already satisfied in virtue of (10.1), and therefore the equations (8.7) are also satisfied.

The Zones and Two Concentric Conducting Shells

11. Some of the features of the complete model system composed of zones, shell and core can be illustrated, with less difficulty in calculation, by substituting a second thin conducting shell for the core. Let the radius of this inner shell be denoted by $Qa(Q<1)$ and its conductivity by K (corresponding to K' for the outer shell).

The magnetic potential in the regions outside the inner shell will be expressed in the same terms as before, except that $K_{n,m}$ will be written in place of $C_{n,m}$. Inside the inner shell the typical term in the magnetic (scalar) potential will be denoted by

$$i_m \exp(-l_m t) k_{n,m} (r^n/a^{n-1}) P_n(\cos\theta). \tag{11.1}$$

The relations between the coefficients will be as follows, where, for brevity, we write

$$\alpha_{n,m}=\frac{4\pi K l_m Q^2 a}{2n+1},\qquad \alpha'_{n,m}=\frac{4\pi K' l_m q' a'}{2n+1}. \tag{11.2}$$

The following relations are analogous to (7.3), (7.4):

$$k_{n,m}=s_{n,m}-\frac{n+1}{n}\,\alpha_{n,m}\left\{K_{n,m}-\frac{n}{n+1}\,s_{n,m}\right\}, \tag{11.3}$$

$$0=K_{n,m}-\alpha_{n,m}\left\{K_{n,m}-\frac{n}{n+1}\,s_{n,m}\right\}, \tag{11.4}$$

$$K_{n,m}=S_{n,m}-\alpha'_{n,m}\left\{S_{n,m}-\frac{n}{n+1}\,z_n\right\}, \tag{11.5}$$

$$s_{n,m}=z_n-\frac{n+1}{n}\,\alpha'_{n,m}\left\{S_{n,m}-\frac{n}{n+1}\,z_n\right\}. \tag{11.6}$$

These lead to the results

$$S_{n,m} = -\frac{n}{n+1}\,\frac{a_{n,m}+a'_{n,m}}{1-a_{n,m}-a'_{n,m}}\,z_n, \tag{11.7}$$

$$s_{n,m} = \frac{1-a_{n,m}}{1-a_{n,m}-a'_{n,m}}\,z_n, \quad k_{n,m} = s_{n,m}/(1-a_{n,m}). \tag{11.8}$$

These agree with the corresponding equations of §10 when $a_{n,m}=0$, *i.e.*, when the inner shell is absent.

The time-coefficients l_m are given by (8.5) with the above value of $S_{n,m}$ inserted. The equation is the same as in the case of a single shell if $KQ^2+K'q'^2$ is substituted for $K'q'^2$ in (10.2), (10.3).

The coefficients i_m are determined by $\sum_{m=0}^{\infty} i_m=0$ and the first of the sets of equations (for each integral value of n)

$$\sum_{m=0}^{\infty} i_m S_{n,m}=0, \quad \sum_{m=0}^{\infty} i_m s_{n,m}=0, \quad \sum_{m=0}^{\infty} i_m K_{n,m}=0, \quad \sum_{m=0}^{\infty} i_m k_{n,m}=0;$$

if one set of these is satisfied, so also are the remaining three sets.

Part II. Numerical Solution for the Zones and Core (only)

12. The problem described in §3 will be solved numerically in the case where the shell is absent ($K'=0$), so that the zones are in the presence of the core alone, for the numerical values of the properties of the zone and core there given. The solution must, of course, be only approximate, and in fact values of n beyond 19, and of m beyond 11, have been ignored.

Substituting the values of a, q_0, θ_0 in (5.1), the following values of $\log z_n$ are obtained (the letter N indicates that the corresponding z_n is negative): the unit in which z_n is reckoned is 1 e.m.u., or $10^5\gamma$.

TABLE II

Values of log z_n

n	$\log z_n$	n	$\log z_n$
1	$\bar{9}.3546$	11	$\overline{10}.6030$ (N)
3	$\bar{9}.5697$	13	$\bar{9}.1865$ (N)
5	$\bar{9}.5942$	15	$\bar{9}.2566$ (N)
7	$\bar{9}.4721$	17	$\bar{9}.1055$ (N)
9	$\bar{9}.1133$	19	$\overline{10}.4864$ (N)

On dividing (9.2) by $10^3\rho_0$, and substituting the values of ρ_0, L_0, M_0, a, q_i and z_n, it takes the form

$$29.461-\sum_{n=1}^{\infty}\epsilon_n\left\{1-\frac{R_n(4\pi k l_m q^2a^2)}{R_{n-1}(4\pi k l_m q^2 a^2)}\right\}=10^{-3}/l_m; \tag{12.1}$$

ϵ_n is always positive; the following are the values of $\log \epsilon_n$:

TABLE III

n	$\log \epsilon_n$	n	$\log \epsilon_n$
1	$\bar{1}.8288$	11	$\bar{2}.2343$
3	0.3645	13	$\bar{1}.3360$
5	0.3881	15	$\bar{1}.4094$
7	0.0941	17	$\bar{1}.0394$
9	$\bar{1}.3179$	19	$\bar{3}.7330$

It is convenient to write

(12.2) $$x^2 = 4\pi k l q^2 a^2,$$

so that (cf. §3)

(12.3) $$l = 5.830.10^{-7} x^2.$$

It is necessary to calculate $R_n(x^2)/R_{n-1}(x^2)$ for a range of n (odd values only) from 1 to 19, and for x^2 from 0 to 4π (as will appear). We have[10] (l.c. §12: note that Table V of p. 477 of this reference is only approximate, and is superseded by the Table below):

(12.4) $$R_0(x^2) = \frac{\sin x}{x}, \qquad R_1(x^2) = \frac{3}{x^2}\left(\frac{\sin x}{x} - \cos x\right),$$

so that

(12.5) $$\frac{R_1(x^2)}{R_0(x^2)} = \frac{3}{x^2}(1 - x \cot x);$$

values of R_n/R_{n-1} for higher values of n may be determined from R_1/R_0 by repeated use of the recurrence formula

(12.6) $$\frac{R_n}{R_{n-1}} = \frac{(2n-1)\,(2n+1)}{x^2}\left(1 - \frac{R_{n-2}}{R_{n-1}}\right)$$

$$= \frac{(2n-1)\,(2n+1)}{x^2} - \frac{2n+1}{(2n-3)\,(1 - R_{n-3}/R_{n-2})},$$

or (for values of x small compared with the value of n) by the formula

(12.7) $$\frac{R_n}{R_{n-1}} = 1 + \frac{x^2}{(2n-1)\,(2n+3)} + \frac{2x^4}{(2n+1)^2(2n+3)(2n+5)} + \dots.$$

Table IV gives the values of $R_1(x^2)/R_0(x^2)$ for values of x from 0 to 4π at intervals of $\pi/10$; it also gives the corresponding values of $10^5 l$ (cf. (12.3)). When $x=0$, $l=0$, and $R_n/R_{n-1}=1$ for all values of n (cf. (12.7)). From Table IV the values of R_3/R_2 were calculated, by means of (12.6), from π to 4π: these are given in Table V.

TABLE IV

TABLE OF $10^5\, l$ AND $R_1(x^2)/R_0(x^2)$ WHERE $x = r\pi + \theta$

θ	$r=0$		$r=1$		$r=2$		$r=3$	
	$10^5\, l$	R_1/R_0	$10^5\, l$	R_1/R_0	$10^5\, l$	R_1/R_0	$10^5\, l$	R_1/R_0
$\pi/10$	.00575	1.009	.6979	−2.422	2.538	−1.331	5.530	−0.916
$2\pi/10$	.02302	1.027	.8287	− .884	2.785	−0.535	5.892	−0.381
$3\pi/10$	.05180	1.096	.9727	− .3525	3.044	−0.244	6.266	−0.182
$4\pi/10$	.09189	1.128	1.1280	− .0665	3.314	−0.077	6.652	−0.065
$5\pi/10$	.1450	1.216	1.295	.1352	3.596	0.049	7.049	0.025
$6\pi/10$	.2071	1.362	1.473	.2865	3.890	0.164	7.457	0.110
$7\pi/10$	.2819	1.612	1.663	.5134	4.195	0.299	7.877	0.210
$8\pi/10$	.3681	2.118	1.865	.8239	4.511	0.518	8.309	0.367
$9\pi/10$	.4660	3.718	2.076	1.631	4.839	1.049	8.752	0.774
π	.5755	∞	2.302	∞	5.179	∞	9.206	∞

From this the following values of R_3/R_2 were calculated, the range required being (approximately) from π to 4π.

TABLE V

VALUES OF $R_3(x^2)/R_2(x^2)$, WHERE $x^2 = \frac{1}{2}r\pi + \theta$

θ	$r=2$	$r=3$	$r=4$	$r=5$	$r=6$	$r=7$
0	1.212					
$\pi/10$	1.279	2.672	−0.528	0.983	−0.747	0.561
$2\pi/10$	1.368	3.691	−0.080	1.480	−0.297	0.878
$3\pi/10$	1.485	11.999	0.212	2.954	−0.034	1.600
$4\pi/10$	1.664	−5.050	0.450	−4.906	0.164	8.204
$5\pi/10$	1.942	−1.447	0.687	−1.939	0.295	−2.112

From Tables IV, V, and similar tables for higher values of n a graph was drawn of

$$y \equiv \sum_{n=1}^{\infty} \epsilon_n \{ 1 - R_n(x^2)/R_{n-1}(x^2) \} \tag{12.8}$$

(cf. (12.1)) summed for odd values of n from 1 to 19, using the values of ϵ_n given in Table III. When $x=0$, $y=0$; y has infinities at $x=\pi, 2\pi, 3\pi, \ldots$ arising from the term $n=1$; it has infinities near $1.85\pi, 2.85\pi, 3.95\pi, \ldots$ (cf. Table V) arising from the term $n=2$; and further series of infinities for each successive term n: each group of infinities starts with a value of x higher than that for the first infinity of the group arising from the preceding term n, and the infinities of each group alternate with those of the preceding group: this is indicated by the recurrence formula (12.6). Since y varies from $-\infty$ to $+\infty$ between each pair of infinities, one root of (12.1), which may be written as

$$y = 29.46 + 10^{-3}/l, \tag{12.9}$$

lies between each pair of infinities of y.

TABLE VI

VALUES OF l_m AND $1-R_n/R_{n-1}$

n	$m=$	1	2	3	4	5	6	7	8	9	10
	$10^5 . l_m=$	0.5738	1.858	2.289	3.171	4.286	5.050	5.244	6.620	8.187	8.994
1	$1-R_1/R_0$	−213.7	0.190	−27.22	1.156	0.652	−1.753	6.266	1.073	0.688	− 0.665
3	$1-R_3/R_2$	− 0.212	−10.08	2.528	0.672	−0.720	4.260	2.572	0.851	− 0.309	6.615
5	$1-R_5/R_4$	− 0.0783	− 0.380	− 0.572	−1.588	3.407	1.056	0.860	− 0.148	6.958	1.692
7	$1-R_7/R_6$	− 0.0415	− 0.163	− 0.217	−0.363	−0.689	−1.175	− 1.390	9.878	0.779	0.294
9	$1-R_9/R_8$	− 0.0258	− 0.0935	− 0.121	−0.185	−0.289	−0.384	− 0.413	− 0.706	− 1.659	− 4.139
11	$1-R_{11}/R_{10}$	− 0.0177	− 0.0618	− 0.0782	−0.115	−0.170	−0.214	− 0.227	− 0.329	− 0.499	− 0.628
13	$1-R_{13}/R_{12}$	− 0.0129	− 0.0441	− 0.0554	−0.0799	−0.115	−0.141	− 0.148	− 0.202	− 0.283	− 0.332
15	$1-R_{15}/R_{14}$	− 0.0098	− 0.0331	− 0.0414	−0.0591	−0.0832	−0.101	− 0.106	− 0.142	− 0.189	− 0.215
17	$1-R_{17}/R_{16}$	− 0.0077	− 0.0258	− 0.0322	−0.0457	−0.0638	−0.077	− 0.080	− 0.103	− 0.138	− 0.156
19	$1-R_{19}/R_{18}$	− 0.0062	− 0.0207	− 0.0258	−0.0364	−0.0505	−0.061	− 0.063	− 0.083	− 0.106	− 0.119

The roots l of (12.1) or (12.9) were determined approximately from the graph of y, intersected by the graph of $29.46+17^{-3}/l$. Afterwards they were evaluated more precisely by numerical calculations, interpolating for values of R_n/R_{n-1} between those tabulated. The roots (x or l) were numbered in order of increasing magnitude. Each l_m, except l_4 and l_5, lies in a region at which one value of R_n/R_{n-1} is varying rapidly and is of outstanding importance.

Table VI gives the first ten roots l_m, and the values of R_n/R_{n-1} for the corresponding values of x^2 (cf. (12.2)) and for odd values of n from 1 to 19.

13. Having found the roots l_m, the next step was to determine the current-factors i_m from the equations (8.6) and (8.7) or (8.8); by (9.1) the latter are equivalent to

$$\sum_{m=1}^{\infty} i_m(1-R_n/R_{n-1})=0,$$

or, by (8.6), to

$$\sum_{m=1}^{\infty} i_m R_n/R_{n-1}=-i_0, \tag{13.1}$$

where the approximate value of l_m is to be inserted in R_n/R_{n-1}; the common factor

$$nq^{2n+1}z_n/(n+1) \tag{13.2}$$

in (8.7) or (8.8)—cf. (9.1)—has been omitted from (13.1). The coefficients of i_m in the equations (13.1) are the values of R_n/R_{n-1}, given in Table VI.

The solutions are of the form

$$i_m=-\frac{\nabla_m}{\nabla}i_0,$$

where ∇ denotes the determinant whose general element $a_{r,s}$, in the r-th row and s-th column, is R_r/R_{r-1} with l_s inserted: while ∇_m is the sum of the minors of the elements of the m-th column. The determinant ∇ and its minors have an infinite number of rows and columns, but in the approximate solution it is necessary to ignore all save the first j rows and columns, so that only the first j values of i_m can be calculated. These approximate solutions may be denoted by $i_{m,j}$, where

$$i_{m,j}=-\frac{\nabla_{m,j}}{\nabla_j}i_0,$$

the suffix j indicating that the rows and columns beyond the j-th are to be ignored in ∇ and ∇_m.

The solution was executed for $j=5$, 6, 7, 8, 9, 10 in turn. It was found that owing to the increasing similarity of the coefficients in the later equations ($n=8$, 9, 10) ∇_j became rapidly smaller as j increased, and the values of $\nabla_{m,j}$ were obtained as small differences between nearly equal large numbers. The values of R_n/R_{n-1} given in Table VI to four significant figures proved to be not sufficiently accurate to afford a satisfactory solution in the case $j=10$; a recalculation for this case, taking the values of R_n/R_{n-1} to a further decimal place in the last two equations ($n=9$, 10) considerably modified the values of the $i_{m,10}$'s. As the labour of obtaining the coefficients in Table VI to the accuracy necessary for this case was prohibitive, the solution for $j=10$ was abandoned. The solutions in the other five cases are as follows, taking $i_0=10{,}000$ e.m.u.

TABLE VII

SUCCESSIVE SOLUTIONS FOR i_m

j = No. of equations used	$i_{1,j}$	$i_{2,j}$	$i_{3,j}$	$i_{4,j}$	$i_{5,j}$	$i_{6,j}$	$i_{7,j}$	$i_{8,j}$	$i_{9,j}$	$\sum_{m=0}^{j} i_m$
5	2248	−3931	−17309	8643	896					547
6	3316	−6432	−25140	16857	3710	−2124				187
7	5142	−17783	−29969	33629	15415	−42455	26050			29
8	5143	−16760	−31512	33710	14628	−36289	21128	−29		19
9	4875	−14857	−30967	30931	12640	−29494	16935	− 7	−30	26

The last column of the Table gives the values of $i_0+\sum_{m=1}^{j} i_{m,j}$, which by (8.6) is zero when $j=\infty$. The exactness with which this equation is satisfied by the approximate solutions perhaps affords an indication of their probable accuracy, and suggests that the solution for $j=8$ is the best. It would seem that the solution for $j=9$ is beginning to be affected by the causes that in the case $j=10$ completely vitiate the results, and that the additional equations $n=9$, 10 will improve the results for $j=8$ only if the coefficients of Table VI are determined more accurately.

The values of $i_m C_{n,m}$ which appear in the expression (6, 4) for the magnetic potential of the external field of the currents induced in the core were then calculated for these five sets of values of i_m, according to (9.1). The following Table gives the results, expressed in 1_γ as unit, for the single value $n=1$, except that for the solution $j=8$ values for $n=3$ and $n=5$ are also given. It will appear that the remaining values, relating to the harmonics from P, onwards, do not appreciably affect the final results.

TABLE VIII

VALUES OF $i_m C_{n,m}$ IN γ (SUCCESSIVE SOLUTIONS)

j = No. of equations used	$m=1$	2	3	4	5	6	7	8	9
5	− 48.1	−0.1	47.2	1.0	0.1				
6	− 70.9	−0.1	68.5	2.0	0.2	0.4			
7	−110.0	−0.4	81.7	3.9	1.0	7.5	16.3		
8	−110.1	−0.3	85.9	3.9	1.0	6.4	13.2	0.0	
9	−104.3	−0.3	84.4	3.6	0.8	5.2	10.6	0.0	0.0
8; $n=3$	−0.2	35.4	−16.7	4.7	−2.2	−32.4	−11.4	0.0	
8; $n=5$	−0.1	1.3	3.8	−11.2	10.4	− 8.0	3.8	0.0	

The factors $\exp(-l_m t)$ occurring in (6.4) are given in Table IX, for values of t at intervals of 3 hours up to 12 hours and at longer intervals up to 4 days. Values less than 0.001 are omitted.

TABLE IX

VALUES OF EXP $(-l_m t)$

Time in hours =	0	3	6	9	12	18	24	36	48	96
$m=1$	1	0.940	0.883	0.830	0.780	0.690	0.609	0.475	0.371	0.138
2	1	0.818	0.669	0.548	0.448	0.300	0.204	0.090	0.040	0.002
3	1	0.781	0.610	0.476	0.372	0.227	0.138	0.051	0.019	
4	1	0.710	0.504	0.358	0.254	0.128	0.065	0.016	0.004	
5	1	0.629	0.396	0.249	0.157	0.062	0.025	0.004		
6	1	0.579	0.336	0.195	0.113	0.038	0.013	0.001		
7	1	0.568	0.322	0.183	0.104	0.033	0.011	0.001		
8	1	0.489	0.239	0.117	0.057	0.014	0.003			
9	1	0.413	0.171	0.070	0.029	0.005				
10	1	0.378	0.143	0.054	0.020	0.003				
11	1	0.369	0.136	0.050	0.019	0.003				

14. The magnetic potential of the field of the currents induced in the core is

$$(14.1) \qquad \Omega_i = \sum_{n=1}^{\infty} \Omega_i^{(n)} \, (a^{n+2}/r^{n+1}) P_n(\cos\theta),$$

where

$$(14.2) \qquad \Omega_i^{(n)} = \sum_{m=1}^{\infty} i_m C_{n,m} \exp(-l_m t).$$

The values of $\Omega_i^{(n)}$ are readily calculated from Tables VIII and IX for the values of n and t occurring in those tables. They are negligible except in the cases $n=1$, 3; the results for $\Omega_i^{(1)}$ are given in Table X for the five solutions $j=5$ to $j=9$, and for the solution $j=8$ the values of $\Omega_i^{(3)}$, $\Omega_i^{(5)}$ are also given.

TABLE X

VALUES OF $\Omega_i^{(n)}$ IN γ (SUCCESSIVE SOLUTIONS)

Time in hours	0	3	6	9	12	18	24	36	48	96
$\Omega_i^{(1)}$; $j=5$	0	−8	−13	−17	−20	−22	−23	−20	−17	−7
6	0	−11	−21	−25	−29	−33	−34	−30	−25	−10
7	0	−23	−37	−46	−52	−56	−55	−48	−39	−15
8	0	−22	−36	−45	−51	−55	−55	−48	−39	−15
9	0	−20	−34	−42	−47	−52	−52	−45	−37	−14
$\Omega_i^{(3)}$; $j=8$	0	5	8	8	8	6	5	2	1	0
$\Omega_i^{(5)}$; $j=8$	0	0	0	0	0	0	0	0	0	0

The entirely outstanding importance of the first harmonic, $P_1(\cos\theta)$, in the magnetic potential of the field of the induced currents, is the most remarkable feature of Table X, and it constitutes one of the chief results of the whole calculation. This harmonic is found [1, 11] to be predominant in the magnetic field of the storm-time variations in middle and low latitudes during magnetic storms; it is of great interest to find that this can probably be explained as due to currents induced (probably in the atmosphere as well as in the earth, though the latter alone is here considered—cf. §12) by primary currents flowing exclusively in high latitudes.

15. The north force at the earth's surface ($r=a$) due to the induced currents is equal to

$$\left(\frac{\partial\Omega_i}{r\partial\theta}\right)_{r=a} = \sum_n \Omega_i^{(n)} \frac{dP_n(\cos\theta)}{d\theta}. \tag{15.1}$$

The first term, $-\Omega_i^{(1)}\sin\theta$, is much larger than the others, and from it and the next term ($n=3$) are calculated the following values of the north force variations, due to the currents in the core, in latitudes 0°, 20°, 40°, 60°, 70°. The results refer to the solution $j=8$, though those for $j=7$, $j=9$ differ only slightly from them.

TABLE XI

APPROXIMATE NORTH FORCE, AT THE EARTH'S SURFACE, DUE TO THE CURRENTS INDUCED IN THE CORE; UNIT 1γ

Time in hours	0	3	6	9	12	18	24	36	48	96
Latitude 0°	0	30	48	58	62	65	62	51	41	15
" 20	0	24	38	48	52	56	54	46	37	14
" 40	0	10	18	25	29	35	36	34	29	11
" 60	0	0	2	6	9	14	18	19	18	7
" 70	0	− 2	−1	1	3	8	10	12	12	5

The corresponding vertical force (measured positive downwards) is given by

$$\left(\frac{\partial\Omega_i}{\partial r}\right)_{r=a} = -\sum_n (n+1)\Omega_i^{(n)} P_n(\cos\theta). \tag{15.2}$$

The term $n=3$ is relatively more important here than in (15.1); it has been included with the principal term in calculating the following approximate values.

TABLE XII

APPROXIMATE VERTICAL FORCE, AT THE EARTH'S SURFACE, DUE TO THE CURRENTS INDUCED IN THE CORE; UNIT 1γ

Time in hours	0	3	6	9	12	18	24	36	48	96
Latitude 0°	0	0	0	0	0	0	0	0	0	0
" 20	0	24	37	45	48	48	45	36	28	11
" 40	0	35	56	68	75	79	76	64	52	20
" 60	0	31	53	68	77	88	89	80	67	26
" 70	0	27	48	63	74	87	90	84	71	28

16. It is also of interest to consider the rate at which the currents increase in the zones, and to compare this with the rate of growth as it would be if the earth's interior had been non-conducting. The results can in the former case

be given only roughly, to the nearest 100 e.m.u. After 48 hours the steady value of the zonal currents is reached and for a time exceeded. In the absence of the core this does not occur.

TABLE XIII

GROWTH OF ZONAL CURRENTS, WITH AND WITHOUT THE CORE; UNIT 1000 e.m.u.

Time in hours	0	3	6	9	12	18	24	36	48	96
Core present	0	0.6	1.5	2.6	3.7	5.9	7.7	9.9	10.8	10.7
Core absent	0	3.1	5.2	6.7	7.7	8.9	9.5	9.9	10.0	10.0

The presence of the core thus retards the growth of the zonal currents during the first day, and afterwards increases the currents beyond their final steady value.

16. When the changes in north force due to the currents in the core, as given in Table XI, are compared with those due to the direct field of the zonal currents, as given in Table I (though these should be increased by 10% to accord with the maximum value of the zonal currents, as indicated by Table XIII), it becomes clear that between latitudes $\pm 60°$, and during most of the first four days from the commencement of the currents, the secondary induced field is much more important (in the north component) than the direct field. In still higher latitudes, however, the secondary field decreases rapidly while the direct field increases considerably, and the latter is much the greater. This confirms the possibility mentioned in §2.1 that the magnetic disturbance observed in low latitudes may be only a secondary consequence of primary currents themselves confined to high latitudes.

It is interesting to compare the north force changes of Table XI with those obtained by analysis of actual magnetic storms[1], which are as follows:

TABLE XIV

AVERAGE NORTH FORCE CHANGES OBSERVED DURING MAGNETIC STORMS: UNIT 1γ

Time in hours from commencement of storm	0	3	6	9	12	18	24	36	48
Group I latitude 22°	0	4	−16	−26	−34	−35	−37	−31	−26
" II " 40°	0	−1	−19	−24	−32	−33	−30	−27	−23
" III " 53°	0	7	− 1	−20	−24	−22	−19	−17	−15

These values are read off from the graphs in Fig. 1 of the paper cited, and are not smoothed in any way. The graphs refer only to the first two days of the storms. As the values are read off only at intervals of 3 hours, the form of the graphs during the first few hours (during which the north force is above its normal value) is not well indicated by Table XIV; in any case this feature is not shown by Table XI, nor was this to be expected.

The absolute magnitude of the north force changes in Table XI depend upon the assumed value of E_0 or i_0; by assuming a value about half as great, Tables XI and XIV might be made to accord fairly well after the first 6 or 9 hours, except as regards sign. The sign in Table XI depends on the assumed (westerly) direction of the zonal currents, and if the direction had been supposed easterly the two tables would agree in sign. But the diminution of north force in high latitudes during magnetic storms, due in our view to the direct field of the zonal currents as in Table I, indicates with considerable certainty that the zonal currents are westerly, as we have assumed. The difference of sign between Tables XI and XIV therefore requires explanation (cf. §17). But the feature in which these two Tables are most accordant is the epoch of maximum force and the rate of subsequent decay of the field. The rate of decay after the first 24 or 36 hours depends almost entirely on the factor l_1 in $\exp(-l_1 t)$. This is $0.5738.10^{-5}$, only slightly less than $0.5755.10^{-5}$, which corresponds (cf. Table IV) to $R_1/R_0 = \infty$, that is, to the slowest rate of free decay of currents in the core, in the absence of the conducting auroral zones. The zones only slightly retard the rate of decay of the currents in the core, and it is probable that the conducting atmospheric shell also has only a slight influence on the rate of decay. It may be concluded, with considerable probability, that the rate of recovery of the earth's field after a magnetic storm (the so-called Nachstörung) is mainly governed by the rate of decay of currents flowing in the earth's core.

17. Besides the difference of sign between Tables XI and XIV, another striking discrepancy between the disturbance-fields here considered, and the storm-time field of a magnetic storm, is shown by Table XII for the vertical force changes. The latter much exceed those of Table XI for the north force, whereas the observed average vertical force changes in middle latitudes[1] are very small (they amount to only a few γ; they are positive except near the beginning of the storm).

These discrepancies seem to be due to the neglect of the conducting atmospheric shell in the present calculations. In a later paper the complete problem may be dealt with, but it seems likely to be far more difficult than the calculations of this paper, already sufficiently complicated. The results here obtained suffice, however, to give a general indication of some features of the complete solution, and there is a likelihood that it will be free from the above two principal discrepancies with observation.

If the core were absent and the shell alone present with the zones, the above solution suggests that the magnetic field of the currents in the shell would be of similar type, outside the shell, to that outside the core. This implies that the magnetic potential outside the shell would depend mainly on a term $\Omega^{(1)} r^{-2} P_1(\cos\theta)$ where, as in Table X, $\Omega^{(1)}$ is negative. The magnetic potential for the secondary field inside the shell (*e.g.*, near the earth's surface) would then consist mainly of a term proportional to $\Omega^{(1)} r P_1(\cos\theta)$. The corresponding north force change at the earth's surface would be negative, and likewise also the vertical force changes (reckoned as in §15). Moreover the ratio of the north force changes to the vertical force changes due to the shell is twice as great as in the case of the core, because the respective potentials for shell and core depend on r (inside the shell) and r^2 (outside the core). Consequently the main observed average changes in north and vertical forces during magnetic storms would be accounted for if the coefficient of $rP_1(\cos\theta)$ in the magnetic potential for the shell, when present with the core as well as with the zones, were about twice as great as that of $(a^3/r^2)P_1(\cos\theta)$ in the magnetic potential for the core; in this case the vertical force changes due to shell and core would almost neutralize one another, while the diminution in the north force due to the shell would be twice as great as the increase due to the core. The ratio 2 for the corresponding harmonic terms $P_1(\cos\theta)$ in shell and core is approximately the same as that found for the principal harmonics in the potential of the diurnal magnetic variations. In the case of magnetic disturbance, it has to be remembered that the currents in shell and core are both secondary, the primary currents being in the auroral region; the atmospheric shell will to some extent shield the core from the changing field of the zonal currents.

It must, of course, be tested by actual calculation whether the effects due to shell and core are in the approximate ratio 2. In a later paper we hope to publish the results of such a calculation, and also to indicate how the results are influenced by reasonably possible changes in the estimated conductivities and other data for the shell, core, and zones. Another conjecture which will be tested is that if L_0 is halved as suggested in (3.2) the principal positive term $i_m C_{n,m}$ (Table VIII) will be for a larger value of l_m than in the present calculation, larger also than that for the principal positive term in the magnetic potential for the shell; this would correspond to a more rapid rise of the induced currents in the core than in the shell (and than in the present calculations). If this occurs, then during the first few hours of a magnetic storm the positive change of north force due to the core might surpass the negative change due to the shell. This seems the simplest likely explanation of the average increase in the north force during the first phase of magnetic storms.

REFERENCES

[1]S. Chapman, Proc. Roy. Soc., A, *95*, 61, 1918.
[2]F. A. Lindemann, Phil. Mag., *38*, 669, 1919.
[3]S. Chapman, The Observatory, London, *42*, 196, §§16, 17, 1919.
[4]S. Chapman, Proc. Camb. Phil. Soc., *21*, 577, 1923.
[5]S. Chapman, Quar. Jour. Roy. Met. Soc., *52*, 225, §5, 1926.
[6]C. Størmer, numerous papers from 1904 onwards, the latest being in Geofysiske Publ., Oslo, 4 (7), 1926.
[7]E. A. Milne, Monthly Notices, Royal Ast. Soc., *86*, 459, 1926.
[8]S. Chapman, Proc. Royal Soc., A, 115, p. 243, 1927; cf. also The Observatory, London, *48*, 105, 1925.
[9]A. Schuster, Phil. Trans., A, *180*, 467, 1889; S. Chapman, *ibid.*, *218*, 1, 1919.
[10]S. Chapman and T. T. Whitehead, Trans. Camb. Phil. Soc., *22*, 463, 1922.
[11]S. Chapman, Monthly Notices, Royal Ast. Soc., 79, 70, 1918.

LE PROBLÈME ACTUEL DES HORLOGES ÉLASTIQUES—FROTTEMENTS ET VISCOSITÉS

PAR M. JULES ANDRADE,
Professeur à l'Université de Besançon, Besançon, France.

I

Sous le nom d'horloges élastiques, je me propose, dans ce mémoire, d'étudier des instruments de précision capables d'une mesure mécanique du temps dont l'isochronisme peut aujourd'hui être réalisé à *l'approximation relative du millionième.*

Cette précision peut, en effet, être obtenue par l'emploi *d'organes réglants* élastiques attelés à un balancier rigide mobile autour d'un axe vertical, et dont l'axe d'oscillation peut appartenir à 3 types distincts que nous préciserons par la suite.

Insistons tout d'abord sur ce point important: le *balancier est rigide;* ceci équivaut à dire que l'organe vibrant et réglant mesureur du temps, est compensé par les méthodes de M. Guillaume, qui paraissent avoir enfin, *dans la métallurgie des invars et de l'élinvar, conquis une sécurité définitive et complète de fabrication.*

Grâce aux récentes et mémorables découvertes de M. Guillaume, la compensation aux températures n'est plus désormais confiée aux constructions incertaines de balanciers multi-métalliques et déformables qui, à leur heure, déjà au 18[e] siècle, furent la grande œuvre de deux chronométriers anglais dont nous saluons ici la mémoire: Earnshaw et Arnold.

Grâce à M. Guillaume, la compensation, aujourd'hui confiée à la *création métallurgique d'une molécule physique nouvelle,* va nous permettre désormais *de pousser à fond le développement de l'isochronisme du balancier réglant des chronomètres fixes et des horloges élastiques.*

Avec la bague bimétallique *fendue* et *déformable* disparaitra la plus forte des perturbations d'isochronisme qui est due à la très petite déformation de cette bague dans le cours même d'une oscillation simple.

Rappelons d'ailleurs que cette perturbation, longtemps insoupçonnée des horlogers, leur fut révélée par Phillips. C'est grâce à la disparition de cette perturbation, par un balancier réglant *désormais rigide,* que nous pourrons aujourd'hui réaliser l'isochronisme à l'approximation relative de $\frac{1}{10^6}$; et ce résultat est obtenu *par un choix convenable de ressorts hélicoïdaux réglants associés.*

Avant d'exposer et de justifier, dans leur ensemble, ces nouvelles méthodes nées de l'alliance de la géométrie avec la chimie physique des alliages, je rappellerai tout d'abord une remarque préalable que j'ai faite en 1911, dans mon premier brevet.

Dénommons *d'abord* les caractéristiques d'un ressort hélicoïdal:

E, coefficient d'élasticité de ce ressort;

R, le rayon de sa projection sur un plan *transverse*, c'est-à-dire perpendiculaire à l'axe d'oscillation du balancier;

P, son étendue angulaire, en sorte que l'inclinaison des spires sur un plan transverse étant nommée i, l'on a entre l'attache fixe du ressort, ou *piton*, et l'attache mobile ou *virole*, toutes deux respectivement situées dans des plans transverses parallèles, séparées par une distance h, la nécessité de la relation:

$$\text{tang } i = \frac{h}{2\pi RP},$$

en sorte que P est compté avec l'unité d'angle: le radian.

Ensuite, nous rappellerons *l'artifice* par lequel Résal a pu établir la première partie d'un problème qui, au premier abord, n'est nullement réductible à la Mécanique rationnelle, si on ne veut s'intéresser tout d'abord, dans le ressort réglant, qu'aux seuls efforts élastiques transverses, dont le moment, par rapport à l'axe d'oscillation, importe seul à l'évaluation du *moment mécanique* transmis par la virole du ressort, autour de l'axe du balancier oscillant, en l'extrémité fixe d'un rayon du balancier, nommé rayon de virole.

Nous supposerons donc, avec Résal, que l'on peut remplacer l'hélice naturelle et l'hélice déformée par leurs projections sur un plan transverse.

De plus, Résal a prévu, dès le début du calcul, que l'on pouvait regarder comme géométriquement négligeable l'allongement de cette projection de la fibre moyenne du ressort déformé, sur un plan transverse.

On est ainsi conduit à ce simple problème de *géométrie plane:* étudier la flexion d'un ressort primitivement circulaire, mais supposé naturellement déformé dans un plan transverse. Ainsi posé, le problème plan réduit s'intègre par quadratures; il conduit à des intégrales elliptiques dépendant d'un radical

$$\sqrt{1+a\cos z+b\sin z},$$

radical où les constantes a et b, liées à l'angle de rotation u du balancier *sont petites* et où la variable z représente, dans le problème plan réduit, l'inclinaison de la normale à la courbe projection de la fibre moyenne déformée, sur le rayon transverse passant par l'attache fixe du spiral ou piton.

L'intégration de Résal, par le développement en série du radical, arrêté aux premières puissances de a et b a conduit Caspari à des approximations, plus maniables, et par lesquelles les *composantes transverses* de la force isolée complémentaire du couple d'encastrement, exprimées en a et b, permettent, conformément à la méthode introduite par Phillips, de ramener le calcul du moment *statique* transmis par le ressort hélicoïdal autour de l'axe d'oscillation du balancier à la seule connaissance de la position transverse du centre de gravité de la fibre moyenne du ressort hélicoïdal projetée sur un plan transverse. Ce calcul exige d'ailleurs la détermination du centre de gravité de cette projection de la fibre moyenne du ressort hélicoïdal, définie par le problème réduit de Résal.

Malgré une erreur échappée à Caspari, dans la détermination préalable de ce centre de gravité, le *moment mécanique* transmis au balancier garde la valeur assignée par Caspari, dont deux erreurs se sont, ici, *fortuitement et heureusement compensées*, comme je l'ai montré dans mon petit volume sur *Les organes réglants des chronomètres.* (Bienne et Besançon, Magron éditeur, 1922).

Ainsi a pu être sauvée la très intéressante justification de la méthode d'isochronisme de Pierre Le Roy, entreprise par Caspari, tout comme la méthode de Phillips avait conduit cet ingénieur français à la justification des fameuses *courbes tâtées* du chronométrier anglais Arnold; celles-ci, durant près d'un siècle, étaient restées les *fameuses et mystérieuses courbes terminales d'Arnold;* ces courbes, à la plus grande satisfaction des régleurs, avaient été définies en 1861 par Phillips.

En 1911, j'ai déduit de la méthode de Caspari la conséquence suivante, qui fut le point de départ de mes découvertes géométriques et mécaniques, exposées dans ce mémoire et qui vont reposer sur l'association systématique des *spiraux hélicoïdaux.*

Prenez un ressort hélicoïdal homogène, d'étendue angulaire $(4n \pm 1)\pi$; coupez-le en deux portions égales; on obtient ainsi deux nouveaux ressorts hélicoïdaux dont l'étendue angulaire est:

$$P = (2n \pm \tfrac{1}{2})\pi.$$

Attelez ces deux ressorts à un même balancier, de manière que les deux nouveaux spiraux gardent, à deux étages différents, leurs places relatives, de part et d'autre de la coupure intermédiaire primitive qui s'est ainsi dédoublée en deux *viroles* intermédiaires, mais attachées d'une manière arbitraire en attribuant à l'un des ressorts un *mouvement relatif hélicoïdal de glissement* par rapport à l'autre, ce mouvement de glissement étant d'ailleurs quelconque, et les attaches extrêmes du ressort primitif restant attachées aux deux pitons extrêmes sur le bâti du chronomètre.

La Figure 1 montre le premier dispositif expérimental de cette sorte réalisé au Laboratoire de Mécanique de l'Université de Besançon, avant 1913.

Construit par Ernest Jaccard, mort en cette même année, il permet de faire tourner en bloc le spiral inférieur par rapport au spiral supérieur; à cet effet, une plaque circulaire mobile, non visible, glisse concentriquement sur le bâti inférieur fixe et permet de déplacer ainsi, à volonté, le piton inférieur, porteur de l'attache fixe du spiral inférieur. Cet appareil m'a servi à me rendre compte rapidement des orientations relatives des deux spiraux qui rendent les frottements plus ou moins forts ou plus ou moins faibles. On en jugeait d'après les durées d'extinction du mouvement vibratoire naturel, appréciées approximativement à l'aide d'un microscope.

Cet assemblage de deux spiraux jouit des propriétés mécaniques suivantes:

1° Les deux nouveaux spiraux, anciennes portions du spiral primitif, avant le sectionnement de celui-ci, travaillent en sens inverses, l'un par rapport à l'autre; l'un se ferme, quand l'autre s'ouvre.

2° Si l'on fait la somme des deux *moments mécaniques* transmis au balancier commun par ce spiral double, on trouve que la somme de ces moments, lorsque le balancier a tourné de l'angle u a pour valeur une expression de la forme $-Hu$, expression dont le coefficient H comprend deux parties; d'abord une partie constante dont la valeur est $2\frac{EI}{RP}\left(1+\frac{4}{P^2}\right)$ et ensuite une partie irrégulière, fonction de u mais dont la valeur, par rapport à la précédente, est petite de l'ordre de $\frac{1}{P^3}$ voisine du millionième dès que P atteint la valeur de 15 tours, $P=15.2\pi$.

Ainsi, le doublet de ressorts qui précède produit un moment mécanique autour de l'axe d'oscillation, moment dont l'isochronisme est approché avec une erreur relative de l'ordre de $\frac{1}{10^6}$.

J'ai appelé ce doublet: *un doublet sinusoïdal* parce qu'au lieu de produire, comme dans la méthode de Pierre Le Roy, un mouvement isochrone par com-

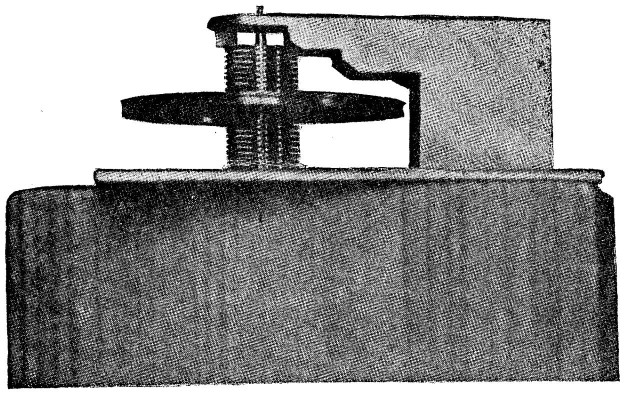


Fig. 1

pensation, sa vibration est isochrone en raison de la régularité du moment transmis au balancier; en sorte que, à cette approximation, le mouvement possède l'isochronisme sinusoïdal. On ne saurait trop insister sur ce point, car j'ai démontré que l'on peut généraliser une remarque très précieuse et déjà ancienne de Villarceau; non-seulement, le frottement constant, mais un assez grand nombre de frottements que nous énumèrerons et que nous pouvons précisément produire avec des ressorts hélicoïdaux convenablement associés, sont rigoureusement isochrones.

Le mécanisme de la méthode d'isochronisme de Pierre Le Roy, telle qu'elle résulte de l'intéressante théorie de Caspari, est le suivant:

Avec un seul spiral, et en tenant compte de la partie irrégulière et non pendulaire du moment mécanique transmis par le ressort réglant au balancier, on forme l'équation différentielle du mouvement; la partie irrégulière du moment mécanique transmis au balancier est petite, et on applique à l'intégration de la dite équation différentielle, la méthode de la variation des constantes en vue de la détermination de la durée d'une oscillation, *au régime d'amplitude normale du chronomètre.*

Ce calcul fournit pour la valeur de cette perturbation cette forme:

$$S(u_0) \cos P,$$

$S(u_0)$ étant une série déjà rencontrée par Bessel, mais ici, *au point de vue de Caspari,* l'intérêt pratique de ce calcul n'était pas le facteur $S(u_0)$, mais au contraire, le facteur $\cos P$, car celui-ci s'annule précisément par l'hypothèse de *construction*

$$P = (2n \pm \tfrac{1}{2})\pi,$$

où Caspari a reconnu sensiblement la caractéristique des spiraux de Pierre Le Roy.

Mais, précisément, ce qui m'a incité à poursuivre mes recherches et à modifier complètement l'objectif à atteindre est l'observation suivante. La justification de la méthode de Le Roy est complète, mais la méthode de Le Roy ne réalise pas l'isochronisme sinusoïdal.

Or, au point de vue pratique, mais en nous dirigeant vers les plus hautes approximations possibles de l'isochronisme, *en fait le millionième, mon but a été de conserver à coup sûr* l'approximation du millionième, même en tenant compte des pressions élastiques génératrices de frottements; or, en adoptant ce but précis, il nous reste à examiner les avantages offerts par les ressorts hélicoïdaux associés.

Je vais maintenant exposer ici comment ce but a été atteint déjà dans la partie mathématique du problème et comment il va l'être dans la partie expérimentale de ces recherches.

II

Dans la voie ainsi ouverte, j'ai reconnu avant tout la nécessité d'avoir une interprétation géométrique simple et maniable de la composante transverse des forces élastiques isolées (complémentaires des couples d'encastrement) qui, sur la virole, se transmet au balancier ou qui, vers le piton, réagit sur le bâti du chronomètre.

Avant même de commencer cette étude, il y avait lieu de compléter et même aussi de corriger l'erreur de Caspari sur la position transverse que la méthode de Résal-Caspari peut assigner au centre de gravité d'un spiral isolé, afin de pouvoir utiliser ces résultats pour la discussion complète des pressions pour les ressorts associés. Voici mes résultats:

1° En ce qui concerne la composante transverse de la force complémentaire:

Soit (Fig. 2), en projection sur un plan transverse, Q le point *piton,* attache fixe du ressort hélicoïdal;

V son attache mobile ou *virole* pour la position d'équilibre du système du balancier et de son ressort réglant;

W la même attache mobile, lorsque le balancier a tourné de l'angle u autour de l'axe d'oscillation du système.

Pour un ressort donné unique, le sens positif des rotations est le sens dans lequel s'enroule le ressort hélicoïdal depuis l'attache fixe Q vers l'attache mobile V.

Dans la situation, où u positif, la virole V du repos étant venue en W, les flèches figurées représentent, en Q la pression du bâti sur le bout piton du ressort; et en W la pression de l'extrémité du rayon de virole sur le bout virole du ressort; en changeant les sens respectifs de ces flèches, on aura respectivement l'attraction

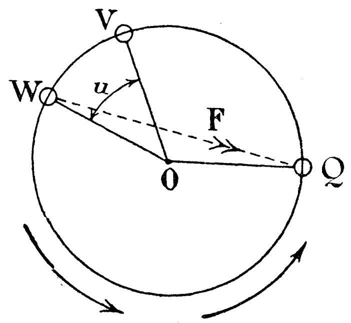


Fig. 2

de l'extrémité W du rayon de virole sur le balancier dirigée vers le piton et l'attraction du bout de piton du bâti vers la virole du balancier.

L'attraction F du point virole du balancier vers le piton Q a pour valeur:

$$\vec{F} = \frac{u}{P^2} . 2 . \frac{E.I}{R^2} . \frac{\overrightarrow{WQ}}{R} . \left[1 + \frac{u - \epsilon \cos u}{P} \right].$$

Dans ces formules:

I désigne le moment d'inertie de l'aire de la section du ressort perpendiculaire à la fibre moyenne hélicoïdale par rapport à l'axe de flexion de cette section; $\epsilon = \pm 1$, lorsque $P = (2n \pm \frac{1}{2})\pi$; et les flèches sur la force $\vec{F}$ ou sur le vecteur $\overrightarrow{WQ}$ représentent deux vecteurs dirigés de la virole W vers Q.

Tout se passe donc en plan transverse comme si la virole du balancier était attirée vers le piton; et si, au contraire, le ressort travaille négativement, l'attraction virole vers piton, se change en répulsion virole depuis le piton Q. Ces formules sont exactes à l'ordre absolu près des termes en $\frac{1}{P^4}$ qui est l'ordre relatif du millionième poursuivi par rapport au moment régulier transmis au balancier.

S'il n'y avait que des forces élastiques transverses, la discussion des pressions génératrices de frottement qui constituent des forces isolées, c'est-à-dire non couplées dans les couples d'encastrement, se réduirait à la discussion des composantes F et il deviendrait nécessaire de résoudre le problème élastique complet dont Résal n'a systématiquement envisagé que l'une des faces. Je reviendrai sur quelques précisions à ce sujet à la fin de ce mémoire de manière à préciser un problème, généralisation du problème de Résal et dont la solution serait digne de tenter nos jeunes géomètres.

Fort heureusement, pour notre but essentiel, le problème complet des horloges élastiques peut recevoir une solution pratiquement satisfaisante, mais qui ne sera pas nécessairement la plus générale possible; nous allons l'exposer sans attendre la solution générale du problème de l'équilibre élastique d'un ressort hélicoïdal.

Mais, avant d'exposer une solution du problème des horloges élastiques, je mentionnerai, sur notre chemin, une hypothèse d'origine assez peu connue; je l'appelle *hypothèse des techniciens*, en souvenir des circonstances où je l'ai rencontrée, pour la première fois, à propos de discussions relatives à un brevet.

Tout ce que la Mécanique rationnelle permet d'ajouter à l'œuvre de Résal est ceci: envisageons l'équilibre purement statique d'un ressort hélicoïdal, après rotation u de son balancier par rapport au point mort général; si même nous négligeons tout d'abord la pesanteur du ressort, nous aurons sur chacune des extrémités de sa fibre moyenne au piton, d'une part, et sur le bout virole de cette même fibre, d'autre part, une force isolée complémentaire et un couple d'encastrement; les deux forces complémentaires et les deux couples complets d'encastrements forment sur le spiral solidifié un système de forces en équilibre, ce qui exige d'abord que les deux forces complémentaires soient équipollentes, c'est-à-dire que leurs composantes transverses soient équipollentes et que leurs composantes longitudinales soient équivalentes.

L'hypothèse, dite des techniciens, consiste à admettre que les deux forces complémentaires s'équilibrent, auquel cas, les deux couples d'encastrement s'équilibrent également entre eux.

Dans ces conditions, les deux bouts libres du ressort seraient soumis pour $u>0$ aux mêmes forces complémentaires que s'ils se repoussaient naturellement proportionnellement à la distance piton-virole et avec l'intensité

$$\frac{u}{P^2}\cdot\frac{2EI}{R^2}\cdot\frac{\overrightarrow{WQ}}{R}\cdot\left[1+\frac{u-\epsilon\cos u}{P}\right],$$

mais le vecteur WQ n'étant plus transverse, mais étant rétabli dans l'espace.

Dans les figures prochaines, et pour fixer les idées, nous prendrons

$$\epsilon=+1;\ P=(2n+\tfrac{1}{2})\pi.$$

Une hypothèse un peu plus générale que l'hypothèse des techniciens consiste à prendre pour la composante transverse:

$$\overrightarrow{F}=\frac{u}{P^2}\cdot\frac{2EI}{R^2}\cdot\frac{\overrightarrow{WQ}}{R}\cdot\left[1+\frac{u-\epsilon\cos u}{P}\right],$$

et pour la composante longitudinale

$$G=\frac{u}{P^2}\cdot\frac{2EI}{R^2}\cdot\left[1+\frac{u-\epsilon\cos u}{P}\right]\frac{h}{R}\cdot m;$$

$m=1$ équivaut à l'hypothèse des techniciens. Dans ces formules h désigne la distance des plans transverses du piton et de la virole.

D'ailleurs $h = 2\pi RP$ tang i.

Comme exemple (Fig. 3), étudions la répartition des pressions élastiques sur le doublet sinusoïdal dans le cas de viroles opposées; en adoptant, par exemple, l'hypothèse des techniciens $m = 1$, l'on séparera les trois groupes des termes de la parenthèse dans les expressions de F et G, et, en dénommant S_1 et S_2 les spiraux associés et W_1 et W_2 leurs viroles diamétralement opposées (sans figures), on obtient:

1° suivant le diamètre commun des viroles W_1, W_2 et estimées positives dans le sens OW_2:

(a) la pression

$$F'_{12} = \frac{u}{P^2} \cdot \frac{4EI}{R^2};$$

(b) la pression

$$G'_{12} = -\frac{u}{P^3} \cdot \frac{4EI}{R^2} \cos u;$$

2° suivant le diamètre fixe des pitons: une force centrifuge:

$$\phi'_{12} = 4 \frac{EI}{R^2} \frac{u^2}{P^3}.$$

III

Si l'on se place ici au point de vue purement théorique, et c'est ce que je ferai pour abréger cette exposition, voici la méthode la plus simple pour éliminer les pressions transverses:

Je m'appuierai d'abord sur la remarque fondée sur le théorème suivant que j'ai signalé dès mes premières recherches sur le doublet sinusoïdal; deux ressorts hélicoïdaux co-axiaux développent des forces élastiques proportionnelles pourvu que ces spiraux associés sur un même balancier aient leurs caractéristiques unies par les relations

$$\left.\begin{array}{l} E = E', \\ \dfrac{I}{R^2} = \dfrac{I'}{R'^2} = \alpha, \qquad \dfrac{h}{R} = \dfrac{h'}{R'}, \\ P = P', \end{array}\right\} \text{spiraux semblables.}$$

Construisons, dès lors, deux doublets semblables autour du même axe d'oscillation avec même plan de virole, mais l'orientation du second doublet se déduisant par une rotation d'un demi-tour par rapport à l'autre.

Dès lors, les pressions transverses se détruisent sur les diamètres homologues renversés et l'extension de la similitude élastique nous donne ce résultat:

Ces quatre ressorts forment un quadruple sinusoïdal qui produit sur le balancier un couple longitudinal dont le bras de levier est $R+R'$ et dont l'intensité ou le moment mécanique du couple d'appui transversal est:

$$4a.\frac{u^2}{P^2}.\operatorname{tang} i.(R+R').E.m.2\pi$$

i étant l'inclinaison commune des spires à l'état naturel.

Si enfin, le quadruple précédent est répété à un étage différent du premier autour de l'axe commun d'oscillation, mais encore après une rotation d'un demi-tour suivie d'une translation, les *deux couples quadratiques* s'équilibrent.

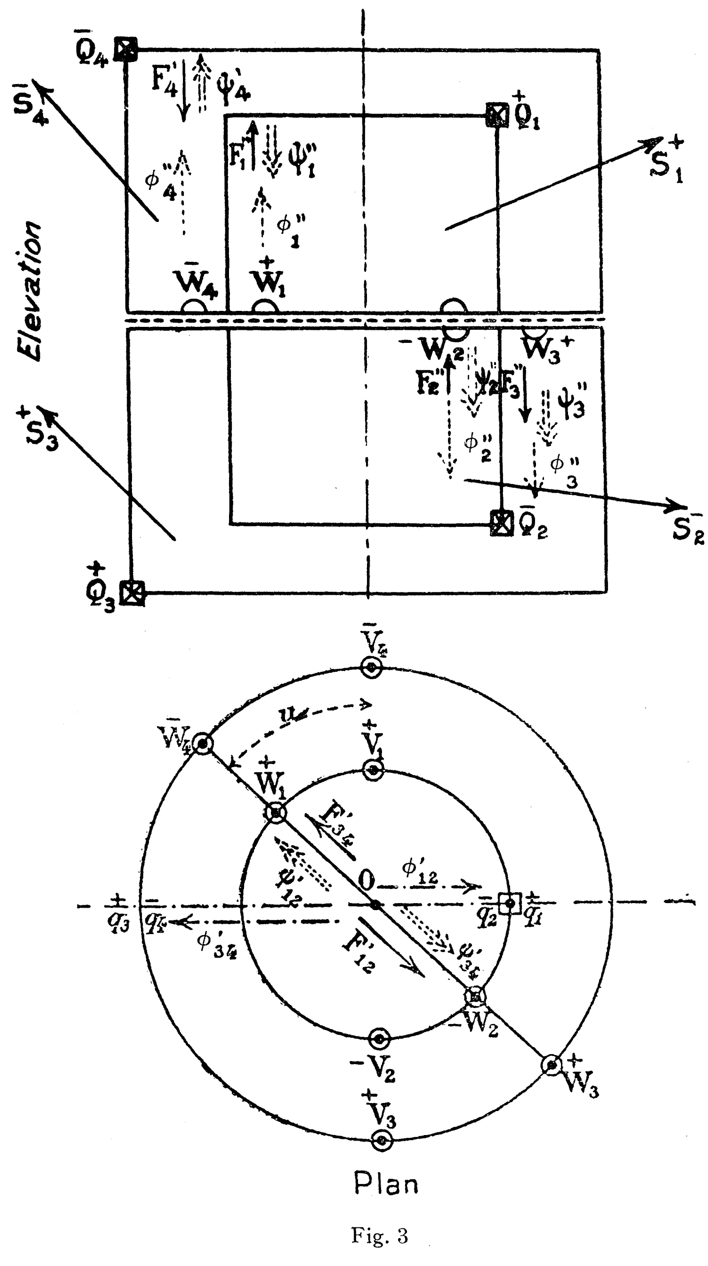


Fig. 3

Cette méthode, reposant sur la fabrication de ressorts de mêmes matières à échelles linéaires distantes, mais ayant des structures identiques dans les sens de leur enroulement hélicoïdal, sinon dans leurs ajustages, *n'est pas celle qui offre le plus de sécurité de fabrication; mais on peut, pour des horloges élastiques,* renoncer aux doublets co-axiaux, mais alors, il faut encore doubler la hauteur totale des logements. On utilise alors seize ressorts au lieu de huit.

En revanche, le premier quadruple formé tout à l'heure peut être conservé, car j'ai montré que: *le frottement quadratique* de notre premier quadruple jouit d'une précieuse propriété que voici:

Deux cas sont à distinguer:

1° ou bien ce frottement quadratique est le seul frottement résistant;

2° ou bien ce frottement quadratique accompagne un frottement constant.

Dans le premier cas, reportons-nous à la valeur du couple d'appui transversal trouvée plus haut; on peut l'écrire, vu la valeur de tang i, sous la nouvelle forme:

$$4\alpha\frac{u^2}{P^2}\operatorname{tang} i\,.\,(R+R')\,.\,Em\,.\,2\pi\,;$$

ceci est petit du 3^e ordre, puisque i est petit du premier ordre; mais la méthode de la variation des constantes appliquée au calcul *de la durée* de l'oscillation simple sous un régime normal de la semi-amplitude u_0 de l'oscillation, montre que la perturbation de la durée est petite de l'ordre σ, c'est-à-dire dépend de $\left(\frac{1}{P^3}\right)^2$ en raison de la forme paire de la fonction précédente de u si on suppose $m=$constante (hypothèse des techniciens) ou même que m soit une fonction paire de u; alors, l'intégrale donnant la valeur de cette perturbation sera très au-dessous de l'ordre absolu $\frac{1}{P^4}$ déjà négligé au début de l'évaluation du moment pendulaire de l'organe réglant.

En ce cas, si le centre de l'action de l'échappement est au point mort: $u=0$, la *vibration entretenue*, dans le cas 1°, reste isochrone à l'ordre de $\frac{1}{10^5}$ près, soit à l'ordre relatif pris, au-dessous du millionième.

Il nous reste maintenant à examiner le cas 2°.

L'équation du mouvement de l'oscillation prend la forme

$$\frac{d^2u}{dt^2}=-k^2(u+f)-k\lambda^2u^2;$$

λ restant petit au 3^e ordre, mais alors, posons $u=\gamma+z$; l'équation différentielle du mouvement prend la forme

$$\frac{d^2z}{dt^2}=-k'^2z-k^2\lambda(\gamma^2+2c\gamma+g)-k^2\lambda^2,$$

γ et g désignant deux constantes; j'ai montré que l'on peut alors choisir la constante γ de manière à satisfaire à l'équation du second degré

$$\gamma^2+2c\gamma+g=0.$$

Nous sommes alors ramenés au premier cas; si donc le centre d'action de l'échappement quasi-instantané est placé au point $u=\gamma$, ce point correspond à $z=0$ pour le mouvement de la variable z; la discussion de cette équation en γ montre qu'elle a deux racines, l'une petite admissible, et l'autre très grande à rejeter, et nous pouvons alors appliquer au mouvement en z la remarque faite dans le cas 1°. D'où l'important théorème suivant:

Dans le cas d'un mouvement pendulaire troublé par un frottement constant combiné avec un petit frottement quadratique, il suffit de décaler le centre d'action de l'échappement, par rapport au point mort des ressorts réglants, de manière à conserver l'approximation du millionième, *au mouvement vibratoire entretenu,* sous la semi-amplitude de régime du dit échappement.

Cette méthode constitue une précieuse généralisation d'une remarque de Villarceau, sur le frottement constant.

Dans le cas d'un frottement constant, l'échappement agissant au point mort décalé de f respecte l'isochronisme absolu de la vibration entretenue.

Dans le frottement quadratique, le décalage γ de l'échappement respecte l'isochronisme très approché de la vibration entretenue; ce décalage γ exige la connaissance préalable des deux coefficients f et λ du frottement constant et du frottement quadratique.

Ceci, soit dit en passant, *montre l'importance de mesures plus précises des coefficients de frottement*; j'ai donné pour celles-ci de nouvelles méthodes.

Les résultats obtenus par un quadruple de deux doublets co-axiaux de ressorts dont les enroulements de fabrication sont de même sens peuvent être obtenus aussi par l'emploi de deux ressorts homogènes de même sens, mais d'enroulements symétriques; mais nous n'entrerons pas ici dans des détails à cet égard. Ces détails seront mieux à leur place dans un volume qui va paraître très prochainement (*Horlogerie et chronométrie* chez J. B. Baillière, Paris).

Nous nous contenterons d'indiquer (Fig. 3) sur le quadruple formé de deux doublets sinusoïdaux co-axiaux, la disparition des pressions élastiques transverses F'_{12}, F'_{34} et de ψ'_{12}, ψ'_{34}, la disparition des pressions transverses ϕ'_{12} et ϕ'_{34}, la disparition des pressions longitudinales F_4', ψ_4'; F_3'', ψ_3'' et F_1'', ψ_1''; F_2'', ψ_2'', et l'apparition des couples longitudinaux quadratiques

$$\begin{cases} \phi_1'', \phi_2'', \\ \phi_4'', \phi_3''. \end{cases}$$

IV

Les frottements spontanément isochrones.

Ces curieux frottements sont liés indirectement à une nouvelle conception *des différents types d'axes d'oscillation*;

Je distingue trois types *d'axes d'oscillation, tous verticaux*:

1° axe, appuyé par son *pivot inférieur* contre un plan horizontal en pierre fine, nommé contre-pivot, mais *appuyé* latéralement par un collier principal *supérieur*, avec un jeu ou ébat très réduit;

2° axe à pivot supérieur, ayant sous lui un contre-pivot jeté en pont horizontal, et allégé par un flotteur *d'allégement*, destiné à réduire seulement la pression verticale due au poids du balancier; cet axe est appuyé latéralement par un *collier inférieur* avec un jeu ou ébat très réduit;

3° l'axe d'oscillation, *sans pivot*, mais appuyé latéralement sur deux colliers, à deux étages horizontaux distincts.

Cet axe est solidaire d'un flotteur complètement libre, dont la partie supérieure est solidaire d'une aiguille latéralement appuyée, avec jeu très réduit, contre un collier supérieur solidaire d'un bâti supérieur.

A la partie inférieure, l'axe lui-même porte une aiguille d'appui latéral sur le collier inférieur. Ces derniers axes qui permettent un mouvement longitudinal centré sur l'axe géométrique des colliers sont essentiellement destinés soit à des horloges élastiques de très haute précision, soit à des balances spirales d'une haute sensibilité et destinée à mesurer des couples naturels horizontaux.

Une condition essentielle de leur emploi, condition que l'on ne saurait trop mettre en évidence est la suivante: *les ressorts hélicoïdaux réglants*, qui produisent sur l'appareil un moment mécanique *strictement pendulaire* à l'approximation du *millionième*, doivent comporter un équilibre des pressions élastiques longitudinales strictement assuré, surtout si l'emploi de la balance impose au flotteur de révolution des dimensions transverses relativement restreintes.

Après avoir défini, avec précision, nos trois types d'axes d'oscillation, au point de vue de leurs réalisations expérimentales, et en observant toutefois que, seul, le premier type a été réalisé et n'est autre que le type des chronomètres marins (sauf la prévision nouvelle de ressorts réglants associés) nous pouvons aborder avec une égale simplicité la définition (maintenant purement mathématique) de nos diverses classes de frottements isochrones.

1° Le mouvement *pendulaire troublé* par un *frottement constant*, c'est-à-dire dont *le moment mécanique résistant*, par rapport à l'axe d'oscillation, *est constant.*

Ce frottement a été signalé pour la première fois, avec ses précieuses propriétés, par Villarceau (Annales de l'Observatoire de Paris):

Enonçons-les en les complétant.

Tant que les oscillations ne sont pas encore éteintes, elles sont rigoureusement isochrones; leurs semi-amplitudes, comptées à partir de la position du balancier dite au point mort (celle où le moment pendulaire du système réglant est nul) ont leurs valeurs absolues consécutives, décroissant en progression arithmétique, et cette propriété constitue la meilleure mesure des frottements constants, *par l'inscription photographique des oscillations.*

De plus, si l'entretien du mouvement a lieu par *chocs instantanés, mais pourvu que le choc d'entretien soit donné,* non plus en la position du balancier qui correspond au point mort du système réglant, mais en une position *convenablement décalée* et un peu *antérieure* au point mort du système réglant, position que, seule l'inscription photographique permet de définir dans un azimut précis du système oscillant.

Le meilleur dispositif expérimental consiste à produire *le choc réparateur d'entretien par un couple d'électro-aimants*, de manière que la réalisation d'un choc instantané par la réalisation *d'un couple mécanique*, électriquement produit n'offre aucun inconvénient, comme dans l'entretien du mouvement par des rouages d'entretien.

2° *Le mouvement pendulaire troublé par une résistance visqueuse* (au sens du langage de Cornu), c'est-à-dire par une résistance proportionnelle à la vitesse.

Si cette résistance est unique, les semi-amplitudes des oscillations simples $u_0, u_1, u_2, \ldots, u_n$ décroissent en progression géométrique, en sorte que $u_0 = u_n . Q^n$, le nombre $Q > 1$ étant lié au coefficient expérimental q de la viscosité;

3° Le mouvement pendulaire troublé, à la fois par un frottement constant, et par une résistance visqueuse.

J'ai fait connaître, dans mon petit volume dejà cité, les formules très simples qui permettent l'étude photographique des semi-amplitudes successives: $u_0, u_1, u_2, \ldots, u_n$ de cette 3[e] catégorie de frottements isochrones, laissant les vibrations rigoureusement isochrones, jusqu'à l'extinction complète du mouvement après N oscillations simples.

Ces formules se résument en la suivante:

$$(u_0+\phi) = (u_n+\phi)Q^n,$$

d'où j'ai déduit la valeur maximum N de n en fonction de u_0 lorsque n oscillations sont achevées; la constante ϕ positive désigne une constante dépendant à la fois du coefficient de viscosité et du coefficient de frottement.

Quant à la constante Q c'est la même que celle qui correspond à la viscosité du liquide flotteur, en admettant toutefois que la viscosité n'est pas accompagnée d'une friction humide dont nous dirons plus loin quelques mots.

Dans le cas d'un axe à flotteur du 2[e] type indiqué plus haut, mais où le liquide du flotteur étant vidé, le frottement s'exerce sur la tête du pivot du balancier, il peut y avoir une variation de la pression génératrice du frottement, en raison d'une pression longitudinale électrique modifiant le poids de charge du balancier.

En ce cas, le mouvement pendulaire troublé a une équation du type:

$$\frac{d^2u}{dt^2} = -k^2u - k^2(f+\alpha u)\epsilon$$

α étant une constante de signe déterminé, et ϵ désignant

$$sgn . \frac{du}{dt} = \pm 1.$$

Ce mouvement est rigoureusement isochrone, bien que les portions d'une même oscillation simple n'aient pas même durée.

4° Dans le cas où la pression proviendrait d'un frottement sur un collier, frottement latéral dû à une pression transverse élastique proportionnelle à u

en valeur absolue et accompagnant d'ailleurs ou non un frottement de viscosité, on obtient dans ces deux cas une oscillation simple dont l'équation différentielle change de forme dans les deux moitiés de l'oscillation simple envisagée, et est:

$$\frac{1}{k^2} \cdot \frac{d^2u}{dt^2} = -u - \eta\epsilon\lambda u - q\frac{du}{dt}\ ;\ \eta = sgn.\, u,\ \epsilon = sgn.\, \frac{du}{dt},$$

λ et q étant deux constantes positives: u_0 étant la semi-amplitude initiale de l'oscillation en cours et u_1 étant la semi-amplitude finale de cette même oscillation, l'on trouve

$$u_1 = u_0 Q' . \sqrt{\frac{1-\lambda}{1+\lambda}},$$

Q' désignant une constante moindre que 1.

Ce cas serait réalisé pour un axe à flotteur dont la friction humide se réduit à la simple viscosité. Et la durée d'une oscillation simple est rigoureusement isochrone.

5° La combinaison des cas 3° et 4° réunis, ne parait pas devoir être spontanément isochrone, je veux dire en mouvements vibratoires non entretenus.

V

Comment sera tranchée, expérimentalement la réponse à l'hypothèse des techniciens?

En attendant, énoncé d'un problème posé à nos jeunes géomètres:

En attendant qu'un nouvel Huygens ait trouvé, pour les horloges élastiques la solution complète que le calcul intégral exige, *une* solution pratique est déjà réalisée; nous souhaitons vivement que la solution complète la plus générale soit prochainement apportée; c'est dans cet espoir que je porte le problème général devant le Congrès International Mathématique de Toronto.

Pour ne pas allonger outre mesure cet exposé d'un problème d'ailleurs urgent pour des réalisations techniques prochaines, j'ai, pour la briéveté et la simplicité de l'exposition d'une solution déjà acquise, insisté tout d'abord sur un moyen d'obtenir, au moyen d'un groupement co-axial de deux doublets sinusoïdaux concentriques, l'équilibre préalable de toutes les pressions élastiques transverses, transmises par le système réglant au balancier.

Au lieu d'adopter cette réduction fondamentale préalable, fondée sur l'hypothèse des techniciens, on peut prendre au contraire comme doublet fondamental un doublet non plus sinusoïdal, mais symétrique; cette symétrie des deux ressorts réglants du doublet étant d'ailleurs susceptible de deux formes:

ou bien une symétrie des deux doublets par rapport au plan transverse commun de leurs viroles;

ou bien une symétrie complétant la symétrie précédente par une rotation d'un demi-tour relatif du second doublet par rapport au premier, ce qui revient à dire encore que la symétrie des deux ressorts du doublet est une symétrie par rapport à un centre de symétrie coïncidant avec le pied de l'axe d'oscillation sur le plan transverse commun des viroles.

Cette solution a l'avantage de se débarrasser de *l'hypothèse des techniciens*, mais elle a l'inconvénient d'exiger l'emploi de ressorts hélicoïdaux d'échelles linéaires égales, mais à enroulements symétriques, ce qui, au point de vue de l'exécution, parait pouvoir compromettre quelque peu les sécurités de fabrication.

Au contraire, la solution particulière que j'ai sommairement exposée, contient des échelles linéaires distinctes, mais, comme je l'ai dit, il est vrai, qu'on peut supprimer la disposition co-axiale, sauf à multiplier les étagements, en sorte que l'on n'emploie que des fractions d'un ressort primitivement unique; mais cet avantage parait payé de la subordination de nos réalisations à la validité affirmée de l'hypothèse des techniciens; cet avantage est peut être payé trop cher.

L'avantage d'adopter comme élément de doublet le groupement symétrique (à viroles opposées) nous débarrasse, semble-t-il, de l'obligation de nous plier à l'hypothèse encore injustifiée que j'ai appelée l'hypothèse des techniciens, mais il a l'inconvénient d'une fabrication par doublets à enroulements symétriques l'un de l'autre, fabrication à peine encore en usage.

Dans ces conditions, il y a un problème de géométrie infinitésimale qui me parait mériter l'attention de nos jeunes géomètres familiers à la fois avec l'emploi des coordonnées curvilignes et avec le problème de Saint-Venant; ce problème, je l'énoncerais volontiers ainsi: chercher, pour un ruban hélicoïdal unique, soit un mode de généralisation du problème de Saint-Venant, soit une solution de la déformation du ruban hélicoïdal, qui puisse s'accommoder avec une approximation suffisante de la partie du problème déjà résolu par Résal et continué par Caspari, à savoir: en admettant que la projection de la *fibre moyenne* du ressort hélicoïdal ait sa projection sur un plan perpendiculaire à l'axe d'oscillation du balancier auquel est attelé le ruban coïncidant avec la courbe définie par la méthode de Résal-Caspari, compléter le problème ainsi amorcé, au point de vue des déformations élastiques longitudinales des éléments du ressort hélicoïdal.

En attendant une solution du problème ainsi posé, et que je crois digne de provoquer les efforts de nos jeunes géomètres, je tiens à dire que dans un volume prêt à paraître* j'ai amorcé une méthode expérimentale qui va me permettre de continuer l'étude commencée au laboratoire sur ce problème nouveau de mécanique physique et qui rapproche deux problèmes en apparence très différents, mais au fond réductibles l'un à l'autre, le problème des horloges élastiques isochrones au millionième, et le problème des balances spirales dont la construction intéresse aujourd'hui d'importants problèmes sur les frictions tangentielles entre liquide et solide flotteur de révolution autour de son axe d'oscillation.

*Chez J. Baillière, Paris, 19 rue Hautefeuille, *Horlogerie et Chronométrie.*

A SYSTEM OF "DEFINITIVE UNITS" PROPOSED FOR UNIVERSAL USE

By Dr. G. A. Campbell,
American Telephone and Telegraph Company, New York, N.Y., U.S.A.

Introduction

It is proposed that physicists discard the C.G.S. and Heaviside units, which are responsible for much confusion and needless mental effort, and employ exclusively a single system of "definitive units". The units suggested include the international meter, the international kilogram, the second, the mechanical watt, the international mercury ohm; and they conform with the other international units as closely as is compatible with self-consistency. The system is not only definite and absolute, but also comprehensive, readily visualized and, in large part, already employed under well-established names. The system was proposed in 1901 by Giorgi and in 1904 by Robertson but received scant attention on account of the artificial prestige of the C.G.S. system. A single universals ystem of units should be the ultimate goal; and the "definitive units" are chosen with a view to their adoption not only by all scientists but also by the butcher, the baker and the candlestick-maker.

The Problem of Universal Units.

The purpose of this paper is to advocate the use of a single system of units meeting these two essential requirements:

1. Universal applicability for all purposes from the common everyday practical needs to the most exacting scientific uses.
2. Ready transition to the single system involving minimum change and inconvenience.

It is unnecessary here to stress the importance of a common meeting ground for all commercial countries in the matter of units. The existing confusion and waste due to the multiple, redundant units which are still in use is quite generally recognized, and progress is being made towards universal units. It is most important that a proposed universal system of units should not call for radical changes; otherwise, its chance of securing adoption is small. For the proposed system what has been done has been to choose from among the multiplicity of existing units, the ones best adapted to form a single, comprehensive and final system for universal use. These units are shown by Table I in the column headed "Definitive Unit"; each unit, with the exception of the (10^5 dyne) unit for force, is now in common use. Furthermore, the international prototype meter and kilogram became the fundamental standards of length and mass in the United States in 1893, while eight of the remaining units (or their international electrical equivalents) were legally established by law in 1894. It is believed that the proposed "definitive" system may be advantageously adopted by individuals, and without inconvenience, even in the absence of further legislative or other general adoption.

Even though we recognize the difficulties in the way of securing the universal adoption of any one system of units, it is nevertheless extremely worth while to determine the best system of units for this purpose, and to secure for it the widest possible adoption as rapidly as possible. It is especially important that physicists should show the way by exclusively employing the system adapted for universal use, before custom and legislation have established additional units which cannot be absorbed into such a system. This change will directly benefit physicists themselves, for they are now handicapped by the intermingled and varying use of many systems of units, including the C.G.S. electrostatic, the C.G.S. electromagnetic, the practical, the international, the rational, the ampere-turn and the gravitational systems in addition to the metric system. Upon the first four of these systems the British Association Committee on Electrical Standards placed the seal of its approval without, apparently, any serious consideration of the possibility and desirability of a single universal system. The C.G.S. system has served a great purpose. It was a long step towards complete unification, but it was not intended to be and is not adapted to be the single universal system, and for that reason it should give place to a truly universal system as soon as possible.

It is important that the subject of units should now be considered from a comprehensive, universal point of view because of possible legislation extending the use of the metric system. Fortunately scientists are in a position, on the basis of a long and essentially successful experience with metric, absolute C.G.S. and other units, to specify the complete system of units which may be confidently fixed as the universal goal. All changes which are introduced should be consistent with and aid in bringing about the ultimate adoption of a universal system of units. Proposals looking toward the adoption of the metric system in the United States and elsewhere should consider the question of units broadly, and not merely adopt the meter-liter-gram units because these were the original metric units. As shown by Table I, the kilogram is required for consistent interrelation with the meter, second and watt. Furthermore, the consistent unit of volume is the stere or meter cube, which is almost exactly a kiloliter.

It is important that the proposed universal system of units bear a simple, distinctive name, especially since many of the individual units brought together in the system differ little, if at all, from units which are already associated with one or more existing systems of units. The name proposed is "definitive", and this name is appropriate because these units are the best final selection from among all metric possibilities. Justification for the name does not require proof of absolute finality; it is recognized that with the advance of science some new basis for units superior to the metric basis may be discovered. In that event the definitive units should be superseded. In the meantime, however, every worker in science, engineering, industry, commerce and every day affairs may confidently turn to the definitive system for his units.

Proposed Definitive Units.

The proposed units are shown in Table I, together with the conversion factors for reducing several other systems of units to definitive units. The meter and the second are the two units for the space time framework; the watt and the joule are the units for power and energy, the essential binding links in all physical transformations; the (10^5 dyne) and the (kilogram) are the dynamical units for force and mass. The definitive system includes the international prototype meter, the international prototype kilogram, the international mercury ohm, the international farad and the international henry. The definitive watt, however, is the mechanical, not the international, watt, the latter being 340 parts in a million larger. The definitive ampere, volt and coulomb are, consequently, each smaller than the corresponding international unit by 170 parts in a million. These differences are negligible except in refined scientific work.

The definitive units are so simple and convenient that it is surprising that they were not adopted originally in place of the C.G.S. units; as far as I know no extended use has ever been made of them as a comprehensive system. The system was proposed, however, in 1901 by Giorgi under the name of the "absolute practical" system, and in 1904 by Robertson as the "complete practical" system. The solid advantages which they pointed out were almost ignored, due to conservatism and the artificial prestige of the C.G.S. system. Emde recommended the adoption of Giorgi's proposal but only as a temporary transition system. The absolute C.G.S. units have continued to be regarded as the fundamental, complete system of units to be extensively supplemented in practice, nevertheless, by the intermingled use of metric, practical, rational, international and ampere-turn units. No uniformity of practice has prevailed, the units from different systems being combined by different authors in different ways.

A consistent system employed to a very limited extent in electrical engineering satisfies the connecting relations of the definitive units but makes use of the centimeter-(10^7 dyne)-(10^4 kilogram) in place of the meter-(10^5 dyne)-(kilogram). For a universal system, however, the meter is to be preferred to the centimeter and the kilogram to the (10^4 kilogram). The meter and the kilogram are the units now in actual use in commerce and practical life in metric countries; the centimeter is recognized as a basic unit by scientists but not to the extent of influencing the nomenclature; 10,000 kilograms as an every-day unit of mass is much too large since it lies far beyond ordinary sense experience. No other possible combination of metric multiples is to be preferred, but it may be noted that the combination of the dekameter-(10^4 dyne)-10 gram) has the attractive feature of making weight and mass numerically equal to each other within 2 per cent.

Advantages of Definitive Units.

The proposed definitive units combine, to a remarkable degree, the best features of all existing systems of units. The following may be specially noted:

1. Definitive units are independent of place and time. They are absolute in this and every other respect and the peers of the C.G.S. units for the most refined scientific work.
2. Readily visualized magnitudes of convenient size for every-day needs characterize the definitive units of the more important quantities. The meter and kilogram are the common metric units of length and mass. The definitive unit of force (10^5dyne), which is approximately the force of gravity on 0.1 kilogram, is a force which is directly apprehended; whereas the dyne, the C.G.S. unit, corresponds more nearly to the proverbially negligible weight of a feather.
3. Convenient names are already in common international use for most definitive units or their practical equivalents. Eventually a few new names should be adopted so that no basis definitive unit will be encumbered with a numerical prefix as in (10^5dyne), (kilogram) and (centare). Names are quite secondary to the adoption of the universal system. Apparently the adoption of definitive units would be hindered rather than helped by the simultaneous consideration of new names for units, if we are to judge by the discussion elicited by Giorgi and Robertson twenty years ago.
4. A complete, coherent system of definitive units is obtained by so extending Table I as to avoid useless coefficients, in essentially the same way as for C.G.S. units. The complete table would include units for area (centare), volume (stere$=1.000027^{-1}$ kiloliter), velocity (meter/second), electric force (volt—meter), temperature, entropy, etc.
5. The entire range of each quantity encountered throughout all nature is conveniently referred to the single corresponding definitive unit by the use of numerical prefixes. These prefixes are to be regarded as a part of the numeric and not as creating an independent basic unit. Arithmetical notation supplies different methods of expressing these prefixes, and further developments are doubtless possible. Table II illustrates the use of powers of mega and micro prefixes to cover the entire range of nature in categories of a million; the known range of lengths extends over only seven of these categories. As far as I have looked into the matter twelve of these categories suffice for any one of the quantities of physics as known to-day.
6. The minimum changes in commerce, in legislation and in scientific instruments are called for by definitive units since the largest possible number of international units are retained.
7. Conversion of existing results into definitive units is made simple by the use of Table I. The table may be extended so as to include English units and the algebraical as well as the arithmetical relations between units. A grounding in the definitive units is sufficient for everyone; familiarity with the C.G.S. units or with the historical role which they played as progenitors of definitive units is not essential. Results expressed in the older units are readily reduced to definitive units by the mechanical use of the table.

8. Definitive units do not artificially exalt any four units as being pre-eminently the basic primary units for all purposes. Any choice may be made which will simplify a particular problem under discussion. Among the first twelve units of Table I, four dimensionally independent units may be chosen in 299 different ways.
9. The ten fundamental relations of Table I, Footnote 2, hold also for Heaviside's units. Since the only present use for Heaviside's units is to obtain these relations and others based upon them, his specific units may be abolished, thereby materially simplifying the connection between theoretical electromagnetism and practical measurements.
10. Complete, homogeneous, physical equations (which are best adapted for theoretical physics) are naturally employed with definitive units, since no dimensional constant of nature has by advance agreement any specific numerical value, such as unity or 4π. Some of the more important dimensional constants, expressed in definitive units, are included in Table II. The choice of units need introduce no changes in the use of convenient ratios, such as permeability and dielectric constant.

Conclusions

The definitive system of units makes it perfectly feasible to employ a single system much more generally than has ever been the case in the past; the natural ultimate goal is the universal use of these units for all purposes. In the attempt to extend the application of metric units in the United States the meter-stere-kilogram, rather than the meter-liter-gram, should form the basis for legislation, in order to give definitive units their proper legalized status, and to secure to the full the advantages of a comprehensive system consistently inter-related with the legalized international electrical units. The gradual discarding of C.G.S. and other redundant units would inevitably follow. Even in the absence of official recognition of definitive units, individuals may advantageously employ the system. This would cause confusion neither to authors nor readers, since these units, in the main, have already acquired vital existence through world-wide, daily use under familiar, well-established names.

TABLE I—FACTORS FOR CONVERSION INTO DEFINITIVE UNITS[1]

Quantity	Definitive Unit[2] and Symbol	Value of Unit in Terms of Definitive Unit[3]					
		International Unit	Practical Unit	C.G.S. Electromagnetic Unit	C.G.S. Electrostatic Unit	Gaussian-Heaviside-Lorentz Unit	Mechanical Metric K-M-S Gravitational Unit
Length	s Meter	1.	10^7	0.01	0.01	0.01	1
Time	t Second	1.	1.	1.	1.	1	1
Power	P Watt	1.00034	1.	10^{-7}	10^{-7}	10^{-7}	0.1019716^{-1}
Energy	W Joule	1.00034	1.	10^{-7}	10^{-7}	10^{-7}	0.1010716^{-1}
Force	F (10^5dyne)	....	....	10^{-5}	10^{-5}	10^{-5}	0.1019716^{-1}
Mass	M (Kilogram)	10^{-3}	....	10^{-3}	10^{-3}	10^{-3}	1
Resistance	R Ohm	1.	1.00052^{-1}	$1.00052^{-1}10^{-9}$	$1.1130^{-1}10^{12}$	1.1290×10^{13}	
Current	I Ampere	1.00017	1.00026	1.00026×10	$2.9974^{-1}10^{-9}$	$1.0626^{-1}10^{-10}$	
Potential	V Volt	1.00017	1.00026^{-1}	$1.00026^{-1}10^{-8}$	2.9974×10^{2}	1.0626×10^{3}	
Electricity	Q Coulomb	1.00017	1.00026	1.00026×10	$2.9974^{-1}10^{-9}$	$1.0626^{-1}10^{-10}$	
Capacity	C Farad	1.	1.00052	1.00052×10^{9}	1.1130×10^{-12}	$1.1290^{-1}10^{-13}$	
Inductance	L Henry	1.	1.00052^{-1}	$1.00052^{-1}10^{-9}$	$1.1130^{-1}10^{12}$	1.1290×10^{13}	
Magnetomotive force	F (Ampere-turn)	....	....	1.25631^{-1}	....	2.82168	
Magnetic flux	ϕ (Weber)	....	....	$1.00026^{-1}10^{-8}$	....	$2.82168^{-1}10^{-7}$	
Magnetic pole	m (Weber)	....	....	1.25631×10^{-7}	....	$2.82168^{-1}10^{-7}$	
Reluctance	R (Amp.-turn/weber)	....	....	$1.25598^{-1}10^{8}$	....	$1.25598^{-1}10^{8}$	

[1] A numerical magnitude based upon any unit is transformed to the corresponding definitive unit by multiplying by the tabulated value of the original unit. A literal relation $f(A, B, C\ldots)=0$ transformed to $f(A/a, B/b, C/c\ldots)=0$ holds for definitive units, a, b, c being the tabulated values of the original units for A, B, C, respectively.

[2] The definitive units are connected by the ten simple, fundamental energy relations:

$$W=Pt=Fs=\tfrac{1}{2}M\dot{s}^2=RI^2t=VIt=\tfrac{1}{2}CV^2=\tfrac{1}{2}QV=\tfrac{1}{2}LI^2=\tfrac{1}{2}F\phi=\tfrac{1}{2}R\phi^2$$

[3] Based upon the experimental constants: velocity of light $=v=299.82\times10^6$ meter/sec.; acceleration due to gravity $=g=9.80665=0.1019716^{-1}$ meter/sec.^{2}; international ohm $=r=1.00052$ "practical" ohm; international ampere $=a=0.99991$ "practical" ampere. Therefore, $a^2r=1.00034=1.00017^2$; $v^2r^{-1}=1.1130^{-1}10^{17}=2.9974^2\,10^{16}$; $4\pi v^2r^{-1}=1.1290\times10^{18}=1.0626^2\times10^{18}$; $r/4\pi=12.5598^{-1}=0.282168^2$; $4\pi^{-\frac{1}{2}}=12.5631$; $r=1.00026^2$, from which follow the tabulated values.

TABLE II. PHYSICAL CONSTANTS IN TERMS OF DEFINITIVE UNITS

	Numeric	*Unit*
Diameter of positive electron	2.0 micro^{3}	meter
Diameter of negative electron	3.8 micro$^{2.5}$	"
Wave length shortest γ-ray	5.7 micro^{2}	"
Wave length cadmium red No. 1	643.84696 micro$^{1.5}$	"
Earth to sun	0.1495 mega^{2}	"
Light year	9.4614 mega$^{2.5}$	"
Parsec	30.838 mega$^{2.5}$	"
Radius curvature of space (Silberstein)	1.0 mega^{4}	"
Mean solar year	31.556926 mega	second
Velocity of light in free space	299.82 mega	(m./sec.)
Capacity of meter cube of free space	8.8572 micro^{2}	farad
Inductance of meter cube of free space	1.25598 micro	henry
Iterative impedance of free space	376.57	ohm
Resistance of one kilogram of mercury in a column one meter long	12.7898 milli	"
Force between two kilogram point masses separated one meter	66.58 micro^{2}	(10^{5}dyne)
Force between two coulomb point charges separated one meter	8.9845 mega$^{1.5}$	"
Force between two weber point poles separated one meter	63.359 kilo	"
Charge of negative electron	0.1593 micro^{3}	coulomb
Mass of negative electron	0.89 micro^{5}	kilogram
Mass of sun	2.0 mega^{5}	"

SUR L'APPLICATION DES MÉTHODES DU CALCUL TENSORIEL A LA THÉORIE DES MOINDRES CARRÉS

PAR M. J. HAAG,

Professeur à l'Université de Clermont-Ferrand, Clermont-Ferrand, France.

1. Nous ne nous occuperons nullement, dans le présent travail, de la justification de la méthode classique des moindres carrés. Nous voulons seulement montrer comment les procédés du calcul tensoriel, qui se sont surtout répandus ces dernières années, grâce à la théorie de la relativité, permettent d'en simplifier considérablement l'exposition.

2. NOTATIONS. Soient $x_1, x_2, \ldots, x_n$ les inconnues; nous les regarderons comme des variables *covariantes*. Les observations directes, que nous supposerons toutes également précises,* donnent les valeurs b_λ de $m=n+p$ formes linéaires†

(1) $$u_\lambda = a_\lambda^i x_i; \quad \lambda=1, 2, \ldots, m; \quad i=1, 2, \ldots, n;$$

dont les coefficients a_λ^i sont exactement connus et vont être regardés comme des variables *contrevariantes*; de sorte que chaque forme u_λ sera un *invariant.* Nous faisons, dès maintenant, la convention que les indices grecs se rapporteront toujours au rang des observations et ne seront jamais des indices de covariance.

La forme quadratique

(2) $$\Theta(x) = \sum_\lambda u_\lambda^2$$

est aussi un invariant. Nous l'appellerons la *forme quadratique fondamentale* et nous supposerons qu'elle est définie.‡ Si l'on pose

(3) $$A^{ij} = \sum_\lambda a_\lambda^i a_\lambda^j,$$

elle s'écrit

$$\Theta(x) = A^{ij} x_i x_j.$$

A tout système de variables covariantes x_i, nous ferons correspondre les *variables adjointes contrevariantes*

$$x^i = \frac{1}{2}\frac{\partial \Theta}{\partial x_i} = A^{ij} x_j;$$

*On peut toujours se ramener à ce cas, en multipliant les u_λ par des facteurs convenables.

†Nous adoptons la convention bien connue, d'après laquelle tout indice figurant deux fois dans la même expression doit être considéré comme un indice *muet*, c'est-à-dire qu'on devra le remplacer successivement par $1, 2, \ldots, n$ et faire la somme des quantités obtenues.

‡Cela équivaut à supposer que, parmi les m formes u_λ, n sont indépendantes. S'il n'en était pas ainsi, les inconnues seraient algébriquement indéterminées.

de sorte que l'on a aussi

$$\Theta(x) = x^i x_i = A_{ij} x^i x^j.$$

Nous désignerons par $\Theta(x, y)$ la *forme polaire*:

$$\Theta(x, y) = x^i y_i = x_i y^i.$$

3. Identités. Nous allons établir tout de suite quelques propriétés importantes des coefficients a_λ. On a

$$\sum_\lambda a^i_\lambda a_{\lambda j} = \sum_\lambda a^i_\lambda a^k_\lambda A_{kj} = A^{ik} A_{kj} = A^i_j, \tag{4}$$

ce qui est égal à zéro, si $i \neq j$ et à un, si $i = j$. D'où l'on déduit immédiatement la formule

$$\sum_\lambda \Theta(a_\lambda) = \sum_\lambda a^i_\lambda a_{\lambda i} = n. \tag{5}$$

Élevant au carré et posant*

$$A = \sum_\lambda \Theta^2(a_\lambda), \tag{6}$$

il vient

$$\sum_{\alpha, \beta} \Theta(a_\alpha)\Theta(a_\beta) = \tfrac{1}{2}(n^2 - A), \tag{7}$$

la sommation portant sur les combinaisons d'indices différents.

Nous avons maintenant, en remplaçant x_i par $a_{\alpha i}$ dans (2),

$$\Theta(a_\alpha) = \sum_\lambda (a^i_\lambda a_{\alpha i})^2 = \sum_\lambda \Theta^2(a_\lambda, a_\alpha). \tag{8}$$

En sommant par rapport à α et tenant compte de (5), il vient

$$\sum_{\alpha, \beta} \Theta^2(a_\alpha, a_\beta) = \tfrac{1}{2}(n - A). \tag{9}$$

4. Valeurs probables des inconnues. On sait qu'elles sont obtenues en rendant minimum la quantité†

$$\theta = \sum_\lambda (u_\lambda - b_\lambda)^2 = \Theta - 2\sum_\lambda b_\lambda a^i_\lambda x_i + \sum_\lambda b^2_\lambda. \tag{10}$$

*Les relations qui vont suivre, jusqu'à la fin du présent paragraphe, ne sont utiles que pour le paragraphe 8 et pourraient être laissées de côté, si l'on s'en tenait à l'essentiel de la théorie.

†Si l'on admet la loi de Gauss pour les erreurs, on a, par exemple,

$$\bar{x}_i = K \int x_i e^{-k\theta} dx_1 dx_2 \ldots dx_n,$$

K et k désignant deux constantes et le champ d'intégration étant infini. Soient $\alpha_1, \alpha_2, \ldots, \alpha_n$ les coordonnées du centre de l'hyperquadrique $\theta = 0$. Portons l'origine en ce point et soient $X_1, X_2, \ldots X_n$ les nouvelles coordonnées. Il vient

$$\bar{x}_i = K \int (\alpha_i + X_i) e^{-k(\Theta + C)} dX_1 dX_2 \ldots dX_n = \alpha_i,$$

car le terme en X_i disparait évidemment et le coefficient de α_i égale 1, d'après le principe des probabilités totales. Cf. H. Poincaré, *Calcul des probabilités*, page 239.

En annulant $\dfrac{\partial\theta}{\partial x_i}$, nous obtenons immédiatement*

(11) $$\bar{x}^i = \sum_\lambda b_\lambda a^i_\lambda,$$

d'où

(12) $$\bar{x}_i = \sum_\lambda b_\lambda a_{\lambda i}.$$

5. Valeurs probables des carrés des erreurs. Soit, d'une manière générale, l'expression linéaire

(13) $$f = g^i x_i + \sum_\lambda c_\lambda b_\lambda,$$

où les c_λ sont des coefficients donnés. Sa valeur probable est, d'après (12),

$$\bar{f} = \sum_\lambda b_\lambda (g^i a_{\lambda i} + c_\lambda).$$

Comme les b_λ sont déterminés par des observations *indépendantes et de même poids*, le carré de l'erreur sur $\bar{f}$ a pour valeur probable

(14) $$E(f) = h \sum_\lambda (g^i a_{\lambda i} + c_\lambda)^2 = h[\Theta(g) + \sum_\lambda c^2_\lambda + 2\sum_\lambda c_\lambda \Theta(g, a_\lambda)],$$

h désignant la valeur probable *a priori* du carré de l'erreur commise dans chaque observation.

En particulier,

(15) $$E(x_i) = hA_{ii};$$

(16) $$E(u_\lambda - b_\lambda) = h[\Theta(a_\lambda) + 1 - 2\Theta(a_\lambda)] = h[1 - \Theta(a_\lambda)].$$

6. Précision des observations. La quantité h ci-dessus introduite est la valeur probable *a priori* du carré de chaque erreur d'observation. Quand on ne la connaît pas, on calcule sa valeur probable *a posteriori*, en utilisant les résultats mêmes de la méthode des moindres carrés.

On appelle *résidus* les quantités

$$r_\lambda = a^i_\lambda \, \bar{x}_i - b_\lambda = \bar{u}_\lambda - b_\lambda.$$

Si les observations étaient rigoureusement exactes, ces résidus seraient nuls. Il s'ensuit que r_λ, par exemple, peut être considéré comme l'erreur commise sur $u_\lambda - b_\lambda$. Par conséquent, la valeur probable *a priori* de r^2_λ n'est autre que $E(u_\lambda - b_\lambda)$, donné par (16). Si l'on pose

$$S = \sum_\lambda r^2_\lambda,$$

on a donc, d'après (5),

(17) $$\bar{S} = h(m-n) = hp.$$

*La valeur probable d'une quantité sera distinguée par un trait la surlignant, suivant une notation fréquemment employée.

D'où l'on déduit la règle classique qui consiste à calculer h comme le *quotient de la somme des carrés des résidus par le nombre des équations surabondantes.*

7. La somme S est évidemment le minimum de la fonction θ. On peut l'écrire sous différentes formes. Nous avons d'abord, d'après (10) et (11),

$$S=\Theta(\bar{x})-2\bar{x}^i\bar{x}_i+\sum_\lambda b_\lambda^2=\sum_\lambda b_\lambda^2-\bar{x}^i\bar{x}_i=\sum_\lambda b_\lambda^2-\Theta(\bar{x}). \tag{18}$$

On a aussi, en utilisant (11) et (1),

$$S=\sum_\lambda(b_\lambda^2-\bar{u}_\lambda b_\lambda)=-\sum_\lambda b_\lambda r_\lambda. \tag{19}$$

Nous possédons maintenant toutes les formules classiques essentielles de la théorie. Nous allons les compléter par la résolution de quelques questions supplémentaires traitées par Gauss*.

8. Erreur a craindre dans le calcul de la précision†. Cherchons la valeur probable V du carré de l'erreur commise sur h quand, dans la formule (17), on remplace $\bar{S}$ par la valeur numérique effectivement trouvée pour S. On a

$$p^2V=\overline{(S-ph)^2}=\overline{S^2}+p^2h^2-2ph\bar{S}=\overline{S^2}-p^2h^2.$$

Nous avons maintenant, d'après (18),

$$S=\sum_\lambda b_\lambda^2[1-\Theta(a_\lambda)]-2\sum_{\alpha,\beta}b_\alpha b_\beta\Theta(a_\alpha,a_\beta).$$

Elevons au carré et prenons les valeurs probables. En appelant H la valeur probable de la quatrième puissance de l'erreur, nous avons d'abord

$$H\sum_\lambda[1-\Theta(a_\lambda)]^2=H(m-2n+A),$$

d'après (5) et (6).

Nous avons maintenant la valeur probable de

$$4\sum_{\alpha,\beta}b_\alpha^2b_\beta^2\Theta^2(a_\alpha,a_\beta)$$

qui est $2h^2(n-A)$, d'après (9).

Nous avons enfin la valeur probable de

$$2\sum_{\alpha,\beta}b_\alpha^2b_\beta^2[1-\Theta(a_\alpha)]\,[1-\Theta(a_\beta)],$$

qui est $h^2[m(m-1)+n^2-A-2(m-1)n]$, en utilisant (7) et (5). Finalement,

$$p^2V=H(p-n+A)+h^2(3n-p-3A);$$

$$V=\frac{H-h^2}{p}+\frac{n-A}{p^2}(3h^2-H),$$

ce qui est bien la valeur trouvée par Gauss.

*Gauss: *Méthode des moindres carrés*, traduction française par J. Bertrand; Mallet-Bachelier 1855.

†Loc. cit., pages 62 à 69. Les calculs de Gauss sont très pénibles.

Si les erreurs obéissent à la loi de Gauss*, on a $H=3h^2$; il ne reste plus que le premier terme et l'on obtient la même erreur probable que si l'on avait utilisé la règle bien connue de la moyenne quadratique, en répétant p fois la mesure d'un même b_λ.

9. Introduction d'une observation supplémentaire†. Supposons qu'une nouvelle observation nous donne la valeur numérique b pour la nouvelle forme linéaire $u=a^i x_i$. Il s'agit d'évaluer les corrections à faire subir aux $\bar{x}_i$ précédemment calculés par la formule (12).

Convenons de désigner par des lettres accentuées toutes les variables relatives au nouveau système d'équations. Les variables contrevariantes sont toujours données par la formule (11), qui devient

$$\bar{x}'^i = \sum_\lambda b_\lambda a^i_\lambda + ba^i = \bar{x}^i + ba^i. \tag{20}$$

Mais, ce sont les valeurs des variables covariantes qui nous sont demandées. La nouvelle forme quadratique fondamentale est

$$\Theta' = \Theta + u^2; \tag{21}$$

on a donc

$$x'^i = x^i + a^i u; \tag{22}$$

puis,

$$x_i = A_{ij} x^j = A_{ij}(x'^j - a^j u). \tag{23}$$

Ceci est une identité. Remplaçons-y les x_i par leurs nouvelles valeurs probables $\bar{x}_i'$; il vient, d'après (20),

$$\bar{x}_i' = A_{ij}(\bar{x}^j + ba^j - a^j \bar{u}') = \bar{x}_i + ba_i - a_i \bar{u}'. \tag{24}$$

Il nous reste à calculer $\bar{u}'$. Nous avons, d'après (22),

$$u = a_i x^i = a_i x'^i - u a_i a^i;$$

d'où

$$u(1+\alpha) = a_i x'^i, \tag{25}$$

en posant

$$\alpha = a_i a^i = \Theta(a).$$

Ceci est encore une identité; si nous y remplaçons les x_i par leurs nouvelles valeurs probables $\bar{x}_i'$, nous obtenons

$$\bar{u}'(1+\alpha) = a_i \bar{x}'^i. \tag{26}$$

D'autre part, nous tirons de (20)

$$a_i \bar{x}'^i = a_i \bar{x}^i + b\alpha.$$

*Dans le cas d'une loi quelconque, Gauss cherche une limite supérieure simple de A et trouve $n+p$. Mais, l'identité (9) donne la limite plus rapprochée n.

†Gauss, op. cit., pages 53 à 56. Nous allons suivre, dans ce paragraphe, l'analyse de Gauss, mais en simplifiant l'exposition.

Or, $a_i\bar{x}^i$ est ce que devient u lorsqu'on y remplace les x_i par leurs anciennes valeurs probables $\bar{x}_i$. Soit $b+r$ le résultat de cette opération, de sorte que r désigne le résidu de la nouvelle équation dû à la substitution des anciennes valeurs probables. Nous avons alors

$$a_i\bar{x}'^i = b(1+a)+r$$

et, en comparant avec (26),

$$\bar{u}' = b + \frac{r}{1+a}. \tag{27}$$

Portant cette valeur dans (24), nous obtenons enfin la *correction demandée*:

$$\Delta x_i = x_i' - \bar{x}_i = -a_i \frac{r}{1+a}. \tag{28}$$

10. Nous avons maintenant, d'après (15),

$$E(x_i') = hA'_{ii}.$$

Or, A'_{ij} peut être obtenu en remplaçant u par sa valeur tirée de (25) dans (23) et prenant le coefficient de x'^j; on a ainsi

$$A'_{ij} = A_{ij} - \frac{A_{ik}a^k a_j}{1+a} = A_{ij} - \frac{a_i a_j}{1+a}.$$

En particulier,

$$A'_{ii} = A_{ii} - \frac{a_i^2}{1+a}.$$

La *correction à faire subir à la valeur probable du carré de l'erreur sur* x_i est donc

$$\Delta E(x_i) = -h\frac{a_i^2}{1+a}. \tag{29}$$

11. Calculons enfin la correction ΔS à faire subir à la somme des carrés des résidus. D'après la formule (19), on a

$$\Delta S = -\sum_\lambda b_\lambda \Delta r_\lambda - br'.$$

Or, d'après (28),

$$\Delta r_\lambda = -a_\lambda^i a_i \frac{r}{1+a};$$

d'où

$$-\sum_\lambda b_\lambda \Delta r_\lambda = \frac{ra_i}{1+a}\sum_\lambda b_\lambda a_\lambda^i = \frac{ra_i\bar{x}^i}{1+a} = \frac{r(b+r)}{1+a}$$

en utilisant la formule (11) et se rappelant la définition de r.

Nous avons ensuite, d'après (27),

$$r' = \frac{r}{1+a}.$$

Donc,

$$\Delta S = \frac{r^2}{1+\alpha}. \tag{30}$$

La *correction à faire subir à h* s'en déduit immédiatement:

$$\Delta h = \frac{hp + \Delta S}{p+1} - h = \frac{\Delta S - h}{p+1}. \tag{31}$$

12. CHANGEMENT DE POIDS D'UNE OBSERVATION*. Supposons que le poids de la dernière observation, qui était primitivement 1, devienne $1+\epsilon$. Il revient au même de garder l'ancien poids et d'ajouter une nouvelle observation identique à la dernière, en lui donnant pour poids ϵ. Si, pour simplifier l'écriture, nous supprimons l'indice λ de la m^e observation, nous sommes ramenés exactement au paragraphe 9, avec cette seule différence que b et les a^i doivent être multipliés par $\sqrt{\epsilon}$. Si r désigne l'ancien résidu $\bar{u}-b$, le résidu de l'observation supplémentaire devient $\sqrt{\epsilon}\, r$ et c'est lui qui doit être substitué à r dans les formules (28), (29) et (30), qui deviennent ainsi

$$\Delta x_i = -\epsilon a_i \frac{r}{1+\alpha\epsilon},$$

$$\Delta E(x_i) = -h\epsilon \frac{a_i^2}{1+\alpha\epsilon},$$

$$\Delta S = \frac{\epsilon r^2}{1+\alpha\epsilon},$$

et pourraient d'ailleurs être établies directement, en répétant les calculs des paragraphes 9 à 11, à partir de la forme quadratique fondamentale $\Theta' = \Theta + \epsilon u^2$.

13. THÉORIE RELATIVE AU CAS OÙ L'ON CONNAIT LES ÉQUATIONS DE LIAISON ENTRE LES GRANDEURS OBSERVÉES†. Ces équations sont celles qu'on obtiendrait en éliminant les x_i entre les équations (1). Elles sont au nombre de p. Ecrivons-les sous la forme

$$\sum_\lambda c^i_\lambda u_\lambda = 0, \quad i = 1, 2, \ldots, p; \tag{32}$$

les c^i_λ étant des coefficients connus. Soit $v_\lambda = u_\lambda - b_\lambda$ l'erreur véritable commise dans la mesure de u_λ. En remplaçant u_λ par $b_\lambda + v_\lambda$ dans (32), on obtient

$$\sum_\lambda c^i_\lambda v_\lambda = t^i, \tag{33}$$

*Cf. Gauss, op. cit., page 57. Gauss indique seulement les résultats, avec le principe de la méthode, qu'il dit être analogue à celle du problème précédent. Mais, il ne remarque pas que le deuxième problème se ramène au premier.

†Cf. Gauss, op. cit., pages 70 à 98. La méthode que nous allons exposer est entièrement différente de celle de l'illustre géomètre.

où les t^i ont des valeurs numériques connues, en fonction des b_λ. Cherchons la valeur probable de v_λ. On a*

$$\bar{v}_\lambda = K\int v_\lambda e^{-k\theta} dv_1 dv_2 \ldots dv_m, \tag{34}$$

en posant $\theta = \Sigma\, v_\lambda{}^2$ et en prenant pour champ d'intégration la région d'hyperespace comprise entre les plans (33) et les plans infiniment voisins correspondant aux valeurs $t^i + dt^i$ des seconds membres.

Considérons d'abord le *cas particulier* où ces plans sont deux à deux perpendiculaires, les c_λ^i étant les cosinus directeurs de leurs normales. Faisons un changement de coordonnées rectangulaires, en prenant lesdits plans comme premiers plans de coordonnées. La formule (34) conserve la même forme avec les nouvelles coordonnées v_λ', en vertu de l'invariance de θ et de $dv_1\, dv_2 \ldots dv_m$. Mais, dans le nouveau champ d'intégration, $v_1', v_2', \ldots, v_p'$ ont les valeurs constantes respectives $t^1, t^2, \ldots, t^p$ et les autres v_λ' peuvent varier de $-\infty$ à $+\infty$, il s'ensuit évidemment que $\bar{v}_i' = t^i$ et que $\bar{v}_\lambda' = 0$, si $\lambda > i$. En revenant aux anciennes coordonnées, on a

$$\bar{v}_\lambda = \sum_i c_\lambda^i\, t^i. \tag{35}$$

14. Arrivons maintenant au *cas général*. Par une substitution linéaire convenable effectuée sur les c_λ^i et sur les t^i, on peut se ramener au cas particulier précédent†. Désignons par des lettres accentuées les nouvelles valeurs de ces variables. La formule (35) leur est applicable et nous donne

$$\bar{v}_\lambda = \sum_i c_\lambda{}'^i\, t'^i. \tag{36}$$

Il s'agit maintenant de revenir aux anciennes variables. A cet effet, nous allons mettre (36) sous une forme invariante. Considérons la forme quadratique fondamentale

$$\phi(y') = \sum_\lambda \Big(\sum_i c_\lambda{}'^i y'^i\Big)^2 = \sum_i (y'^i)^2,$$

d'après les relations qui existent entre les cosinus $c_\lambda{}'^i$. Les variables adjointes covariantes y_i' se confondent actuellement avec les y'^i; de sorte que (36) peut s'écrire sous la forme invariante

$$\bar{v}_\lambda = c_\lambda'^i t_i' = c_{\lambda i}' t'^i.$$

Revenons maintenant aux anciennes variables; nous obtenons

$$\phi(y) = \sum_\lambda (c_\lambda^i\, y_i)^2, \tag{37}$$

$$\bar{v}_\lambda = c_\lambda^i\, t_i = c_{\lambda i}\, t^i = \phi(c_\lambda,\, t), \tag{38}$$

*Cf. J. Haag, *Sur la méthode des moindres carrés*: Comptes Rendus Acad. Sciences, Paris, tome 178, 14 avril 1924, page 1356.

†Cela équivaut à dire que, dans le réseau linéaire défini par les plans (33), on peut trouver p plans deux à deux rectangulaires; cela est évidemment possible d'une infinité de manières, si, comme nous l'avons admis implicitement, les p équations (32) sont indépendantes.

On pourrait employer un artifice analogue dans la première méthode. Mais, cela n'apporterait aucune simplification.

les variables covariantes se déduisant des variables contrevariantes comme variables adjointes à la forme quadratique fondamentale $\phi(y)$. Les formules (38) nous donnent les *corrections les plus plausibles** à faire subir aux b_λ.

15. Cherchons maintenant la *valeur probable du carré de l'erreur commise sur* u_λ *après cette correction.* La nouvelle erreur est

$$v_\lambda - \bar{v}_\lambda = v_\lambda - c_{\lambda i}\sum_\alpha c^i_\alpha v_\alpha = v_\lambda - \sum_\alpha v_\alpha \phi(c_\lambda, c_\alpha).$$

La valeur probable de son carré est

$$E(u_\lambda) = h\left\{[1-\phi(c_\lambda)]^2 + \sum_{\alpha \neq \lambda} \phi^2(c_\lambda, c_\alpha)\right\}.$$

Or, la forme quadratique ϕ est formée avec les c^i_λ comme θ était formé, au paragraphe 2, avec les a^i_λ. Elle donne lieu aux mêmes identités, sous la seule condition de remplacer n par p. En utilisant (8), la formule ci-dessus devient

(39) $$E(u_\lambda) = h[1-\phi(c_\lambda)].$$

16. La valeur *a posteriori* de h peut être obtenue au moyen de *la somme* S *des carrés des corrections*†. On a, comme au paragraphe précédent,

(40) $$\bar{S} = h\sum_{\lambda, \alpha} \phi^2(c_\lambda, c_\alpha) = hp,$$

d'après (9). On retrouve donc la formule (17).

Pour le calcul direct de S, on peut remarquer que, d'après (37) et (38), on a

(41) $$S = \sum_\lambda (c_{\lambda i} t^i)^2 = \phi(t) = t_i t^i.$$

17. La formule (41) peut s'écrire

$$S = t_i\Big(\sum_\lambda v_\lambda c^i_\lambda\Big) = \sum_\lambda v_\lambda \bar{v}_\lambda,$$

si les v_λ constituent un système quelconque d'erreurs satisfaisant à (33). On déduit de là

$$\sum_\lambda v^2_\lambda = S + \sum_\lambda (v_\lambda - \bar{v}_\lambda)^2 \geq S.$$

Donc, les corrections (38) sont, parmi toutes les corrections compatibles avec (33), celles qui rendent minimum $\sum v^2_\lambda$. On retrouve le *principe des moindres carrés*. D'ailleurs, en appliquant la méthode des multiplicateurs de Lagrange, il est aisé de retrouver directement les formules (38). Lesdits multiplicateurs ne sont autres que nos t_i.

18. Comme au paragraphe 8 et plus simplement encore, on peut *calculer* V On a, d'après (41) et (33),

$$S = \sum_\alpha v^2_\alpha\ \phi(c_\alpha) + 2\sum_{\alpha, \beta} v_\alpha v_\beta \phi(c_\alpha, c_\beta).$$

*Nos v_λ sont exactement les ϵ de Gauss. On peut les obtenir aussi très facilement sous la forme (38), en suivant l'analyse du grand géomètre avec les notations tensorielles.

†Ces corrections ne sont autres que les résidus de la première théorie.

Élevons au carré et prenons les valeurs probables. Il vient

$$\overline{S^2} = HA + 2h^2(p - A) + h^2(p^2 - A),$$

en posant

$$A = \sum_{\lambda} \phi^2(c_\lambda)$$

et utilisant les identités (9) et (7). D'où*

$$V = \frac{HA + h^2(2p - 3A)}{p^2} = \frac{2h^2}{p} + \frac{H - 3h^2}{p^2} A. \tag{42}$$

*Gauss donne cette formule sans démonstration. Si les erreurs suivent la loi de Gauss, le second terme disparaît. Dans les autres cas, il est limité supérieurement par $\frac{H-3h^2}{p}$, puisque $A \leq p$.

ABSTRACTS OF COMMUNICATIONS

SECTION III

ASYMPTOTIC SOLUTIONS IN THE PROBLEM OF THREE BODIES

BY PROFESSOR DANIEL BUCHANAN,
University of British Columbia, Vancouver, Canada.

The paper deals with the orbits which are asymptotic to the straight line and equilateral triangle points of libration and also with those which are asymptotic to the periodic orbits in the vicinity of these equilibrium points. Cases are considered when the third body is infinitesimal and the finite bodies move (1) in circles and (2) in ellipses; also when the three bodies are finite.

THE ORBIT OF THE EIGHTH SATELLITE OF JUPITER

By Professor E. W. Brown,
Yale University, New Haven, Conn,, U.S.A.

This paper is a general account of the main features of the orbit of Jupiter's eighth satellite. With an eccentricity of .4, an inclination of over 30° and a period about one-sixth that of Jupiter round the Sun, the problem of its motion under the attractions of Jupiter and the Sun presents difficulties owing to slow convergence which are different from and much greater than those of the ordinary lunar and planetary theories. The retrograde motion causes no trouble. A literal theory seems to be out of the question and this attempt at a numerical theory has involved somewhat extensive calculations. The equations of motion with the true longitude in the moving orbital plane, as adopted in the author's theory of the Trojan group, are used, being simplified by the adoption of a "departure point" as an origin of reckoning. An intermediate orbit, which includes all powers of the ratio of the mean motions of Jupiter and the satellite and the eccentricity of the latter with the principal effects of the inclination, has been obtained, so as ultimately to give the longitude relative to Jupiter within about a minute of arc. The positions depending on the eccentricity of Jupiter and on the ratio of the mean distances are in general much smaller and it is shown how they may be calculated. No evidence of instability of the orbit has so far appeared.

THE THEORY OF A THIN ELASTIC RECTANGULAR PLATE CLAMPED AT THE EDGES

By Professor A. C. Dixon,
Queen's University, Belfast, Ireland.

The discussion of this subject introduces functions of a complex variable with singularities of a special kind occurring at the points of the lattice to which the rectangle belongs. Green's function, that is, the deflection produced at any point (x,y) of the plate by unit load at any other point (x_0, y_0) is a function of x, y with the singularity of

$$\tfrac{1}{2}(z-z_0)(\bar{z}-\bar{z}_0)\{\log(z-z_0)+\log(\bar{z}-\bar{z}_0)\}$$

and may be expressed in four ways, namely,

$$\phi(z)-\psi(z)-\phi(\bar{z})+\psi(\bar{z})+(\bar{z}-z)\{\phi'(z)-\psi'(\bar{z})\},$$
$$\phi(z)-\psi(z-2ib)-\phi(\bar{z}+2ib)+\psi(\bar{z})+(\bar{z}+2ib-z)\{\phi'(z)-\psi'(\bar{z})\},$$
$$-\phi(z)+\psi(-z)+\phi(-\bar{z})-\psi(\bar{z})+(\bar{z}+z)\{\phi'(z)+\psi'(\bar{z})\},$$
$$-\phi(z)+\psi(2a-z)+\phi(2a-\bar{z})-\psi(\bar{z})+(\bar{z}+z-2a)\{\phi'(z)+\psi'(\bar{z})\},$$

where $z=x+iy$, $\bar{z}=x-iy$, and the edges of the plate are the lines $z=\bar{z}$, $z=\bar{z}+2ib$, $z=-\bar{z}$, $z=-\bar{z}+2a$. By the comparison of these forms it is found that if

$$\chi(z)=\frac{\partial^3}{\partial z\partial z_0\partial\bar{z}_0}\{\phi(z)-\alpha\psi(a-z)-\beta\psi(z-ib)+\alpha\beta\phi(a+ib-z)\}$$

where $\alpha=\pm1$, $\beta=\pm1$, then $\chi(z)$ has in the strip $0<x<a$ only the singularities of

$$f(z)\equiv-\frac{1}{2ib}\left\{\varpi_\beta\frac{z-\bar{z}_0-ib}{ib}-\alpha\varpi_\beta\frac{z+z_0-a-ib}{ib}-\beta\varpi_\beta\frac{z-z_0}{ib}+\alpha\beta\varpi_\beta\frac{z+\bar{z}_0-a}{ib}\right\}$$

where $\varpi_\beta(u)$ is the function denoted by

$$-\int_{-\infty}^{\infty}\frac{e^{\lambda u}d\lambda}{e^{\lambda}-e^{-\lambda}-2\beta\lambda}$$

when $-1<Ru<1$, and the function* χ satisfies the conditions:

$$\chi(a+ib-z)=-\alpha\beta\chi(z);$$
$$\chi(z)-\chi(ze^{i\pi})=2\alpha z\chi'(a-z)=2\beta z\chi'(z+ib);$$
$$-\alpha\chi(a-z)+\alpha\chi(a-ze^{-i\pi})=\beta\chi(z+ib)-\beta\chi(ib+ze^{-i\pi})=2z\chi'(z).$$

*For the theory of this function see Proc. London Math. Soc., (2), 21, (1922) pp. 277-8.

Taking

$$\int \left\{ \varpi_\beta \frac{t-z-ib}{ib} \chi(t) + \varpi_\beta \frac{t-z}{ib} \chi(t-ib) \right\} dt$$

round the rectangle, we find

$$\chi(z) = f(z) + \int_0^{ib} \chi(t) J(t, z) dt \tag{1}$$

where

$$-2\pi b J(t, z) = -\varpi_\beta \frac{t-z-ib}{ib} + \varpi_\beta \frac{t+z-ib}{ib} + \alpha\beta\varpi_\beta \frac{t+z-a}{ib}$$

$$-\alpha\beta\varpi_\beta \frac{t-z+a}{ib} + 2\beta\varpi_\beta \frac{t-z}{ib} - 2\alpha\varpi_\beta \frac{t+z-a-ib}{ib} - 2\beta \frac{ib-t}{ib} \varpi'_\beta \frac{t-z}{ib}$$

$$-2\alpha \frac{t}{ib} \varpi'_\beta \frac{t+z-a-ib}{ib}.$$

There is a similar equation

$$\chi(z) = g(z) + \int_0^a \chi(t) K(t, z) dt. \tag{2}$$

In (2) t is real and in the interval $(0, a)$, and in (1) z may lie in this same interval: on the other hand in (1) t lies on the straight path between 0 and ib and in (2) z may also lie on this segment. Thus (1) and (2) are integral equations to be satisfied by the function χ on two of the sides of the plate. When they are solved, they also determine χ as a function of the complex variable z, and thus yield the solution of the mathematical problem to find G. The integral equations do not fall under the ordinary theory since the nuclei J, K are not bounded. The problem of their solution is discussed*.

*The theory discussed in the present paper has been considered more fully by the author in a paper presented to the London Mathematical Society, April 23, 1925, under the title *The functions involved in the theory of a thin elastic rectangular plate, clamped at the edges, and certain integral equations satisfied by such functions.* This paper has since appeared in the Proceedings of the London Mathematical Society, Second Series, vol. 25, p. 417, 1926.

A FINITE WORLD-RADIUS AND SOME OF ITS COSMOLOGICAL IMPLICATIONS

By Dr. Ludwik Silberstein,
Research Laboratory, Eastman Kodak Company, Rochester, New York, U.S.A.

Assuming deSitter's world, a rigorous formula for the Doppler effect, for star and observer in any inertial motion, is given by the equation

$$\frac{\lambda}{\lambda+\delta\lambda} = \gamma \sec^2\sigma \left[1 \mp \sqrt{1 - \frac{\cos^2\sigma}{\gamma^2}\left(1 + \frac{p^2}{R^2 \sin^2\sigma}\right)}\right]$$

($\sigma = r/R$, p, γ integration constants), and with the aid of its approximate form,

$$\left(\frac{\delta\lambda}{\lambda}\right)^2 = \left(1 - \frac{r_m^2}{r^2}\right)\left(\frac{r^2}{R^2} + \frac{v_m^2}{c^2}\right),$$

where r_m, v_m are constants, the data of eighteen clusters, the Magellanic clouds and one spiral nebula are shown to bear evidence for a finite curvature radius R, with $7 \cdot 7.10^{12}$ astr. units as most probable value.

The equations of motion in a radial gravitation field are discussed, with special attention to a single star and to a uniform globular galaxy of stars. In the former case the star, of mass $M = c^2L$, is shown to be surrounded by a neutral sphere of radius $(LR^2)^{\frac{1}{3}}$ and in the latter case a necessary condition of permanence (recurrence) is deduced, according to which, if Lc^2 be the galactic mass, $2(LR^2)^{\frac{1}{3}}$ is the upper limit to the diameter of the globular system. This criterion is illustrated on globular clusters, properly so-called, and our own galaxy. The former objects satisfy amply the said criterion, whereas our galaxy is much too large to be permanent.

The deSitterian inertial behaviour is imitated approximately by a fourd imensional Euclidean rotation around the observer as centre, with the "cosmic day" $\frac{2\pi R}{c}$ as period.

A NEW DEDUCTION OF THE ELECTROMAGNETIC EQUATIONS

By Professor W. F. G. Swann,
Director, Sloane Laboratory, Yale University, New Haven, Connecticut, U.S.A.

Suppose E is a vector, the volume integral of whose divergence

$$\rho = \text{div } E \tag{1}$$

is supposed to be conserved in the sense of the equation of continuity:

$$\text{div } \rho u + \frac{\partial \rho}{\partial t} = 0.$$

It immediately follows that

$$\text{div}\left(\rho u + \frac{\partial E}{\partial t}\right) = 0,$$

so that there exists a vector h, such that

$$\rho u + \frac{\partial E}{\partial t} = \mu \text{ curl} h, \tag{2}$$

where μ is a constant.

Write

$$\text{div } h = -\sigma_0 \tag{3}$$

and suppose that the volume integral of div h is conserved in the sense given by

$$\text{div } \sigma_0 \nu + \frac{\partial \sigma_0}{\partial t} = 0,$$

so that

$$\text{div}\left(\sigma_0 \nu - \frac{\partial h}{\partial t}\right) = 0,$$

and a vector ϵ exists such that

$$\sigma_0 \nu - \frac{\partial h}{\partial t} = \mu \text{ curl} \epsilon, \tag{4}$$

where ν is a quantity transforming like a velocity under restricted relativity.

It then appears that, in order that the equations 1-4 shall transform under the restricted relativity transformation in such a manner as to be mutually consistent, we must have $\epsilon = \frac{c^2}{\mu^2} E$.

Finally, defining $H = \frac{\mu}{c} h$ and $\sigma = \frac{\mu}{c}\sigma_0$, we arrive at the ordinary electro-magnetic equations:

$$\frac{1}{c}\left(\rho u + \frac{\partial E}{\partial t}\right) = \text{curl } H,$$

$$\rho = \text{div } E,$$

$$\frac{1}{c}\left(\sigma v - \frac{\partial H}{\partial t}\right) = \text{curl } E,$$

$$\sigma = -\text{div } H,$$

generalized, however, by the addition of the σ terms corresponding to real magnetic charge, the density of which may, of course, be zero.

NEW PHYSICAL THEORIES

By Professor F. M. da Costa Lobo,
Director of the Royal Astronomical Observatory, Coimbra, Portugal.

In this paper an attempt is made to provide an explanation of physical phenomena, which shall agree with all observations and be based upon the most general principles.

Conceptions introduced in elucidation of these principles may be defined as follows:

Material point. The minimum of matter considered independently of any property we may attribute to it.

Space. The actual and possible positions of material points.

Physical universe. The totality of material points.

Distance between two points. The minimum of points necessary to establish connection between them without interruption or repetition.

Orientation. The conjuncture of two distances.

Phenomenon. An aggregate of points together with their distances and orientations.

Time. Abstract notion associated with the continuous succession of phenomena defined by certain points; it is a function of the relation between the space determined by the assembling of the successive positions of a set of points and the space corresponding to a certain conjuncture in the successive positions of the said set.

Instant. A cut in the succession corresponding to a definite phenomenon.

Velocity. The succession of phenomena comprised between definite instants.

Furthermore with reference to the material points, already regarded as the limit of matter's disintegration, the author propounds the following principle:

The universe is constituted by an aggregate of material points conceived as possessing and maintaining indefinitely the minimum of matter and the maximum of velocity.

NOTE ON THE FORCE EQUATION OF ELECTRODYNAMICS

By Mr. S. J. Jacobsohn,
University of Chicago, Chicago, Illinois, U.S.A.

The equation which is usually taken as determining the motion of the electron

$$\iiint \rho \left(E + \frac{v \times H}{c} \right) d\tau = 0,$$

where (E, H) is the resultant field of the electron and all other charges, the integration extending over the electron, is shown to be non-covariant under the transformation of the special theory of relativity. It is covariant, when assumed in the case of a Lorentz electron, to hold for the inertial system which has the instantaneous velocity of the centre of the electron. In this form, however, it leads to expressions for the momentum and energy equations which are different from those ordinarily employed, the difference necessitating considerable care in the meanings to be attached to electromagnetic force and rate of doing work.

ON THE ELECTROSTATIC POTENTIAL ENERGY OF THE CALCITE CRYSTAL

By Professor S. Chapman,
Imperial College of Science and Technology, London, England.

The calcite crystal has the form of a cube shortened along a diagonal (its trigonal axis). The CO_3 ions are in planes perpendicular to this axis. The crystal arrangement is such that distortion is possible by a contraction or elongation of the dimensions along the trigonal axis (accompanied by elongation or contraction at right angles) without destroying the type of the grating structure. It is shown that, on certain simple assumptions in accordance with modern views on crystals of this type, the electrostatic potential energy has a minimum value corresponding to a rhombohedral angle of the crystal nearly equal to, though actually 4° greater than, the observed rhombohedral angle of the crystal. The difference of 4° can be accounted for in a general way as due to the over-simple nature of the assumptions underlying the calculation. The *variation* of the angle, throughout a group of carbonate crystals of similar structure, can be accounted for with great accuracy by allowing for the different sizes of the metallic ions.

SOLVED AND UNSOLVED PROBLEMS IN DYNAMICAL METEOROLOGY

By Professor V. Bjerknes,
Director, Geophysical Institute, Bergen, Norway.

The motion of the atmosphere should be one of the most obvious applications of theoretical hydrodynamics. But up to date it has not been so. The main reason is that attention has been directed almost exclusively to fluids of which the properties have been too strongly idealized. In this connection the omission of friction is unimportant. The main thing is that the equation

$$\rho = f(p)$$

has generally been adhered to (ρ density, p pressure). Thereby the connection with thermodynamics is cut off; the consideration of the origin of atmospheric motions is completely excluded; the general principle of the absolute conservation of vortex motion is obtained, which leads to such a striking contradiction to everyday meteorological experience.

The deciding step will therefore always be to start with the true equation

$$\rho = f(p, \theta, h),$$

(θ temperature, h humidity). After this hydrodynamics can exert its influence upon the development of meteorology along the following two lines:

I. By furnishing meteorologists with hydrodynamic laws in a form in which they can be applied directly to the actual meteorological situations or processes.

II. By solving special problems which contribute to the general understanding of the dynamics of the atmosphere.

In line I. the barometric formula, the principle of the "geostrophic wind," and the laws of the vortex formation (which contain as special cases the Helmholtz-laws of their conservation) have been important hydrodynamic contributions to meteorological work. The ultimate aim of this kind of work will be to arrive—at least in theory—at a mathematical method of weather forecasting, based upon a direct application of the hydrodynamic and thermodynamic laws to the situations represented by the weather chart. Even if this aim should never be reached, the attempt within the limits of possibility to treat the problem of the weather forecasting as if it were a mathematical problem can only react favourably upon the progress of this work. In reality it has already done so to a considerable extent (cf. the progress connected with the discovery of the polar front phenomena and their utilization in the forecasting work).

For the mathematician the progress along line II. will be the more interesting, as it will furnish him with important problems. The main thing is to formulate the problems in such a way that from the meteorological side we reach our goal and from the mathematical side we do not encounter difficulties impossible to overcome. This may be realized in the following way:

In the first approximation we consider the great atmospheric motions as steady motions going on very slowly. This is legitimate because the barometric formula, although a formula of hydrostatics and not of hydrodynamics, is fulfilled with a very high degree of approximation during all great atmospheric motions.

In the second approximation we consider the consequence of small disturbances of this steady state of motion.

In this way we arrive at problems defined by linear equations with which we can hope to proceed mathematically; and what is equally important, it can be foreseen by elementary qualitative methods that in this way we shall arrive at a rational treatment of the most central problem in dynamical meteorology, that of the extratropical moving cyclones and anticyclones.

The paper was illustrated by diagrams and models.

NOTE: For more detailed information both with regard to subjects treated in the paper whose abstract is herewith printed and with regard to the contents of the other paper presented to the Congress under the title "The Forces that Lift Aeroplanes" the reader may consult the following two papers by Professor Bjerknes:

Le problème des Cyclones, Jour. de Physique, Paris, Nov., 1924.

Sur les forces qui portent les aéroplanes et leur relation avec les actions hydrodynamiques à distance, ibid., Paris, Dec., 1924.

The complete contents of the author's two Congress papers together with further developments will also be found included in a course of lectures delivered under the title "Physical Hydrodynamics" at the California Institute of Technology in Pasadena during the Fall of 1924.

THE MATHEMATICAL ANALYSIS OF THE EARTH'S MAGNETIC FIELD

By Dr. Louis A. Bauer,
Director, Department of Terrestrial Magnetism, Carnegie Institution, Washington, D.C., U.S.A.

The Department of Terrestrial Magnetism of the Carnegie Institution at Washington has under way a mathematical analysis of the Earth's magnetic and electric fields, based on homogeneous and reliable data obtained chiefly by the Department since 1905, and supplemented effectively by various co-operating countries.

This analysis is being made with strict regard to the refinements which enter into the spherical harmonic series by taking into account the actual shape of the Earth and the latest developments resulting from the theory of relativity.

The preliminary results indicate the need of making the analysis, at the start, as free as possible from any hypothesis as to the physical cause or causes of any constituent portions of the magnetic or electric field to be studied. In brief the numerical coefficients first obtained are strictly the product of the mathematical analysis as applied to the observed data. Anyone is thus enabled to start with the so-called "fundamental magnetic or electric constants" of the Earth and make any further attempt he may desire at a physical analysis or physical interpretation in the light of the most modern developments of theory and experiment.

Additional observations in the polar regions will be needed for the final solution of some of the outstanding questions of theory.

(Illustrated by lantern slides).

CALCULATION OF THE VELOCITY OF VERTICAL OCEAN CURRENTS IN THE SAN DIEGO REGION FROM THE ACCOMPANYING TEMPERATURE REDUCTION BELOW "NORMAL" VALUES

By Dr. G. F. McEwen,

Scripps Institution for Biological Research, La Jolla, California, U.S.A.

Considerable attention has been given to the computation of ocean currents regarded as effects of various measured forces, winds, differences in specific gravity, earth's deflecting force, etc. In contrast to the use of a dynamical basis, currents may be regarded as causes of changes in "normal" gradients of temperature, salinity, or other properties of the water. Qualitative investigations of circulation based upon the distortion of isotherms, isohalines, etc., have been extensively made, since the beginning of the science of oceanography.

A mathematical theory of temperature gradients in bodies of water having only a convective circulation due to evaporation and the absorption of solar radiation has been developed by the author and applied to lake data, with very satisfactory results. The present paper presents an application of the same equations, so amended as to include the effect of an average vertical flow of water, to observations of serial temperatures and salinities in the Pacific off Southern California. Preliminary numerical estimates are presented of the rate of "upwelling" in this region, and of the temperature reduction due to any rate of flow parallel to a temperature gradient of given value.

ON THE DIRECT NUMERICAL CALCULATION OF ELLIPTIC FUNCTIONS AND INTEGRALS

By Professor Louis V. King,
McGill University, Montreal, Canada.

Although the numerical evaluation of elliptic functions and integrals in terms of the Theta-functions is extremely rapid by the use of the highly convergent series involving the *nome* q, these are not usually the functions which make their appearance directly in physical or astronomical problems. The necessary transformations required to express the results in terms of the q-series require in many cases somewhat complicated analysis, while numerical computation by this method necessitates the exact calculation of a large number of auxiliary quantities.

The monograph briefly described under the above title includes the entire theory of elliptic functions associated with the arithmetico-geometrical scales, thus completing and adding to the original developments of Lagrange, Legendre, Gauss, Jacobi and others, extending them in several directions and placing before mathematical physicists, in accessible form, a mode of approach to elliptic function theory directly related to the art of machine computation. While many new formulae and methods of computation are derived, no claim is made as to novelty in fundamental analytical treatment. It is interesting to note that this treatment was to have formed the content of the last volume of Halphen's *Fonctions elliptiques,* unfortunately left incomplete at the time of his death.

NOTE ON THE REDUCTION OF PARALLAX PLATES

By Professor H. C. Plummer,
Artillery College, Woolwich, England.

Professor Schlesinger's method of reducing photographic plates for the determination of stellar parallaxes has proved simple and valuable. In this note the method is explained on slightly different lines, and the order of accuracy to be expected from the results is discussed.

COMMUNICATIONS

SECTION IV

ENGINEERING

RESEARCH IN MECHANICS AND SOUND AT THE BUREAU OF STANDARDS

By Dr. Lyman J. Briggs,

Chief of Division of Mechanics and Sound, Bureau of Standards, Washington, D.C., U.S.A.

The purpose of this paper is to describe briefly some investigations which have recently been carried out by the staff of the Division of Mechanics and Sound, Bureau of Standards. The work of this Division relates particularly to the subjects of applied mechanics, sound, ballistics, aerodynamics and aeronautic and engineering instrument development.

Compression Members.

The Bureau's equipment of large testing machines is almost continually in demand to determine experimentally the carrying capacity of structures for whose design no adequate theory has been worked out. This has included the testing of compression bridge members under loads up to 6 million pounds, web plates for the new Camden Bridge over the Delaware, girders for the airship Shenandoah, and other similar structures.

Tests of this nature, planned to determine the adequacy of a given design for a specific purpose, do not readily fit into any systematic programme of study of the theory of column design. We have, however, endeavoured to arrange the work so that all the compression tests shall not only serve their immediate specific purpose, but shall also contribute something of permanent value to the theory of column design.

From the standpoint of applied mathematics, the compression members tested at the Bureau fall into three groups.

The first group includes the sturdy columns, such as are common in steel structures, ranging in slenderness from $L/R=40$ to $L/R=90$. The tests on these columns have confirmed the results of other investigators that the largest factor in determining their strength is the quality of the material and especially its yield point. For example, in tests of two columns identical in construction and as nearly identical in dimensions as mill practice allows, one failed under a stress of 26,000 lb./in.^{2} while the other stood a stress of 41,000 lb./in.^{2} On the other hand, tests made of a group of six solid rolled H columns showed that none of them differed in efficiency $\left(\frac{\text{column strength}}{\text{yield point of material}}\right)$ by as much as 4% from the mean of the group, although their yield points differed by more than 25%

and the type of construction and manner of failure were markedly different. (See Tech. Paper No. 328, U.S. Bureau of Standards, 1926).

Interpretation of results is made difficult by the fact that it appears to be technically impossible to secure homogeneous rolled material in large sections. In some large columns recently tested, yield points were observed differing by as much as 30% over the cross-section. In addition, the strength of a column will differ somewhat with the rate of loading.

The column theory which seems to fit best the results of these tests is the "double modulus" theory, first stated by Considère and elaborated by Karman and Southwell. In particular it accounts for the "pick up" of load by a badly bent column noted by Lilly and Karman on small struts and which was observed on all six of the columns referred to above. It also accounts for the failure in the direction of greatest moment of inertia shown by some of them.

Dr. L. B. Tuckerman is endeavouring to derive from this "double modulus" theory an approximate column formula for eccentric loading, not too complex for practical use and yet valid over a wide range of sizes, shapes, and materials. Mathematically, the difficulty lies in the fact that even with the simplest assumptions the curve of the column axis appears in terms of elliptic integrals whose modulus depends upon the empirical constants of the material, the dimensions of the column, and the applied load. The determination of the maximum load is then a problem in elliptic modular functions, which has so far proved intractable to analytic treatment. Graphical and tabular integrations are, however, being used which seem to promise results.

From this investigation four fairly definite results have come which are in agreement with available experimental data:

1. For columns of all materials showing a well-defined yield point closely similar curves should be obtained connecting the dimensionless variables

$$\lambda = \frac{1}{\pi\sqrt{\frac{E}{S}}} \cdot \frac{L}{R} \text{ and } \sigma = \frac{s}{S},$$

L = length of the equivalent round end column,
R = least radius of gyration,
E = Young's Modulus,
S = yield point of material in compression,
s = column strength.

2. For values of $\lambda < 0.40$ an axially loaded column may be expected to carry loads well beyond the yield point of the material, with no falling off in the neighbourhood of the yield point.

3. For values $0.40 < \lambda < 0.50$ pick-up in varying degrees may be expected.

4. For values of $0.50 < \lambda$ the strength of the column can be closely represented by either of the two empirical formulae

$$\lambda = \frac{1}{\sqrt{\sigma}\sqrt{1+\frac{1}{2}\frac{\alpha^2}{(1-\sigma)^2}}}, \quad \lambda = \frac{2}{\sqrt{\sigma}\left(1+\sqrt{1+\frac{\beta^2}{(1-\sigma)^2}}\right)},$$

α and β being constants readily obtainable from the stress-strain curve of the material but differing somewhat for different shapes of cross-section. These formulae asymptotically approach Euler's formula as λ becomes infinite. It seems probable that with different values of α and β the same formulae will represent columns with eccentric loading fairly well.

An extended series of experiments with small specimens of widely different materials (steels, wrought iron, brass, duralumin, monel metal) has been started to see over what range and to what accuracy formulae of this type can be used. If it proves possible to predict accurately the strength of sturdy columns from coupon tests on the material, we will be more nearly in a position to estimate the practical significance of failures in columns which are apparently not sufficiently sturdy.

At the other extreme of the family of compression members lies the second group, the group of very light latticed columns and girders for use in aircraft. An extended series of such tests made by Mr. R. S. Johnston on type members of the framework of the Shenandoah resulted in each instance in a typical flexural failure, due to establishing within the elastic limit of the material a state of critical instability either in the girder as a whole or in some of its parts. Consequently, the loads such girders will carry are determined solely by their geometrical dimensions and the moduli of elasticity of the material. The ultimate strength or yield point of the material has no influence on the strength of the girders so long as it is greater than the critical stresses.

Theoretically, the critical loads of these girders should be calculable by pure elastic theory, such as Westergaard's* general theory of buckling of elastic structures. Practically, however, the geometry of the structure is too complex, but a semi-empirical approximation has been developed by Dr. Tuckerman. The curve obtained by plotting stress against slenderness ratio for columns of this type is evidently similar to those obtained from sturdy columns (Fig. 1). A good approximation to this curve is given by the formulae for sturdy columns presented above, provided the apparent modulus used is decreased slightly with decreasing slenderness ratio due to shear deflection and the stress-strain curve of the material is replaced by the "stress-strain curve" of a short section of the column. Due to eccentric action of the lattice members this stress-strain curve shows a well-defined, practically horizontal portion or a "yield point" which varies but little even for widely different sizes of columns of similar construction. This "yield point" of the column evidently depends upon: (1) the torsional rigidity, (2) the two flexural rigidities of the channels, and (3) the arrangement of the lattices.

An experimental investigation has been started for the Bureau of Aeronautics, Navy Department, to determine the relation between this "stress-strain curve" with its yield point and the shape of the channels used. There is room for considerable improvement in design since the material has a yield point of over 30,000 lb./in.^{2} while the girders uniformly failed under a stress of about 22,000 lb./in.^{2}

*Proc. Am. Soc. Civil Engineers, vol. 47, pp. 455-533, 1921.

Between these two extremes lies the third broad group of columns, those which are not sufficiently sturdy to exclude secondary or detail failure and not sufficiently light to develop critical instability within the elastic limit. For such structures no real theory exists. The theory which is used consists of semi-empirical interpolation or extrapolation from test results. Perhaps the most promising method of treating them lies in Pettingill's theory of classification presented in Southwell's* interesting paper on the strength of struts.

By a proper system of classification the result of each column test becomes more valuable by being placed in a logically related group.

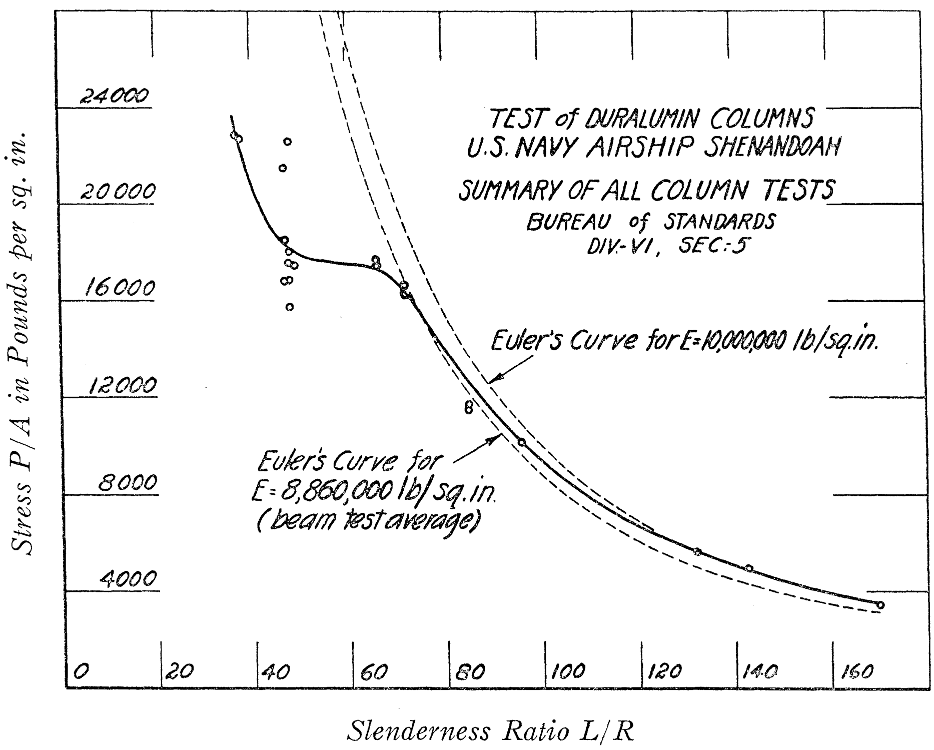


Fig. 1
Stress vs. Slenderness Ratio for Duralumin Latticed Columns.

We have seen in the case of sturdy columns that pure elastic theory may furnish a basis for approximate interpretation even up to the yield point; so also in secondary failure elastic theory may point the way to correct design. In tests conducted by Mr. Johnston of the web plates for the towers of the Delaware River Bridge at Camden, buckling occurred after the limit of elasticity had been passed, but he found the theory of Bryan and Southwell on the elastic buckling of plates useful in interpreting the results.

*Aircraft Engineering, vol. I, pp. 135-138, 1920.

Ballistic Investigations

Solenoid Coil Chronograph.

A chronograph for the accurate measurement of the velocity of projectiles has been developed by Eckhardt and his associates in co-operation with the technical staff of the Ordnance Department, United States Army. This apparatus is based upon the use of solenoid coil screens, an important development in the field of exterior ballistics due to Mr. F. E. Smith of the British Admiralty Office. Two or more of these coil screens are connected in series to an oscillograph element. When a magnetized projectile passes through a solenoid coil the element is deflected and its movement is recorded photographically upon a film carried on the periphery of a high speed drum. During the exposure this drum is not only rotated rapidly but is also translated slowly along its axis by means of a screw cut on the axis so that the oscillograph element traces a spiral record on the film many feet in length. Time intervals are recorded on the film by means of a second oscillograph element which is connected with the plate circuit of a 500 cycle tuning fork equipped with an electron tube drive. This second element has a large amplitude and sweeps across the trace made by the first element at high speed, giving a series of well-defined lines representing time intervals of one-thousandth of a second. With this apparatus a clear-cut record may readily be obtained in which a time interval of one-thousandth of a second is represented by a length of one centimeter on the photographic film. The time required for the projectile to pass from one coil to the next can thus be accurately determined. This time, supplemented by an exact determination of the distance between the coils (about 5 meters), determines the velocity. Oscillograph cameras of this type have been supplied to the proving grounds of both the Army and the Navy.

The distance between the solenoid coils may be definitely fixed by mounting the coils so that they form the ends of a rigid cage, and when mounted in this way the coils may be quickly elevated and oriented to conform to high-angle fire. This equipment has recently been used in extensive range tests of a 154 mm. howitzer at the Aberdeen Proving Ground under the direction of Colonel Tschappat, in which 200 rounds were fired at different elevations in 8 hours, and the muzzle velocity and range of each round was determined.

Piezo-Electric Gauge.

A piezo-electric pressure gauge, owing to the absence of inertia effects, is particularly well-adapted to the measurement of rapid pressure changes such as those in the breech of a gun. Unlike spring gauges, the displacements involved in the piezo gauge are due entirely to the elastic yielding of the materials. Sir J. J. Thomson* and D. A. Keys† have employed the piezo-electric property of tourmaline for this purpose, the record of the variation of pressure with time being obtained by means of a cathode ray oscillograph.

*J. J. Thomson: Engineering, 107, 1919, pp. 543-544.

†D. A. Keys, *A piezo-electric method of measuring explosion pressures:* Phil. Mag., October, 1921, pp. 473-488.

At the request of the Ordnance Department, the development of a piezo-electric gauge of quartz was undertaken. A thin plate cut from a quartz crystal so that the plane of the plate is perpendicular to an electric axis of the crystal and subjected to a compressional force in the direction of the electric axis will liberate a quantity of electricity at its faces proportional to the force applied. If a stack of these plates is built up with due regard to the polarity of the faces, using thin metal electrodes between adjacent plates, with alternate electrodes connected together in condenser fashion, the charge developed with a given load will be proportional to the number of plates used. However, quartz, unlike tourmaline, yields no charge when exposed to a uniform hydrostatic pressure, so that in measuring pressure the stacked plates must be inserted in a gas-tight housing and the pressure must be applied in the direction of the axis of the stack by means of a piston plunger.

The cathode-ray oscillograph is ideal in being free from inertia effects, but its use as a recording instrument for ballistic purposes is seriously hampered by the necessity of pumping out the instrument to a cathode-ray vacuum each time a plate is exposed. The recording apparatus which has been developed by Dr. Karcher* for use with the piezo-electric gauge in ballistic work consists essentially of a recording ballistic galvanometer having a long period and small damping. The throw of the galvanometer when the gauge is discharged through it is recorded by a camera similar to the oscillograph camera described above. Karcher has shown that when the damping constant and the restoring constant of the galvanometer are sufficiently small, the angular velocity of the galvanometer mirror is proportional to the pressure. The differential of the deflection-time curve obtained, multiplied by a suitable constant, gives the pressure-time curve. The constant is determined by loading the gauge in a testing machine and recording the throw of the galvanometer when the load is suddenly released. The load-throw curve obtained in this way is a straight line passing through the origin.

Direct Measurement of the Air Resistance of Projectiles.

The way in which the air resistance of a projectile changes with speed, particularly for speeds approximating that of sound is of great interest from the standpoint of aerodynamics as well as ballistics The experimental determination of the air resistance of projectiles has heretofore been based almost exclusively upon the measurement of the time required for the projectile to pass in succession through a series of screens, from which the change in velocity, and consequently the resistance, can be found.

Direct measurement of the resistance of projectiles in an air stream have recently been made for the Ordnance Department by Lieut. Col. G. F. Hull and the writer. Models of projectiles up to 4 inches in diameter have been used and resistance measurements have been carried out at varying speeds up to the speed of sound. The air stream was obtained by allowing a continuous supply of compressed air from a large centrifugal compressor to expand through a cylindrical nozzle 12 inches in diameter located at the top of a vertical stand-

*J. C. Karcher: Bureau of Standards Scientific Paper No. 445, August 4, 1922.

pipe. The measurements were made in the centre of the free jet, the nose of the projectiles being from 6 to 12 inches above the upper end of the nozzle. The speed and temperature of the free jet were computed from the speed, temperature and pressure of the air in the stand-pipe, assuming an isentropic expansion. The measurements have also included drag and cross-wind forces on projectiles and head-resistance measurements of bomb models, stream-line bodies and spheres.

For head-resistance measurements a piston type of balance has been found to be the most satisfactory. The cylinder of the balance is supported from a stream-lined bar, coaxial with the vertical air-stream. A spindle projecting downward from the piston is attached to the trailing end of the model. To measure the force on the model, the pressure in the cylinder is adjusted until the model floats in the wind stream. The force on the model can then be computed from the weight of the moving system, the diameter of the piston, and the differential pressure. A correction for the resistance of the spindle is made by disconnecting it from the model and holding the latter in the same position relative to the spindle on an independent support. The large centrifugal compressors for supplying the air-stream have been made available through the courtesy of the General Electric Company.

The measurements show that the rate of change of the resistance coefficient as the speed of sound is approached depends upon the form of the projectile, and for some projectiles departs materially from that given by the Gavre curve, upon which many ordnance tables are based. The resistance coefficient varies greatly with the form of the projectile, boat-tailing in particular resulting in a marked decrease in the resistance.

A few measurements have also been made at speeds above that of sound, using an expanding orifice designed for a specified air speed. The force on the projectile was found to vary slightly with its distance from the orifice, indicating the existence of stationary pressure waves in the jet. This condition was never encountered when working at speeds below that of sound, and an investigation of small expanding orifices is now under way to determine if possible the conditions under which parallel lines of flow in the issuing jet may be secured.

At speeds above that of sound, a stationary sound-wave appears in the jet in front of the projectile. Its presence can be readily shown by projecting it on a white screen by means of sunlight or an electric arc, and under favourable lighting conditions it can also be seen with the unaided eye as a beautiful shimmering surface.

Aerodynamical Characteristics of Airfoils at High Speeds.

The high speed wind stream has also been utilized in measurements of the characteristics of airfoils with particular reference to propeller design*. Six airfoils with camber ratios ranging from 0.1 to 0.2 were tested at speeds from 550 to 1000 feet per second, which includes the speed range of the effective

**Aerodynamical Characteristics of Airfoils at High Speeds.* L. J. Briggs, G. F. Hull, and H. L. Dryden. Technical Report No. 207, Nat. Adv. Comm. for Aeronautics, Washington, 1924.

sections of most air propellers. The chord length was 3 inches and the span 17 inches, so that the ends of the airfoil extended well beyond the boundary of the 12-inch jet.

A special balance was designed for these measurements, which is represented diagrammatically in Fig. 2. The airfoil *m* is held by its ends in a fork *f* which is outside the air stream. The fork is supported on a pivot *c* which passes through the arm *a*, the latter being restrained from rotating about the centre pivot *c* by a sylphon element *d*. This sylphon, which is a thin-walled cell of circular cross-section with deep accordion-like corrugations in its walls

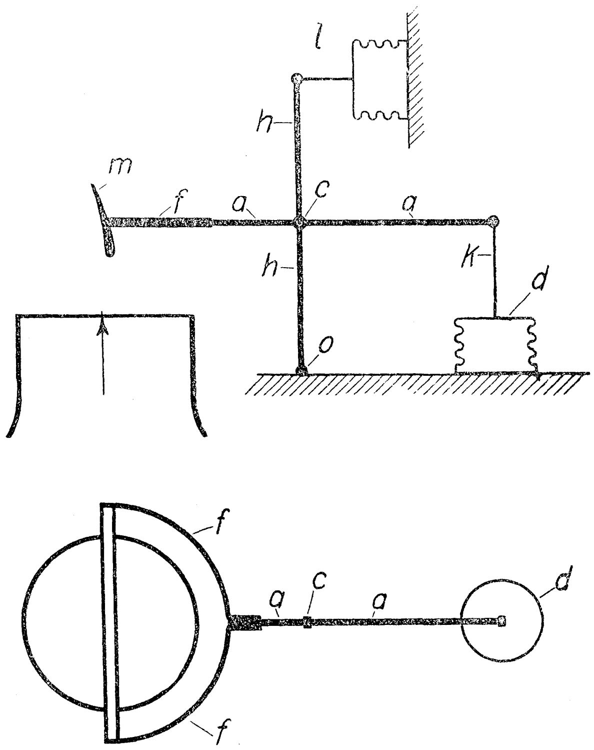


Fig. 2

Schematic Diagram of Balance Used in Measuring the Lift and Drag Coefficients of Airfoils at High Speeds.

is filled with oil and connected with a pressure gauge. The drag component of the force which tends to rotate the fork arm about its centre pivot is resisted by the sylphon, the resulting pressure being read off on the gauge. A small correction to the drag reading is necessary, depending on the position of the centre of pressure. The column *h* which supports the pivot *c* is also free to rotate about an axis *o*, and is restrained in such motion by a second sylphon *l*, which measures the lift component directly.

Means are provided for rotating the airfoil to vary the angle of attack. The distance of an axis of rotation from the leading edge can also be shifted. By rotating the airfoil in the wind-stream a position can be found where the torque on the airfoil is zero. For this angle of attack the line of resultant force passes through the axis, so that the centre of pressure can be determined. By changing the position of the axis of rotation with respect to the leading edge, a new angular position of (unstable) equilibrium is found. In this way the shift of the centre of pressure as the angle of attack changes can be readily determined.

The results of our experiments show that airfoils of the form commonly used as propeller sections have the following characteristics at high speeds:

1. The lift coefficient for a fixed angle of attack decreases very rapidly as the speed increases.

2. The drag coefficient increases rapidly.

3. The centre of pressure moves back toward the trailing edge.

4. The speed at which the rapid change in the coefficients begins is decreased (1) by increasing the angle of attack, and (2) by increasing the camber ratio.

5. The angle of zero lift shifts to high negative angles up to the "critical speed" and then moves rapidly toward 0°.

Fabric Tension Meter

In covering the framework of a rigid airship with fabric, it is desirable to know the tension in the envelope, since transverse stresses on the girders sufficient to crush the ship can easily be applied in this way. On the other hand, unnecessary looseness in the fabric increases the tendency to flap. An interesting and useful instrument for measuring the tension of an airship envelope in place has been developed by Messrs. Keulegan, Tuckerman and Eaton.

The base of the instrument consists of a flat plate which rests against the fabric. In the centre of the base is an elliptical hole 4×16 inches opening into the inner chamber of the instrument. The portion of the base plate between the elliptical opening and the outer edge of the plate is perforated with a large number of small holes which open into a second outer chamber surrounding the inner one. By applying suction to this outer chamber the instrument may be firmly attached to the envelope, the part of the fabric opposite the elliptical opening being left free.

Within the inner chamber is a lever arm ending in a ball which rests on the fabric at the centre of the elliptical opening, the position of the ball being shown by a deflection gauge. When the instrument is attached to the fabric in the desired position by means of the suction rim, the pressure in the inner chamber is slowly reduced until the fabric has been deflected into the chamber by a standard amount, and the pressure difference required to produce this deflection is read from a specially designed gauge. In single ply fabric one reading is made with the minor axis of the ellipse parallel to the warp of the fabric and a second reading is taken with the instrument turned 90°. Suitable safety devices are provided so that excessive pressures cannot be applied to the fabric.

The relation between the applied hydrostatic pressure P and the tension in the fabric is given by the equation

$$P=\frac{S_1+s_1}{R_1}+\frac{S_2+s_2}{R_2}$$

where R_1 and R_2 are the radii of curvature of the fabric (determined by the dimensions of the elliptical chamber and the deflection of the fabric), S_1 and S_2 are the initial tensions in the fabric in the corresponding arcs, and s_1 and s_2 are subsidiary tensions arising from the stretching of the fabric in the instrument. A description of the various procedures which may be employed to evaluate S_1 and S_2 cannot be undertaken here, but it is evident that a solution can be obtained if observations are made of the pressures corresponding to two standard deflections a and b in each position, and if the elastic modulus of the material in tension is known. In practice a nomogram has been found very useful in computing the tensions.

The above discussion refers to single-ply fabrics. The stresses in multiple-ply fabrics may be found by taking readings in four positions, oriented 0°, 45°, 90° and 135° to any arbitrarily chosen reference line.

The value of the instrument has been demonstrated by the Bureau of Aeronautics in connection with measurements on the Shenandoah (See Tech. Paper No. 320, U.S. Bureau of Standards, 1926).

Suspended Head Airspeed Meter

Another useful airship instrument is the suspended head electrical airspeed indicator developed by Messrs. Frymoyer, Henrickson and Eaton for use on the Shenandoah. The head of the instrument is suspended 40 feet or more below the car so as to be free from any disturbances in air flow produced by the hull of the airship, and is connected by a cable to the indicating instruments which are located at convenient points in the navigator's car. The streamline head is equipped with stabilizing surfaces to keep it pointed into the wind and carries in its nose a commutator driven by a small propeller. The commutator as it rotates alternately charges a condenser from a source of constant E.M.F. and discharges it through one or more indicating instruments connected in series, which indicate the speed directly in knots. The pointer of the indicator is steady for speeds as low as two knots, and the readings can be depended upon to one knot over the whole scale (70 knots) at all altitudes at which the ship can fly.

Radio-Acoustic Method of Locating a Ship's Position

The U.S. Coast and Geodetic Survey is now engaged in an extensive survey of the ocean floor off the Pacific coast. The position of the ship when taking a sounding has heretofore been based upon visual observations of the bearing of known shore stations. Such observations cannot be made in the presence of fog or haze, and in consequence sounding operations have often been limited to five or six days a month. Furthermore, a smooth sea which is highly advantageous in depth sounding work is very likely to be accompanied by fog.

The limitations placed upon the work through poor visibility were so serious that Messrs. Eckhardt and Keiser undertook, in co-operation with Mr. Heck of the Coast Survey*, to develop a method of locating the ship which would be

*Special Publication No. 107, U.S. Coast and Geodetic Survey, *A Radio-Acoustic Method of Position Finding for Use in Hydrographic Surveying*, by N. H. Heck, E. A. Eckhardt and M. Keiser.

independent of visibility conditions. It was essential that the ship should be the base of operations. The procedure which has finally been adopted is to send from the ship an under-water sound signal, recording the instant of its departure. This sound signal, upon its arrival at a hydrophone station near the shore, automatically starts the transmission of a radio signal which is in turn received and recorded on a chronograph on board the ship.

The hydrophone, which is simply a microphone adapted to under-water service, is supported near the sea bottom in 60 to 100 feet of water and is connected by a cable to a shore station, where necessary auxiliary equipment is installed.

Assuming for the moment that the apparatus is free from lag, the elapsed time between the firing of the bomb and the reception of the radio signal, multiplied by the speed of sound in sea water, gives the distance of the ship from the hydrophone station. If the same sound signal is used to measure the distance of the ship from each of two hydrophone stations, suitably located along the shore in known positions, the ship's position can be immediately determined. The inclusion of a third station provides a check on the accuracy of the measurements.

The sound signal is produced by the detonation of a small bomb of T.N.T. suspended from a raft towed by the ship. The signal from a three-pound bomb has been recorded on hydrophones sixty miles away. At the instant the bomb is exploded, a record is automatically made on the ship's chronograph, and a warning radio signal is sent out to the shore station operators.

When the sound wave strikes the hydrophone it changes the intensity of the current in the hydrophone circuit leading to the shore station. This change in current strength is amplified to operate a relay, which in turn starts an automatic sending key driven by clockwork. The radio signal is, however, not sent out immediately, a known time lag purposely being introduced to avoid signals from two stations arriving at the same time. The initial signal is followed by a code signal which identifies the sending station. To determine the overall time lag from each station, the shore apparatus is so arranged that it may be actuated by a wireless signal sent from the ship in place of the sound signal, the equipment otherwise being unchanged throughout. This time lag is subtracted from the time recorded when a bomb is used, which gives the time required for the sound wave to travel from the raft to the hydrophone. On clear days check measurements of the speed of sound in sea water may be made by determining the position of the ship independently from bearings on known shore stations. This method of finding the position of a ship at sea is now in daily use by the Coast Survey S.S. Guide with gratifying results.

THE USE OF THE INTEGRAPH IN THE PRACTICAL SOLUTION OF DIFFERENTIAL EQUATIONS BY PICARD'S METHOD OF SUCCESSIVE APPROXIMATIONS

BY DR. THORNTON C. FRY,

Research Laboratories of the American Telephone and Telegraph Company and the Western Electric Company, Inc., New York, N.Y., U.S.A.

INTRODUCTION

1. Many engineering problems are easily formulated in terms of differential equations. Frequently these are of a simple type and easily solved, although this is by no means universally true. When they are more difficult they can generally be solved by the use of suitable series expansions, but this entails labour and expense which may not be justified by the economic importance of the information obtained.

Bessel's equation may serve as an illustration. It has no simple solution in terms of the elementary functions, and if it were necessary to start from the beginning and develop a solution without reference to past studies, the cost might very well be prohibitive and the problem might have to go unsolved.

There are many equations, inherently no more difficult than Bessel's, which are of less general interest and have not been studied so extensively. If such equations are to be solved at all, so far as industrial mathematics is concerned, some simple method of procedure must be developed for the purpose. It is not necessary that this method yield absolute accuracy, for few of the magnitudes with which the engineer has to deal are more than approximations. It is sufficient for it to compare favourably with the best standards of engineering accuracy.

The method explained in this paper has been found valuable in treating problems of this kind. It is founded upon the Picard method of successive approximations, from which it differs only in certain minor details.

Theoretically, therefore, it is capable of yielding any desired degree of accuracy, if the computer is persistent enough; but practically it requires the use of some form of mechanical integration, and the accuracy of the results is therefore limited by that of the mechanical device.

What means of integration is used is a matter to be decided on the basis of convenience. From the standpoint of speed the most satisfactory results are obtained with the integraph, and as the errors thus produced are comparable with the inaccuracies of curve plotting they are sufficiently small to be unimportant in most phases of applied mathematics.

Outline of the Picard Method of Successive Approximations

2. Suppose, to start with, that the problem presents itself, not as a single equation defining a single function y, but as a set of n equations defining n functions, $y, y', \ldots, y^{(n-1)}$. Suppose, moreover, that all these equations are of the first order*, so that they may be written in the form

$$
(1) \qquad
\begin{aligned}
\frac{dy}{dx} &= f\,(x, y, y', \ldots, y^{(n-1)}), \\
\frac{dy'}{dx} &= f'\,(x, y, y', \ldots, y^{(n-1)}), \\
&\cdots\cdots\cdots\cdots\cdots\cdots \\
\frac{dy^{(n-1)}}{dx} &= f^{(n-1)}(x, y, y', \ldots, y^{(n-1)}),
\end{aligned}
$$

and that the desired solution is required to satisfy the conditions

$$
(2) \qquad y=\eta,\ y'=\eta',\ \ldots\ldots,\ y^{(n-1)}=\eta^{(n-1)},
$$

at $x=\xi$.

If the true solution of these equations were known, it would be a set of n explicit functions of x; and if all the y's in the right-hand members of (1) were replaced by these known functions, the equations would reduce to an exceedingly simple form, in which the x-derivatives of each of the y's would be given directly as functions of x. The equations could then be integrated directly, and made to satisfy the boundary conditions. Formally, the result would be:

$$
(3) \qquad
\begin{aligned}
y &= \eta + \int_\xi^x f(x, y, y', \ldots, y^{(n-1)})dx, \\
y' &= \eta' + \int_\xi^x f'(x, y, y', \ldots, y^{(n-1)})dx, \\
&\cdots\cdots\cdots\cdots\cdots\cdots \\
y^{(n-1)} &= \eta^{(n-1)} + \int_\xi^x f^{(n-1)}(x, y, y', \ldots, y^{(n-1)})dx.
\end{aligned}
$$

*This restriction is, of course, largely a formal one, for it is possible to write any set of ordinary differential equations in the form of a set of first order equations. Thus, the equation

$$
\frac{d^3y}{dx^3}+\frac{d^2y}{dx^2}+\left[\frac{dy}{dx}\right]^2+2y-x=0,
$$

may be thrown into the form

$$
\frac{dy}{dx}=y', \quad \frac{dy'}{dx}=y'', \quad \frac{dy''}{dx}=x-2y-y'^2-xy'',
$$

which is a special case of (1); and a similar transformation can obviously be applied to a set of equations as well. The generality of (1) is therefore very great.

This manipulation seems to have resulted in nothing new, for the y's defined by (3) are exactly the same as the y's that occur in the integrands. If they are known, nothing has been accomplished; if they are not known, no means of finding them has been presented. Nevertheless, the equations (3) are of value, in that they afford a means of checking a set of trial functions to determine whether or not they satisfy (1). For if these trial functions are substituted in the right-hand members of (3), and the indicated integrations are performed, a set of y's will be obtained which may or may not be identical with the trial functions. If they are, the trial functions are indeed the solution of (1); otherwise they are not.

The equations (3) are useful in another and less obvious respect, in that they make it possible to obtain from an incorrect set of trial functions a second set which is better than the first. As this second set can be similarly used in deducing a third, and so on, the process can be continued until it leads to a result which is so nearly identical with the true solution that the differences may be ignored. This is the essential idea back of the Picard method of successive approximation.

For the first approximation the functions $y, y', \ldots, y^{(n-1)}$ are assumed to be constants, and, in order that the boundary conditions may be satisfied, these constants are chosen as $\eta, \eta', \ldots, \eta^{(n-1)}$, respectively. These first approximations are then substituted in the right-hand members of (3), and the indicated integrations are performed, giving a new set of functions,

$$y_2,\ y_2',\ \ldots,\ y_2^{(n-1)},$$

which may or may not be identical with the η's. If it is, it follows that the η's were indeed the correct solution, and there is no need of further manipulation. But if not, the η's are incorrect, and they are therefore abandoned forthwith.

The y_2's are then themselves adopted as a second approximation and substituted in the right-hand side of (3). When the integrations have been performed, still another set of functions, $y_3, y_3', \ldots, y_3^{(n-1)}$, appears. This is the third approximation. Continuing the process, set after set of approximations is deduced, each related to the preceding one by the equations

$$y_{k+1}=\eta+\int_{\xi}^{x} f(x, y_k, y_k', \ldots, y_k^{(n-1)})dx,$$

$$y'_{k+1}=\eta'+\int_{\xi}^{x} f'(x, y_k, y_k', \ldots, y_k^{(n-1)})dx,$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$y_{k+1}^{(n-1)}=\eta^{(n-1)}+\int_{\xi}^{x} f^{(n-1)}(x, y_k, y_k', \ldots, y_k^{(n-1)})dx.$$

It may happen—and sometimes does happen—that one of these sets is the correct solution. If so, the process of obtaining one from the other is such that all succeeding sets are identical. In other words, the process has terminated in the correct solution. This outcome, however, is unusual. Ordinarily each set differs frcm all those which preceded it; but it can be proved that they differ progressively less and less from one another, and ultimately approach a set of limiting values which is itself the true solution of equations (1). In other words, *by starting from the preliminary guess that* $y, y', \ldots, y^{(n-1)}$ *are all constant, and by following a well-established routine involving only the integration of known functions of x, a result may ultimately be obtained which differs as little as we please from the desired solution of the set of equations.*

This process is capable of graphical interpretation, as is illustrated in Fig. 1 for a set of two equations defining the two functions y and y'. The first approximation—in which y and y' are assumed to be constant and equal to the boundary values 1 and 0—is represented by the straight lines y_1 and y_1' in the upper half of the figure. When these constants are substituted in f and f', they give the pair of functions f_2 and f_2' plotted in the lower half of the figure. The second approximations y_2 and y_2' (shown in the upper half of the diagram) are the integrals of these functions, the constants of integration having been so chosen that the boundary values are again satisfied. These second approximations in turn are substituted in the functions f, and lead to a new set of curves f_3 and f_3' which appear in the lower half of the diagram.

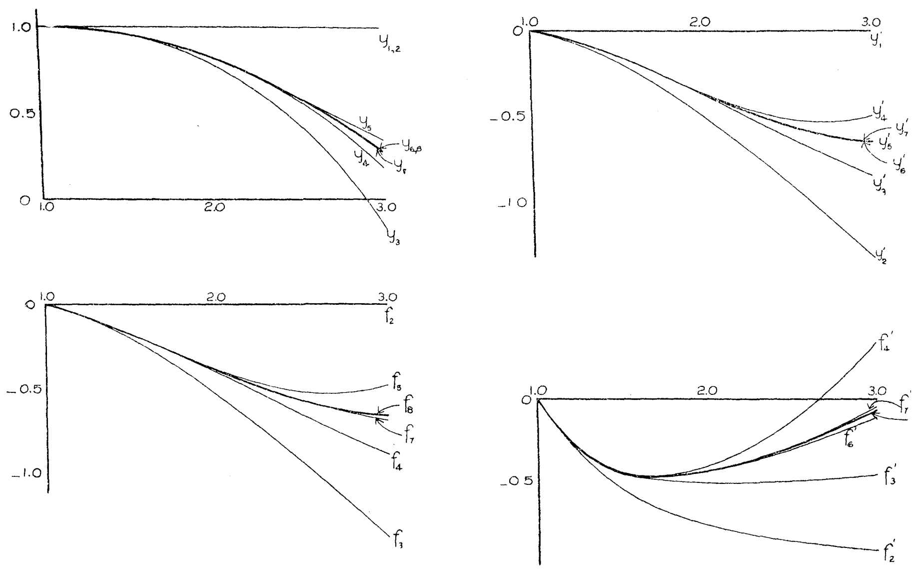


Fig. 1

Thus the process continues. From each set of f-curves a pair of y-curves is obtained by a straightforward process of integration, and then the next set of f-curves is found by substituting the new y's in the right-hand members of (1).

Theoretically there is nothing objectionable about this form of solution. *Practically* it assumes that the integrations can all be performed, which is unfortunately not often the case. That is, although there is usually no difficulty in passing from a set of y-curves to the set of f-curves determined by it, the process of passing from this latter to the next succeeding set of y-curves is often too tedious to be feasible. If by some means this objection can be overcome, the Picard method can be made to serve a useful purpose in obtaining numerical results, as well as in theoretical studies.

In those cases where graphical representation is possible—which means that the equations must be reducible to a form which does not contain unknown parameters—the integraph affords a satisfactory solution of this difficulty. This instrument, which is shown in Fig. 2, is in a sense a planimeter, but it differs from ordinary planimeters in two respects: it performs its integrations in the

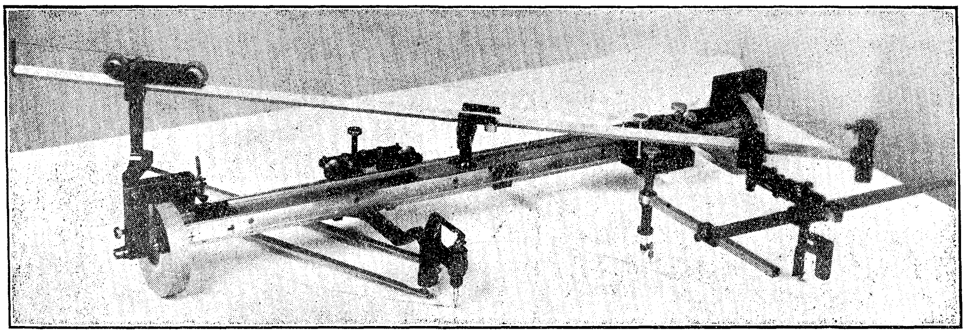


Fig. 2

rectangular coordinate system, and it actually traces the integral curve by means of a ruling pen attached to the machine. Thus, if the tracing point were caused to follow the curve f_2 of Fig. 1, the pen—having been initially placed upon the point specified by the boundary conditions—would actually draw the integral curve y_2.

The instrument is easily controlled, is not difficult to adjust, and seldom departs from the true solution by as much as the width of the ink line which it draws. Its accuracy is therefore sufficient to meet most practical requirements.

Example I. An Illustration of the use of the Integraph in Picard's Method

3. The curves reproduced in Fig. 1 were obtained by this method of procedure. They are solutions of the set of equations:

$$\frac{dy}{dx}=y', \quad \frac{dy'}{dx}=\left(\frac{1}{x^2}-1\right)y-\frac{1}{x}y', \tag{4}$$

subject to the boundary conditions

$$y=1, \quad y'=0,$$

at $x=1$.

It therefore follows, from comparison of (4) with (1), that the functions f are

$$f(x, y, y')=y', \; f'(x, y, y')=\left(\frac{1}{x^2}-1\right)y-\frac{1}{x}y'. \tag{5}$$

The first set of trial curves, in which y and y' are constants, is denoted by y_1 and y_1'. Upon substitution of these constants in (5) the functions f take the forms

$$f_2=0,\ f_2'=\frac{1}{x^2}-1.$$

They are plotted in the lower half of the figure.

It so happens that these functions could be readily integrated without mechanical aid. In fact, in the case of this particular illustration several approximations could be obtained before the integrations become arduous. Nevertheless in the construction of Fig. 1 all the y and y' curves were drawn by the integraph; for it was easier to produce them in that way than to make the computations incidental to plotting them. In particular, y_2 and y_2' are the integrals obtained by tracing f_2 and f_2'.

To obtain f_3 and f_3' these new functions y_2 and y_2' must be substituted in (5). This is done by reading values from the curves and substituting them in (5), continuing the process for as many values of x as are necessary to permit the f-curves to be drawn in. These new f-curves, f_3 and f_3', are then integrated to give y_3 and y_3', from which in turn f_4 and f_4' are deduced by substitution.

This cycle is repeated again and again, until two consecutive sets of y's are so nearly identical as to be indistinguishable on the scale of the drawing. When this stage has been reached, the last curves are accepted as the solution of (4).

As an illustration of the amount of labour required to carry out this process, the complete set of computations is reproduced in Table I. It was first decided to extend the solution as far as $x=3$, and the tabular interval 0.2 was fixed upon as a suitable one from the standpoint of curve plotting. Then the factors $\frac{1}{x^2}-1$ and $\frac{1}{x}$, by which y and y' are multiplied in (5), were computed and recorded in columns 2 and 3, in order that they might be readily available for the work that was to follow.

As f_2' is equal to $\frac{1}{x^2}-1$, which has already been recorded, and $f_2=0$, it was not necessary to give these numbers space in the table. Also, y_2 being constant and equal to 1, there was no need of recording it. But y_2' is variable, and its values were therefore entered in column 4.

By equation (5), $f_3=y_2'$. Hence, the graph of f_3 was obtained by plotting the numbers given in column 4. To get f_3' it was necessary to carry out the processes indicated in the second of equations (5): multiplying corresponding entries in columns 3 and 4, and subtracting them from column 2. The results are shown in column 5, and when plotted give the curve f_3'.

The other columns up to 21 are obtained in a similar fashion, those headed f' being computed from the next preceding y and y', and those headed y or y' being read directly from the curves of Fig. 1. When y_8 had been obtained it was obvious that no further improvement was possible. Therefore y_8 was accepted as the true solution.

TABLE I

COMPUTATIONS INCIDENTAL TO THE SOLUTION OF EXAMPLE I

x	$\frac{1}{x^2}-1=f_2'$	$\frac{1}{x}$	$y_2'=f_3$	f_3'	y_3	$y_3'=f_4$	f_4'
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1.0	0.0000	1.0000	0.000	0.000	1.000	0.000	0.000
1.2	−0.3056	0.8333	−0.033	−0.278	0.998	−0.031	−0.279
1.4	−0.4898	0.7143	−0.113	−0.409	0.984	−0.101	−0.410
1.6	−0.6094	0.6250	−0.224	−0.469	0.951	−0.190	−0.460
1.8	−0.6914	0.5556	−0.354	−0.495	0.894	−0.286	−0.459
2.0	−0.7500	0.5000	−0.498	−0.501	0.809	−0.386	−0.414
2.2	−0.7934	0.4545	−0.653	−0.496	0.694	−0.486	−0.330
2.4	−0.8264	0.4167	−0.814	−0.487	0.548	−0.584	−0.210
2.6	−0.8521	0.3846	−0.982	−0.474	0.370	−0.680	−0.054
2.8	−0.8724	0.3571	−1.155	−0.460	0.156	−0.774	−0.140
3.0	−0.8889	0.3333	−1.330	−0.446	−0.096	−0.864	0.368

x	y_4	$y_4'=f_5$	f_5'	y_5	$y_5'=f_6$	f_6'	y_6	$y_6'=f_7$	f_7'
(1)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)
1.0	1.000	0.000							
1.2	0.998	−0.031							
1.4	0.986	−0.102	−0.410	0.986					
1.6	0.955	−0.190	−0.463	0.955					
1.8	0.910	−0.282	−0.472	0.910	−0.284	−0.471			
2.0	0.843	−0.370	−0.447	0.844	−0.376	−0.445			
2.2	0.756	−0.445	−0.398	0.763	−0.460	−0.396	0.760	−0.460	−0.394
2.4	0.649	−0.500	−0.328	0.668	−0.533	−0.330	0.661	−0.533	−0.324
2.6	0.524	−0.528	−0.243	0.565	−0.591	−0.254	0.549	−0.591	−0.241
2.8	0.379	−0.520	−0.145	0.459	−0.630	−0.175	0.427	−0.636	−0.145
3.0	0.216	−0.471	−0.035	0.359	−0.648	−0.103	0.299	−0.662	−0.045

x	y_7	$y_7'=f_8$	f_8'	y_8	correct value	error
(1)	(18)	(19)	(20)	(21)	(22)	(23)
1.0				1.000	1.000	0.000
1.2				0.998	0.998	0.000
1.4				0.986	0.985	−0.001
1.6				0.955	0.956	+0.001
1.8				0.910	0.908	−0.002
2.0				0.844	0.842	−0.002
2.2				0.760	0.759	−0.001
2.4				0.661	0.659	−0.002
2.6	0.549	−0.588	−0.242	0.549	0.547	−0.002
2.8	0.425	−0.627	−0.147	0.427	0.425	−0.002
3.0	0.296	−0.646	−0.048	0.300	0.297	−0.003

To check the accuracy of the result, it was noted that when the relationship $\frac{dy}{dx}=y'$ is substituted in the second of equations (4), the set reduces to Bessel's Equation

$$\frac{d^2y}{dx^2}+\frac{1}{x}\frac{dy}{dx}+\left(1-\frac{1}{x^2}\right)y=0, \tag{6}$$

the solution of which, subject to the assigned boundary conditions, is

$$y=1.4034\ J_1(x)-0.3251\ Y_1(x).$$

The correct values of y, as computed from this equation, are recorded in column 22. They make it possible to determine the accuracy of the graphical computation, which turns out to have been nowhere in error by more than three units in the third decimal place. This represents less than one fiftieth of an inch on the scale of the original drawing.

Example II. Illustrating an Alternative form of Procedure When the Equations Are Not of the First Order

4. There is one outstanding peculiarity about Fig. 1. It is that the diagrams for f and y' are—as a whole—identical. This comes about through the fact noted in the last paragraph, that the set is equivalent to a single equation of higher order, which fact manifests itself in the simple form of the first of equations (4).

Because of this peculiarity, each y-curve is obtained by *double* integration of an f', the first integration giving a y', and the second a y. But as the cycle of operations now stands, this y and y' do not belong to the same set of approximations. It is not y_k' and y_k which are the successive integrals of f_k', but y_k' and y_{k+1} instead. In other words, although equations (4) are of such a form that the symbol y' must necessarily stand for $\frac{dy}{dx}$, the y' which is assigned to a particular set of approximations is not the derivative of the y which belongs to the same set.

It would be more elegant to have each set of curves stand in their normal differential relations with respect to one another, an improvement which is easily brought about if the first integral of f_k' is denoted by y_k', and its second integral by y_k (not y_{k+1}). These approximations can then be used in exactly the same way as were those previously denoted by the same symbols. But in so doing, the traditional cycle is modified, and as the classical proofs of the convergence of the process depend upon this traditional cycle, the question arises as to whether the new cycle, like the old, can be relied upon to lead to a solution. The answer is that it may, for it is a comparatively simple exercise in pure mathematics to so revise the classical argument as to apply to the modified procedure. Moreover, it is a matter of experience that such problems as are met in engineering practice can frequently be solved in this way with less labour than would be occasioned by the classical procedure.

To compare this modified cycle with the original one equation (6) has been solved by the revised method. The results are shown in Fig. 3. The approximation curves are obviously different, but they lead to substantially the same result in the end.

Each f-curve (as for instance f_3) is computed from the preceding set of y's (in this case, y_2 and y_2'); and from it the next set of y's (y_3 and y_3') is obtained by double integration. The computations are reproduced in Table II. The accuracy of the results obtained by this method is about the same as by the original one, and two less approximations are required. When the equation is of high order, the advantage in favour of the modified method is usually even more pronounced.

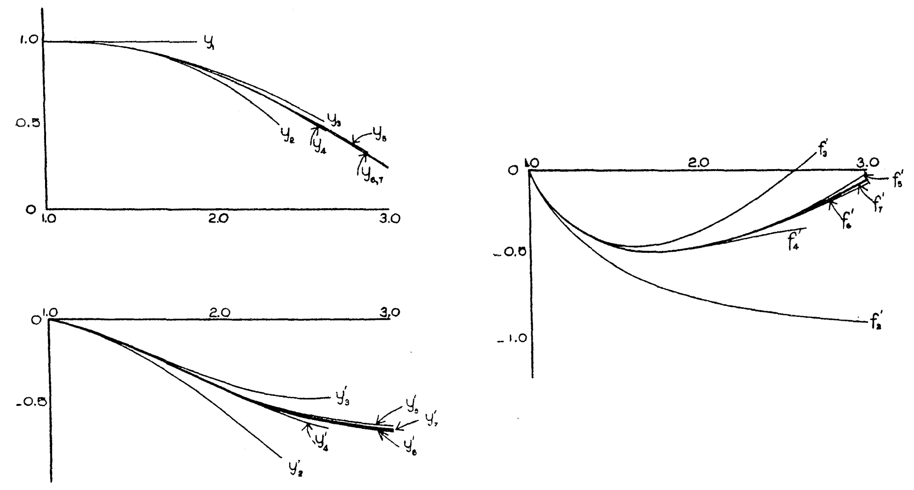


Fig. 3

It will be noted that the earlier approximations in Fig. 3 are not extended all the way to $x=3$. The reason is, that the inaccuracy of the earlier approximations in the vicinity of $x=3$ is so great that the f'-values deduced from them are grossly incorrect. The labour of computing them is therefore virtually wasted. If, instead of dealing with the entire interval at the start, it is partitioned off into a number of sub-intervals, all but the first can be ignored until the solution has been completed over this limited range. The values of y and y' at the end of the first sub-interval then form a new set of boundary values, by means of which the solution may be extended to the second interval. In this way, one sub-interval after another may be dealt with, until the desired range of values is obtained.

If this routine is strictly adhered to, the first approximation in the second region will be a set of constants. But when the straight lines corresponding to these constants are joined to the correct curves already found, they are likely to be obviously incorrect: so obviously incorrect, in fact, that the computer will have no doubt as to his ability to draw a better approximation free hand. Under these circumstances, the labour is simplified by making the best guess possible, and using it instead of the straight lines required by the classical scheme. More-

TABLE II

COMPUTATIONS INCIDENTAL TO THE SOLUTION OF EXAMPLE II

x	$\frac{1}{x^2}-1=f_2'$	$\frac{1}{x}$	y_2'	y_2	f_3'	y_3'	y_3
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1.0	0.0000	1.0000	0.000	1.000	0.000	0.000	1.000
1.2	−0.3056	0.8333	−0.033	0.998	−0.278	−0.032	0.998
1.4	−0.4898	0.7143	−0.115	0.983	−0.399	−0.102	0.985
1.6	−0.6094	0.6250	−0.224	0.950	−0.439	−0.187	0.955
1.8	−0.6914	0.5556	−0.356	0.892	−0.419	−0.275	0.905
2.0	−0.7500	0.5000	−0.502	0.805	−0.353	−0.354	0.840
2.2	−0.7934	0.4545	−0.656	0.690	−0.249	−0.415	0.762
2.4	−0.8264	0.4167	−0.818	0.542	−0.107	−0.451	0.674
2.6	−0.8521	0.3846				−0.451	0.582
2.8	−0.8725	0.3571					
3.0	−0.8889	0.3333					

x	f_4'	y_4'	y_4	f_5'	y_5'	y_5
(1)	(9)	(10)	(11)	(12)	(13)	(14)
1.0	0.000					
1.2	−0.278					
1.4	−0.410					
1.6	−0.463	−0.191	0.955	−0.463	−0.190	0.955
1.8	−0.473	−0.386	0.905	−0.467	−0.283	0.907
2.0	−0.453	−0.379	0.840	−0.440	−0.375	0.840
2.2	−0.416	−0.466	0.755	−0.387	−0.458	0.757
2.4	−0.369	−0.544	0.652	−0.312	−0.528	0.658
2.6	−0.320	−0.613	0.536	−0.221	−0.581	0.546
2.8					−0.615	0.427
3.0					−0.627	0.303

x	f_6'	y_6'	y_6	y (calculated)	error
(1)	(15)	(16)	(17)	(18)	(19)
1.0			1.000	1.000	0.000
1.2			0.998	0.998	0.000
1.4			0.985	0.985	0.000
1.6	−0.463		0.955	0.956	+0.001
1.8	−0.470		0.907	0.908	+0.001
2.0	−0.442	−0.377	0.840	0.842	+0.002
2.0	−0.392	−0.461	0.757	0.759	+0.002
2.4	−0.324	−0.533	0.657	0.659	+0.002
2.6	−0.242	−0.590	0.543	0.547	+0.004
2.8	−0.153	−0.630	0.422	0.425	+0.003
3.0	−0.060	−0.651	0.294	0.297	+0.003

over, this procedure is justified by the fact that the straight lines themselves are pure guesses, the use of which has no other foundation than the *a posteriori* fact that it leads to a solution.

In practice it is generally advisable to make such a guess, not when the first interval has finally been disposed of as accurately as possible, but shortly before. By so doing, the same steps of the process which serve to eliminate the last small errors from the one part of the curves also partially correct the grosser errors in the operator's guess.

It is by this means that the curves of Fig. 3 were obtained. Thus f_3' was extended from $x=2.4$ to 2.6, and f_5' to 3.0, by guess, although the correct result at 2.4 was not obtained until the last approximation had been completed.

Example III. A Problem in Electrical Circuit Theory, Illustrating the Solution of Simultaneous Differential Equations of the Second Degree

5. Sometimes the modified procedure explained in §4 is applicable even when the set of equations is not reducible to a single equation of higher order. An illustration may be obtained from the consideration of the electrical circuit shown in Fig. 4.

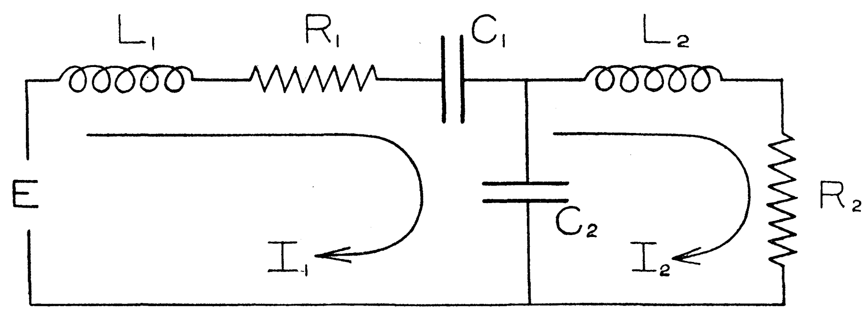


Fig. 4

If x_1 represents the instantaneous charge on the condenser C_1, and x_1-x_2 the instantaneous charge on C_2, the differential equations controlling the behaviour of the circuit are

$$(7)\qquad \begin{aligned} &L_1\frac{d^2x_1}{dt^2}+R_1\frac{dx_1}{dt}+\left[\frac{1}{C_1}+\frac{1}{C_2}\right]x_1-\frac{1}{C_2}x_2=E,\\ &-\frac{1}{C_2}x_1+L_2\frac{d^2x_2}{dt^2}+R_2\frac{dx_2}{dt}+\frac{1}{C_2}x_2=0. \end{aligned}$$

Such a set of equations can be rewritten in terms of first order equations by introducing new variables in place of $\frac{dx_1}{dt}$ and $\frac{dx_2}{dt}$; and in this instance this happens to have physical significance, since these derivatives are the currents I_1 and I_2, respectively. The four equations which result are:

$$
(8) \qquad
\begin{aligned}
\frac{dx_1}{dt} &= I_1, \qquad L_1\frac{dI_1}{dt} = E - R_1 I_1 - \left(\frac{1}{C_1}+\frac{1}{C_2}\right)x_1 + \frac{1}{C_2}x_2,\\
\frac{dx_2}{dt} &= I_2, \qquad L_2\frac{dI_2}{dt} = \frac{1}{C_1}x_1 - R_2 I_2 - \frac{1}{C_2}x_2.
\end{aligned}
$$

To the equations in this form the Picard method might be applied directly; but if this were done the currents assigned to any one approximation would not be the derivatives of the corresponding charges, as they must be because of a fundamental physical law. It is therefore more satisfactory to deal with the equations in the original form (7) instead of transforming them into (8).

Suppose for example, that the constants have the values

$$
\begin{aligned}
L_1 &= 0.1, & L_2 &= 1,\\
R_1 &= 20, & R_2 &= 100,\\
C_1 &= 10^{-5}, & C_2 &= 10^{-5},
\end{aligned}
$$

and suppose that x_1 and x_2 are replaced by x and y respectively, in order that no confusion may arise as to the significance of the subscripts. Then (7) reduces to the form

$$
(9) \qquad
\begin{aligned}
x'' &= 10\,E - 200\,x' + 10^6\,y - 2.10^6\,x,\\
y'' &= 10^5\,x - 10^5\,y - 100\,y',
\end{aligned}
$$

the primes denoting differentiation.

The boundary values, which afford the starting point for the solution, are determined by physical conditions. If it is assumed that no electromotive force is impressed on the circuit prior to $t=0$, they are

$$x = 0, \quad x' = 0, \quad y = 0, \quad y' = 0.$$

As usual these boundary values are taken as the first approximations, and called x_1, x_1', y_1, y_1'. (See Fig. 5.) When substituted in (9) they give

$$x_2'' = 10\,E,$$

from which the second approximations x_2' and x_2 are obtained by double integration; and also

$$y_2'' = 0,$$

from which y_2' and y_2 are similarly obtained. These are again substituted back in (9), and yield x_3'' and y_3'', the entire cycle being identical with that of Example II, except in so far as it involves two variables. All these curves are shown in Fig. 5.

It should be noticed that it is not necessary to have an algebraic definition of E in order to carry out these processes. A graphical representation—such, for instance, as an oscillograph record or as the curve for $10E$ shown in Fig. 5—serves the purpose just as well, and is even more convenient than an algebraic definition which is hard to compute.

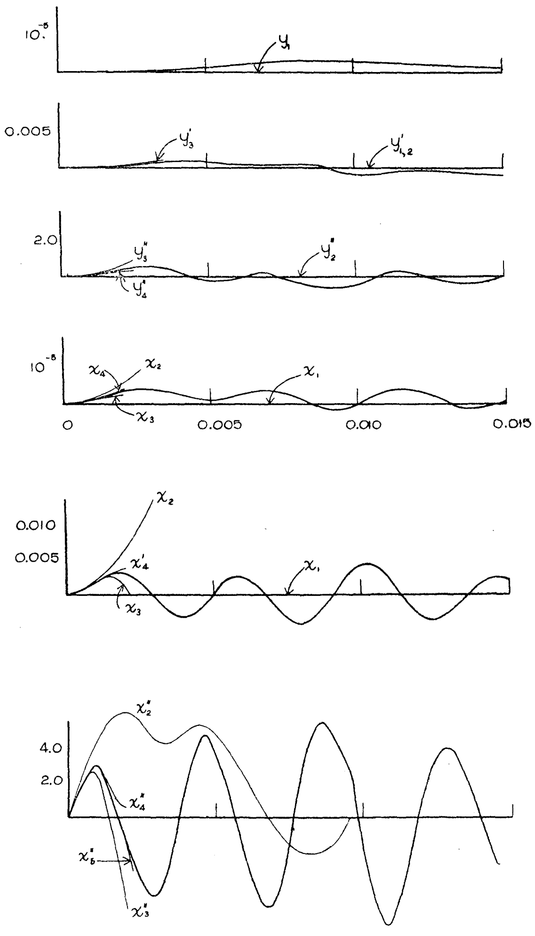


Fig. 5

This freedom from the necessity of algebraic definition is inherent in graphical work, and extends to every phase of it. It is not even necessary that the fundamental equations (1) themselves be expressible algebraically; any form of definition by means of which the numerical values of the f's can be obtained when the numerical values of their arguments are known is equally satisfactory. In the case of one variable, for example, the equation would appear as

$$\frac{dy}{dx}=f(x, y),$$

but the symbol $f(x, y)$ need not connote any algebraic law. Instead, the functional relationship may be given by some form of curve, or even by a more or less extensive double-entry table such as might be obtained from laboratory observation. If this table permits the numerical value of the derivative to be obtained when the numerical values of the arguments are given, it defines the functional relationship sufficiently well for the purposes of the problem.

The electromotive force assumed in the construction of Fig. 5 illustrates this point. It was sketched freehand and is representable by no known formula, yet the labour of solving the problem was no greater than it would have been had an algebraic definition been at hand.

The assumed driving force persists for 0.01 second, during which time it agitates the circuit, and the disturbance thus produced persists even after the driving force is removed. Thus, although the curves of Fig. 5 are extended to the time $t=0.015$, which is 0.005 second after the removal of the driving force, the disturbance has not greatly abated during this time. To extend the solution even farther requires only persistence, but the additional labour was not warranted by the uses to which the solution was to be put.

To carry out the solution as far as it is shown required 25 approximations. In order to avoid undue confusion only the first four of these are shown. The time consumed was about twelve hours. Whether this amount of labour is prohibitive or not depends upon the economic importance of the problem. It would be a comparatively unimportant industrial problem, however, that did not justify two days' work.

It is difficult to determine the accuracy of the result, owing to the indefinite nature of the driving force. In a rough way, however, some confirmation can be obtained by the argument which follows: In any two-mesh circuit such as this, the free oscillation which takes place after the driving force is removed consists of the sum of two damped sinusoids. In the present case, *one* of these is very pronounced, particularly in the x'' and x' curves. The other, which is of longer period, can be detected in the y's, where it has the appearance of a slow drift. As the frequencies of these free oscillations are independent of the nature of the driving force and can be algebraically computed, they afford a check upon the work.

Upon applying this check, it is found that the frequencies should be 33.8 and 227 cycles, the first affected by the damping factor e^{-51t}, and the second by the factor e^{-99t}. Computation from the various arches of the x'' curve indicates that the one frequency should lie between 225 and 231 cycles per second, as it does. The other constants are unobtainable.

Although this check is a very rough one, it is accurate enough to cause the detection of errors of the grosser sort, leaving only the possibility of minor inaccuracies, the probable magnitudes of which may be judged from those revealed by the other examples.

EXAMPLE IV. ILLUSTRATING THE PROCEDURE WHEN THE BOUNDARY VALUES CONSTITUTE A SINGULARITY OF THE f-FUNCTIONS

6. In Examples I and II, which deal with Bessel's Equation (6), the boundary conditions are applied at $x=1$. This is an ordinary point of the coefficients $\frac{1}{x^2}-1$ and $\frac{1}{x}$. If, however, the attempt is made to apply the boundary values at $x=0$, which is a singular point for these coefficients, the right-hand side of (5) becomes indeterminate, and the value of f' cannot be obtained. Similar difficulties may present themselves in a variety of guises, but almost all of them can be overcome by some sort of artifice. Frequently the best plan is to determine in advance by algebraic means a set of functions, y_1, y_1', . . . , to which the true solution is known to be asymptotic in the vicinity of the troublesome point, using this set as the first approximation in the Picard process. When this can be done, the difficulty usually vanishes, and the computation can be completed according to the standard routine.

In the case of equation (6), if the boundary values* are

$$y'=1, \quad y=0,$$

at $x=0$, the three terms

$$\frac{y'}{x}, \quad y, \quad -\frac{y}{x^2}$$

take the limiting forms

$$\frac{1}{0}, \quad 0, \quad \frac{0}{0}.$$

The last term is indeterminate and may either dominate the first or be negligible with respect to it, but the second term is obviously negligible by comparison with the first. Hence, in the immediate vicinity of $x=0$, (6) may be replaced by

$$\frac{d^2y}{dx^2}+\frac{1}{x}\frac{dy}{dx}-\frac{y}{x^2}=0.$$

This equation, however, is of a well-known type, the general solution of which is†

$$y=ax+\frac{b}{x}.$$

*If the boundary values on y and y' were taken the same as in Examples I and II, only the value of x being changed, no solution would exist. For the general solution, $y=AJ_1(x)+BY_1(x)$, may either vanish or become infinite in the neighbourhood of $x=0$, but cannot take a finite value. The attempt to apply these boundary values would therefore fail, not because of any deficiency in the method of solution, but because it is asked to do impossible things. The boundary valus used above are so chosen that a solution is possible.

†Note that this asymptotic solution, like the general solution of (6), cannot take a finite value different from zero at $x=0$.

If the boundary values are to be satisfied the constants a and b must be taken as 1 and 0, respectively. Hence

$$y_1 = x, \qquad y_1' = 1,$$

are approximate solutions in the neighbourhood of $x=0$. Accepting these as first approximations, and substituting them in (5), f_2' is found to be

$$\left(\frac{1}{x^2} - 1\right)x - \frac{1}{x},$$

which reduces to x, and is no longer troublesome.

The curves (y_1, y_1' and f_2') are shown in Fig. 6. When f_2' is twice integrated, the second set of approximations, y_2' and y_2, is obtained, and from these the remaining approximations are deduced by the usual cycle of operations.

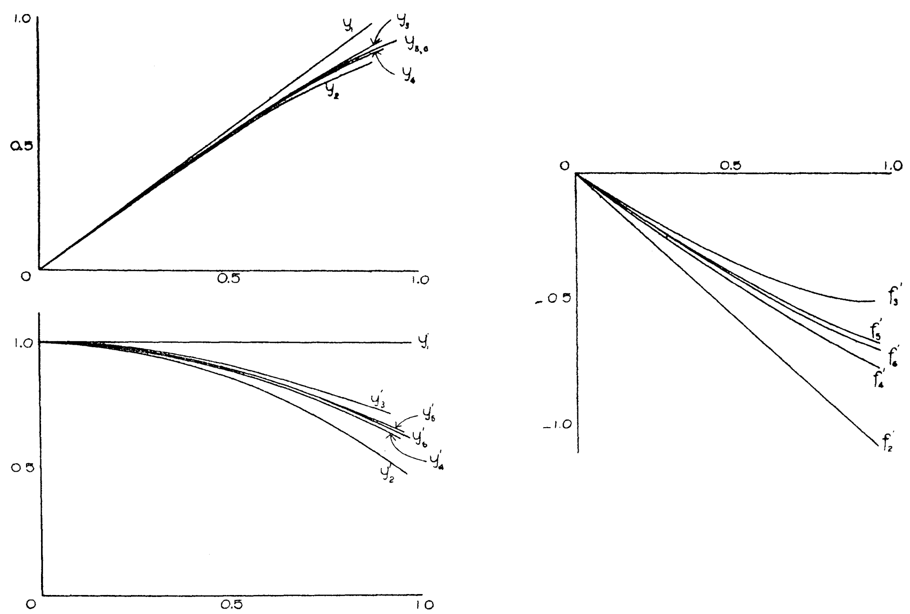


Fig. 6

To show the accuracy of the final results, they are displayed in Table III, along with the true values as computed from the known solution

$$y = 2J_1(x).$$

The maximum error is about one-fiftieth of an inch on the scale of the original drawing, six approximations being sufficient to give this result.

TABLE III

COMPARISON OF RESULTS OF THE GRAPHICAL AND THE ALGEBRAIC SOLUTION OF EXAMPLE IV

x	y (graphical)	y (calculated)	error
0.0	0.000	0.000	0.000
0.2	0.199	0.199	0.000
0.4	0.390	0.392	0.002
0.6	0.574	0.575	0.001
0.8	0.736	0.738	0.002
1.0	0.878	0.880	0.002

EXAMPLE V. ILLUSTRATING THE PROCEDURE WHEN THE BOUNDARY VALUES ARE NOT GIVEN FOR THE SAME VALUE OF x

7. The classical proof of the convergence of the Picard approximations requires that the boundary values be given in the form (2). But there are many other types of conditions which serve to define a solution equally well. For instance,

$$\begin{aligned} y &= \eta, & &\text{at } x = \xi_1, \\ y' &= \eta', & &\text{at } x = \xi_2, \\ &\cdots \\ y^{(n-1)} &= \eta^{(n-1)}, & &\text{at } x = \xi_{n-1}, \end{aligned}$$

is such a set, and the graphical solution can frequently be applied to cases of this sort. As an illustration, Bessel's Equation (6), may again be considered, subject to the conditions

$$\begin{aligned} y &= 0, & &\text{at } x = 3, \\ y' &= 1, & &\text{at } x = 1. \end{aligned}$$

As in Examples I and II, the constants

$$y_1 = 0, \quad y_1' = 1$$

are chosen as the first approximations and substituted in (5), just as if both boundary values were stated for the same value of x. The result is a second derivative curve,

$$f_2' = -\frac{1}{x},$$

which is plotted for integration (see Fig. 7). The first integral of this function is obtained in the usual way, but in tracing the second the pen of the integraph is set on the point (3, 0), and the machine is run *backward* to $x = 1$, thus assuring an integral which will take the proper boundary value. The curves thus obtained are accepted as y_2' and y_2. They are then substituted in the right-hand side of (5), a new second derivative curve f_3' is plotted, and the cycle repeated as often as is required to obtain the desired degree of accuracy. In the case of Fig. 7 nine approximations were required.

The solution of the equation being

$$y = -1.848 J_1(x) + 1.141 Y_1(x),$$

it is possible to determine the accuracy of the graphical results. The check is

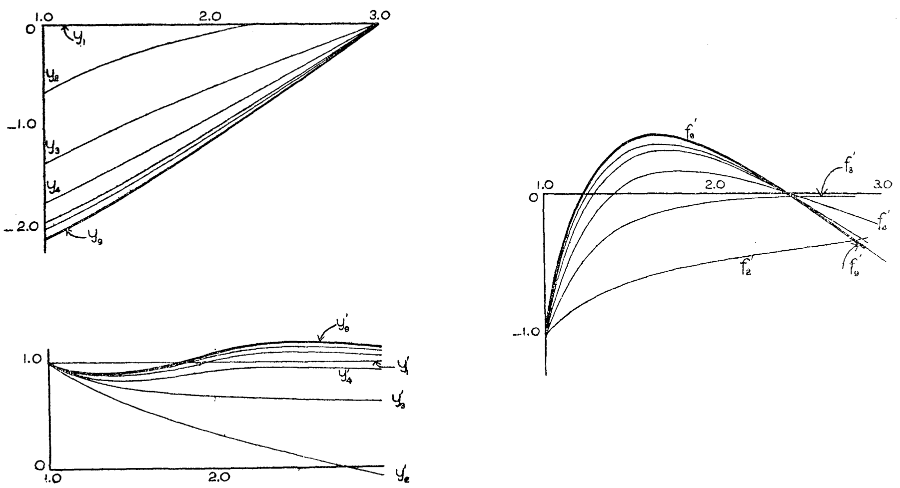


Fig. 7

shown in Table IV. The maximum error, on the scale of the original drawing is about one-fiftieth of an inch.

TABLE IV

COMPARISON OF RESULTS OF THE GRAPHICAL AND THE ALGEBRAIC SOLUTION OF EXAMPLE V

x	y (graphical)	y (computed)	error
1.0	−2.160	−2.155	+0.005
1.2	−1.971	−1.968	+0.003
1.4	−1.794	−1.788	+0.006
1.6	−1.606	−1.601	+0.005
1.8	−1.404	−1.398	+0.006
2.0	−1.186	−1.181	+0.005
2.2	−0.954	−0.951	+0.003
2.4	−0.714	−0.712	+0.002
2.6	−0.471	−0.470	+0.001
2.8	−0.229	−0.231	−0.002
3.0	0.000	0.000	0.000

One point worthy of notice in connection with this type of boundary condition is, that it is not possible to reduce the labour by dividing the entire interval into convenient sub-intervals. Instead, every approximation must be carried through the entire interval, no matter how obviously incorrect it may be over the major part of the range. It is only by so doing that the boundary values can be fitted into the graphical scheme at all.

EXAMPLE VI. ILLUSTRATING THE PROCEDURE WHEN ONE DERIVATIVE IS WITHOUT A BOUNDARY VALUE

8. In the important special case where the equations (1) are equivalent to a single n'th order equation, the curve representing y is frequently required to pass through n preassigned points. In such cases, the boundary conditions take the form:

$$\begin{array}{ll} y=\eta_1, & \text{at } x=\xi_1, \\ y=\eta_2, & \text{at } x=\xi_2, \\ \dots & \dots \\ y=\eta_{(n-1)}, & \text{at } x=\xi_{(n-1)}. \end{array}$$

With such boundary values it is not feasible to start the process of approximation by assuming constant values for $y, y', \dots, y^{(n-1)}$; because in the first place, the constant value assigned to y cannot be so chosen as to satisfy more than one of the n conditions imposed upon it, and because in the second place, nothing at all is known about the various derivatives, and there is therefore no rational means of assigning appropriate constants to them. Hence it is necessary to add to the usual cycle of operations some additional process by means of which the boundary values may be chosen.

Suppose, for the sake of argument, that the $(k-1)$st set of approximations has been derived, and that it has been used to obtain a new k'th approximation for the n'th derivative. The usual cycle of operations would require this n'th derivative to be repeatedly integrated, the results being accepted as the k'th set of approximations. But as the value of $y^{(n-1)}$ is not known for any value of x, it is impossible to determine in advance the constant of integration which should be used in performing the first integration—which means mechanically that the point from which the integraph pen should start is unknown. Under the circumstances, the most that can be done is to choose a constant at random in the hope that the mistake may be corrected later on. The integral $\bar{y}^{(n-1)}$ thus obtained, and the integral $y_k^{(n-1)}$ determined by the boundary values, can differ from each other only to the extent of an additive constant. Hence they must satisfy the relation

$$y_k^{(n-1)} = \bar{y}_k^{(n-1)} + C_k^{(n-1)}. \tag{10}$$

The derivative of next lower order—the $(n-2)$nd—should be obtained by integrating $y_k^{(n-1)}$, but this is impossible mechanically since $y_k^{(n-1)}$ is not known. However, it follows from (10) that

$$\begin{aligned} y_k^{(n-2)} &= C_k^{(n-2)} + \int y_k^{(n-1)} dx \\ &= C_k^{(n-2)} + C_k^{(n-1)} x + \int \bar{y}_k^{(n-1)} dx, \end{aligned}$$

in which the integral of $\bar{y}_k^{(n-1)}$, like $\bar{y}_k^{(n-1)}$ itself, is indefinite, and can be produced by the integraph. That is, no matter what constants of integration are assigned in the two integrations, the curve drawn by the integraph and the curve required by the boundary conditions can differ by a polynomial of the first degree, at most.

By continuing this argument, the relation

$$y_k = C_k + C_k'x + \ldots + C_k^{(n-1)} \frac{x^{n-1}}{(n-1!)} + \bar{y}_k \tag{11}$$

can ultimately be deduced, the $\bar{y}_k$ representing any n-fold integral that might be obtained with the integraph, and y_k the particular n-fold integral which would pass through the required points.

Call the values of $\bar{y}_k$ at the points $\xi_1, \xi_2, \ldots, \xi_n$, as read from the curve drawn by the integraph, $\bar{\eta}_1, \bar{\eta}_2, \ldots, \bar{\eta}_n$. The values of y_k at these same points are known to be $\eta_1, \eta_2, \ldots, \eta_n$. Substituting them in (11), pair by pair, the set of n linear equations

$$\begin{aligned} &C_k + C_k'\xi_1 + C_k'' \frac{\xi_1^2}{2!} + \ldots + C_k^{(n-1)} \frac{\xi_1^{n-1}}{(n-1)!} = \eta_1 - \bar{\eta}_1, \\ &C_k + C_k'\xi_2 + C_k'' \frac{\xi_2^2}{2!} + \ldots + C_k^{(n-1)} \frac{\xi_2^{n-1}}{(n-1)!} = \eta_2 - \bar{\eta}_2, \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ &C_k + C_k'\xi_n + C_k'' \frac{\xi_n^2}{2!} + \ldots + C_k^{(n-1)} \frac{\xi_n^{n-1}}{(n-1)!} = \eta_n - \bar{\eta}_n, \end{aligned} \tag{12}$$

is derived, from which the n unknown constants $C_k, C_k', \ldots, C_k^{(n-1)}$ can all be determined by the usual methods of solving linear equations.

Once $C_k^{(n-1)}$ is known, it can be added to $\bar{y}_k^{(n-1)}$, and it will then give $y_k^{(n-1)}$ in accordance with (10). Similarly, by adding $C_k^{(n-2)} + C_k^{(n-1)}x$ to $\bar{y}_k^{(n-2)}$ the function $y_k^{(n-2)}$ may be obtained. Proceeding in this manner, the entire set of k'th approximations may be deduced from the indefinite integrals given by the integraph.

With this extension the cycle of operations appears to be satisfactory, in the sense that a routine exists by means of which each set of approximations can be used to obtain the next set in order. It is still necessary, however, to overcome the difficulty, to which attention has already been directed, of choosing the first set in the sequence. This is most conveniently done by deriving the curve of $(n-1)$st degree which passes through the n given points, and accepting it, together with its successive derivatives, as the first set of approximations.

If the differential equation is of high order, this method of solution is too tedious to be practical. But most of the differential equations of applied mathematics are of the second order, and few indeed are higher than the fourth. For such equations the computations are not at all difficult.

Take, for example, Bessel's Equation (6). To solve it subject to the boundary values

$$y = 1, \quad \text{at } x = 1,$$
$$y = 0, \quad \text{at } x = 2,$$

a first approximation passing through these points must be obtained, and in accordance with the rule laid down above, this must be a curve of degree $n-1$. As $n=2$, this curve reduces to a straight line, the equation of which is obviously

$$y_1=2-x.$$

This y_1 together with its derivative

$$y_2'=-1,$$

constitutes the first set of approximations.

When they are substituted in (5), they lead to a second derivative

$$f_2'=2\left(\frac{1}{x^2}-1\right)+x,$$

which is plotted in Fig. 8. The integraph is then used to integrate this curve, giving $\bar{y}_2'$ and $\bar{y}_2$. In the first of these the constant of integration was so chosen as to make $\bar{y}_2'=-0.6$ when $x=1$; the sole reason for this choice being that the curve occupied a convenient place on the paper. In the second integration, however, the constant was so chosen as to satisfy one of the two boundary values.

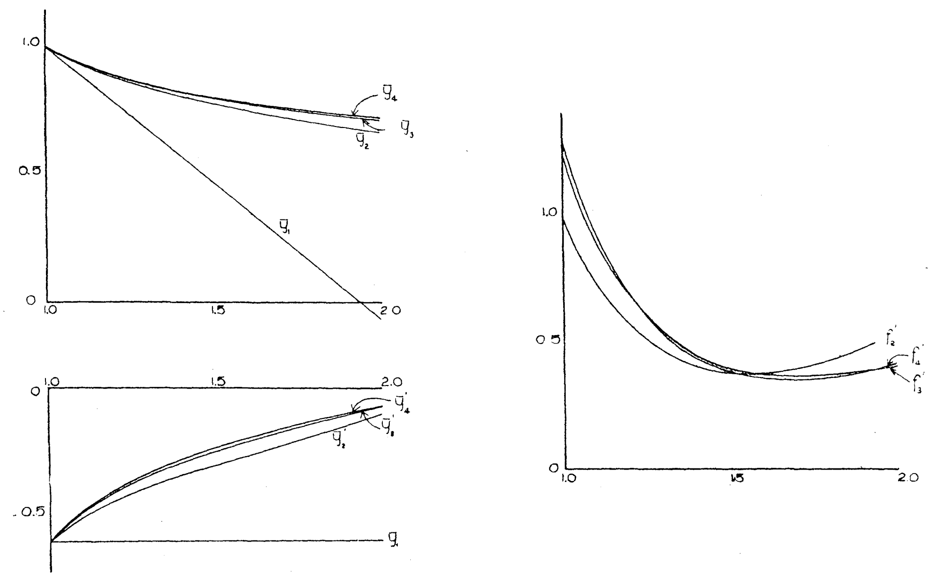


Fig. 8

The polynomial of the degree $n-1$ which converts $\bar{y}_2$ into y_2 being of the form

$$C+C'x,$$

and η_1 having been chosen as unity, the equations (12) reduce to the simple form

$$C+C'=0,$$
$$C+2C'=-\bar{\eta}_2.$$

The solution of these equations being

$$C' = -\bar{\eta}_2,$$

$$C = +\bar{\eta}_2,$$

it follows that

$$y_2 = \bar{y}_2 + (1-x)\bar{\eta}_2.$$

From the figure, the value of η_2 is read of as 0.681, whence

(13) $$y_2 = \bar{y}_2 + 0.681(1-x).$$

Similarly

(14) $$y_2' = \bar{y}_2' - 0.681.$$

Next, the values of $\bar{y}_2$ and $\bar{y}_2'$ corresponding to a suitable set of values of x must be read from the curves. They are then corrected by means of equations (13) and (14), so as to give y_2 and y_2'. These are then substituted back in (5), yielding f_3', from which in turn $\bar{y}_3'$ and $\bar{y}_3$ are obtained by integration, corrected to give y_3' and y_3, and so on.

As the curve is a simple one the process converges so rapidly that only four approximations are required. As a check upon the accuracy of the results obtained, they are recorded in Table V, together with the values computed from the known solution

$$y = 0.2652J_1(x) + 0.9101K_1(x).$$

The maximum error, on the scale of the original drawing, was about one two-hundredth of an inch, the high degree of accuracy being a direct result of the rapidity of convergence.

TABLE V

COMPARISON OF RESULTS OF THE GRAPHICAL AND THE ALGEBRAIC SOLUTION OF EXAMPLE VI

x	y(graphical)	y(calculated)	error
1.0	1.000	1.000	0.000
1.1	0.872	0.873	+0.001
1.2	0.755	0.756	+0.001
1.3	0.646	0.646	0.000
1.4	0.541	0.541	0.000
1.5	0.441	0.442	+0.001
1.6	0.346	0.346	0.000
1.7	0.254	0.254	0.000
1.8	0.166	0.166	0.000
1.9	0.081	0.081	0.000
2.0	0.000	0.000	0.000

Concluding Remarks

9. The general method of solution to which attention has been called above is virtually identical with Picard's method of approximation. So long as the boundary values are given in the form (2), the departures from the Picard process are of negligible consequence, and the usual proof of convergence, with appropriate minor changes, is still valid.

When the form of the boundary values is altered, however, somewhat more extensive changes are required in the Picard process, and the convergence of the sequence of functions is no longer assured. Examples V and VI show that the process sometimes converges, and if it does, it can only converge to the true solution. But it does not follow that it always converges. As a matter of fact, it is not difficult to adduce cases in which the opposite is true.

When such cases occur, it is still possible to fall back upon the expedient of using a cut-and-try process. For instance, if a differential equation were to be solved subject to the boundary values

$$y=0, \quad \text{at } x=0,$$
$$y'=1, \quad \text{at } x=3,$$

and if the process explained in §7 failed, the equation might be solved subject to several different sets of conditions such as

$$y=0, \quad y=0, \quad y=0,$$
$$y'=0, \quad y'=1, \quad y'=2,$$

all applied at $x=0$. From the final curve for each of these cases the value of y' at $x=3$ could then be read off, after which by interpolation, a new value of y' could be so chosen at $x=0$ as to cause the value of y' at $x=3$ to be very close to unity, as desired.

For instance, if when y' is given the values 0, 1 and 2 at $x=0$, it is found to have the values 0.6, 0.9, 1.1 at $x=3$, interpolation suggests that the value $y'=1.5$ at $x=0$ ought to lead to $y'=1.0$ at $x=3$. On trial, the latter value might be found to be 1.02, and then a further interpolation would be resorted to, giving the value 1.42 at $x=0$ as a plausible one to try. By patience, the solution would ultimately be obtained in this way. In general, however, the labour would be prohibitive, and it would be necessary to seek some other method of solution.

In spite of such failings, however, the Picard method, in combination with the integraph, has a wide field of usefulness in the solution of such problems as are met in industrial research. Not the least of its advantages lies in the fact that no extensive knowledge of mathematics is required to carry out the cycle of operations. This makes it possible to turn the actual work of solving the problems over to competent computers, such as are available in most industrial institutions, thus releasing the engineer for other activities.

In conclusion, some mention should be made of several other closely related discussions of graphical and mechanical methods for solving differential equations. The one which bears the closest resemblance to the present paper is contained in the book, *Graphical Methods*, by Carl Runge. Two methods are there presented, one founded upon Picard's approximation scheme, the other upon the method of Cauchy-Lipschitz. Both contemplate the use of purely graphical integration—which is not a satisfactory process in general—but can be equally well treated by means of the integraph. When so regarded, it is identical with the process explained in Example I.

Ernesto Pascal, in a booklet entitled *I miei Integrafi per Equazioni Differenziale*, deals with the problem of designing machines to integrate various types of equations, explaining in detail a large number of such machines. Where the work of a particular industry is such that equations of a certain type arise with great frequency, the possession of a machine especially designed for this type of equation may be advisable. In other establishments it is desirable to have one machine which is universally useful. No existing machine fully satisfies this requirement as to versatility, but the integraph approaches it most nearly.

In addition to these books, which contain a large proportion of original work, the excellent general reviews given by Maurice D'Ocagne in the second volume of his *Cours de Géométrie* and by Runge in the Enzyklopädie der Mathematischen Wissenschaften will be found helpful by those who are interested in the subject of graphical solution of differential equations*.

*Since the abovewas written (Aug. 7, 1924), Professor V. Bush of the Massachusetts Institute of Technology has developed a mechanical device of remarkable versatility in the solution of differential equations. See, for example, the advance press notices as quoted in The Literary Digest for Dec. 17, 1927.

LE PRINCIPE DE RÉCIPROCITÉ DANS LES DIVERSES BRANCHES DE LA PHYSIQUE

PAR M. UMBERTO PUPPINI,
Professeur à l'Université de Bologne, Bologne, Italie.

1. Le Principe de Réciprocité donne une représentation mathématique de divers phénomènes dans plusieurs branches des sciences appliquées. Il a, non-seulement l'avantage de donner des vues d'ensemble, mais encore de permettre d'obtenir des conséquences pratiques et spéciales auxquelles il serait sans doute difficile de parvenir par d'autres procédés.

Ce principe, appelé parfois aussi principe de Maxwell et de lord Rayleigh, a été l'objet de recherches théoriques et de vérifications expérimentales par plusieurs savants italiens.

Je me propose actuellement de donner un résumé très bref des énoncés obtenus en Italie dans diverses parties de la Physique. J'insisterai particulièrement sur l'Hydraulique, branche de la Science Technique à laquelle je consacre spécialement mon activité.

2. Le Principe de Réciprocité synthétise les relations de dépendance mutuelle entre déplacements et forces dans un système physique et donne une formule de comparaison entre deux états différents du système. On donne ici aux mots *déplacements* et *forces* un sens réel ou figuré selon le système physique dont il s'agit.

Pour que le principe soit valable, il faut que l'expérience permette d'admettre une loi linéaire et homogène entre les *déplacements* et les *forces* au voisinage d'un point quelconque du milieu en question.

3. Je commencerai par rappeler l'énoncé du Principe de Réciprocité dans la branche de la Physique appliquée où il a eu le plus grand nombre d'applications pratiques utiles: la théorie de l'Élasticité.

Dans le Nuovo Cimento de 1872, Enrico Betti a donné le théorème suivant: *si dans un corps solide élastique homogène, deux systèmes de déplacements font équilibre à deux systèmes de forces superficielles, la somme des produits des composantes des forces d'un système par les composantes des déplacements de l'autre, pris aux mêmes points reste la même si l'on intervertit les deux systèmes.*

En 1912, M. Gustavo Colonnetti a donné un deuxième énoncé du Principe de Réciprocité: *soit un milieu en équilibre sous l'action de forces données, la composante suivant une direction donnée (ou le moment par rapport à un axe donné) de la force intérieure qui se transmet au travers d'une section donnée, est égale au* *travail qu'accompliraient les forces si on obligeait les deux faces d'une section obtenue en coupant le milieu suivant la direction donnée à accomplir une translation relative unitaire et négative (ou à accomplir une rotation relative unitaire et négative autour de l'axe donné).*

M. Giuseppe Albenga, dans sa Note: *Sur le principe de réciprocité de Land* et M. Silvio Canevazzi, dans sa Note: *Sur les lignes d'influence dans la Science des Constructions*, ont montré que le théorème de Land et celui de M. Colonnetti peuvent être considérés comme de très importantes conséquences particulières du théorème de Betti.

Dans le domaine des applications de la Science des Constructions, il me suffira de rappeler que, en vertu du Principe de Réciprocité, on trace les lignes d'influence des déplacements, des forces, et des moments, comme si elles étaient des courbes de déformations élastiques correspondant à des sections données.

4. En Électrostatique, dans le Nuovo Cimento de 1872, le grand physicien de Bologne, M. Augusto Righi, donna, dans un de ses premiers travaux, des théorèmes analogues à un ancien théorème de Riemann, mentionné par Betti dans le Nuovo Cimento de 1864. Rappelons le suivant: *si l'on a plusieurs conducteurs isolés et non chargés et que l'on communique à un de ces conducteurs A une certaine quantité d'électricité, la valeur du potentiel sur un second conducteur B est la même que celle que l'on aurait sur A si l'on avait placé sur B la même quantité d'électricité.*

5. En Électrodynamique, M. Luigi Donati énonça un Principe de Réciprocité dans un Mémoire lu à l'Académie des Sciences de l'Institut de Bologne en 1899.

Supposons que des courants électriques continus parcourent un réseau de fils conducteurs où se trouvent des nœuds par lesquels le réseau reçoit (ou émet) des courants. Imaginons que ce réseau soit parcouru par deux systèmes différents de courants d'intensités I et I', dus à des forces électromotrices E et E' dans les fils et à des potentiels H et H' aux nœuds par lesquels le réseau reçoit (ou émet) des courants d'intensités I_e et I'_e. M. Donati a établi la relation suivante entre ces grandeurs:

$$\Sigma IE' + \Sigma_1 I_e H' = \Sigma I'E + \Sigma_1 I'_e H, \tag{1}$$

où les sommes Σ sont étendues aux fils et les sommes Σ_1 aux nœuds par lesquels le réseau reçoit ou émet de l'énergie.

La formule (1) se déduit de la combinaison des $n-1$ équations de Kirchhoff relatives aux n nœuds et des r équations d'Ohm relatives aux r fils:

$$\Sigma I + \Sigma I_e = 0, \qquad I_{ab} = \frac{H_a - H_b + E_{ab}}{R_{ab}},$$

où R_{ab} est la résistance ohmique d'un fil (a, b).

Dans deux travaux publiés en 1900 et en 1910, M. Donati a déduit du principe des relations particulières sur la distribution du potentiel dans un réseau de fils conducteurs, et dans son Mémoire de 1917: *Les courants alternatifs* *et la loi de réciprocité*, il a donné une formule analogue à la précédente pour le cas des courants alternatifs (supposés sinusoïdaux, de même fréquence et sans inductance mutuelle entre les fils). Mais, dans cette nouvelle relation, les E, I, H sont des nombres complexes correspondant aux vecteurs représentatifs des grandeurs alternatives, et au lieu des résistances ohmiques, on introduit les nombres complexes représentant les impédances des fils: ainsi, les sommes qui figurent dans la formule sont des sommes de produits de quantités complexes.

A la différence des sommes de la formule (1), ces nouvelles sommes n'ont pas d'interprétation énergétique. Cependant, M. Donati a encore énoncé un autre Principe de Réciprocité pour des courants alternatifs sinusoïdaux et les sommes qu'il introduit ont une interprétation énergétique: en désignant par J la demi-somme du courant correspondant à des valeurs données de E, H et du courant dans le même fil et correspondant aux mêmes valeurs de E, H, mais en supposant changé le signe des réactances de tous les fils du réseau, M. Donati établit la formule:

$$\Sigma EJ' + \Sigma_1 HJ_e' = \Sigma E'J + \Sigma_1 H'J_e \tag{2}$$

où les produits sont scalaires.

On peut donc dire que, pour les courants alternatifs, la loi de Réciprocité se subdivise en deux lois: d'abord une loi introduisant des quantités complexes et analogue à la formule (1), sans interprétation énergique, mais à conséquences cinématiques, et ensuite une loi avec des scalaires d'où l'on peut déduire des conséquences énergétiques.

6. C'est encore dans le domaine de l'Électrodynamique que, tout récemment, en 1923, M. Vittorio Gori a énoncé dans le Bulletin de l'Union Mathématique Italienne le théorème suivant: *étant donnés plusieurs circuits électriques immobiles pouvant subir une action inductive causée par deux systèmes différents de variations de flux d'induction magnétique, à un instant quelconque, la somme des produits des intensités des courants induits dans un système par les variations du flux à travers les circuits de l'autre reste la même si l'on permute les deux systèmes.*

7. Enfin, dans le domaine de l'Électrodynamique encore, j'ai démontré, en 1922, le théorème suivant: *soit une solution électrolytique (ou, plus généralement, un corps conducteur), et supposons cette solution limitée par des surfaces cohibantes d'une part, et par des surfaces conductrices d'autre part; considérons deux systèmes de valeurs pour les potentiels des surfaces conductrices et les systèmes correspondants pour les intensités des courants qui traversent ces surfaces; la somme des produits des intensités des courants dans un système, par les potentiels de l'autre est égale à la somme des produits des intensités des courants dans le second système, par les potentiels du premier.*

De ce théorème résulte le corollaire suivant: *si un courant d'intensité unité, sortant en un point d'un liquide, crée une diminution de potentiel en un second point de ce liquide, inversement, un courant d'intensité unité, sortant par le second point, crée au premier point la même diminution de potentiel.*

Lorsque je parlerai du Principe de Réciprocité en Hydraulique, je reviendrai sur ce corollaire; j'ajouterai seulement que, en collaboration avec M. Gori, nous avons fait une vérification expérimentale de ce théorème en utilisant comme électrolyte une solution de sulfate de cuivre.

8. En Thermodynamique, j'ai énoncé, en 1916, le principe suivant: *étant donnée une masse solide isotrope et deux régimes différents de flux de chaleur entre cette masse et l'extérieur, la somme des produits des flux de chaleur échangés dans le premier régime par les températures correspondants du second est égale à la somme des produits des flux du second régime par les températures du premier.*

C'est là un principe général, qui ne dépend ni de la forme du milieu isotrope, ni de son étendue, ni de ses relations avec l'extérieur, ni de sa structure, homogène ou non. L'énoncé du principe suppose simplement la loi de proportionnalité entre le flux de Chaleur et la variation de température qui, à son tour, implique seulement que le milieu est isotrope et que le coefficient de conductibilité thermique est une constante indepéndante de la température.

De cette loi générale, on peut déduire des relations particulières. Par exemple, si l'on suppose que la masse est *athermane*, c'est-à-dire, imperméable à la chaleur, qu'elle ne contient aucune source de chaleur, qu'il n'y filtre aucun fluide, chaud ou froid, que les surfaces de contact avec l'extérieur sont composées d'un nombre limité de parties, chacune d'elles à une température constante, et si, de plus, enfin, les régimes sont permanents, on obtient la relation très simple:

$$\Sigma \Phi T' = \Sigma \Phi' T, \tag{3}$$

où Φ, Φ' sont les flux dans deux régimes, T, T' les températures sur les surfaces où passent les flux de chaleur. Il est entendu que les sommes doivent comprendre toutes les surfaces d'entrée et de sortie des flux de chaleur.

J'ai fait une application immédiate de ce Principe de Réciprocité en Thermodynamique à la question de l'échauffement de l'eau dans les conduites souterraines.

Rappelons que si l'on a une conduite métallique souterraine de rayon extérieur r, dont l'axe est à une profondeur h, et dans laquelle circule de l'eau à la temperature T, le coefficient de conductibilité du terrain supposé homogène étant K, la quantité de chaleur reçue dans l'unité de temps par l'unité de longueur de la conduite est:

$$\phi = \frac{2\pi K(T_0 - T)}{\log\left(\frac{h}{r} + \sqrt{\frac{h^2}{r^2} - 1}\right)} \; .$$

M. Forchheimer a proposé de prendre pour T_0 la température qu'aurait le terrain à la profondeur h, si la conduite n'existait pas, règle qui serait rigoureuse si la température du terrain était uniforme et constante, mais qui ne semble pas satisfaisante car la température du terrain varie périodiquement, l'amplitude de la variation décroissant, et le retard de phase croissant avec la profondeur. Cependant, l'application à ce cas du Principe de Réciprocité conduit à reconnaître que la règle de M. Forchheimer donne des résultats différant des résultats rigoureux de quantités négligeables parce que d'un ordre de grandeur très petit.

Considérant ensuite le cas de deux conduites parallèles et voisines, j'ai montré que le Principe de Réciprocité ramène l'étude de l'échauffement de l'eau dans ces conduites à l'étude plus simple et déjà connue de l'échauffement de l'eau dans le cas d'une seule conduite.

9. En Hydraulique, le Principe de Réciprocité a fait l'objet de recherches récentes.

J'ai énoncé le Principe en 1911, relativement à un réseau de tuyaux dans lequel le mouvement du liquide suit la loi de Poiseuille, c'est-à-dire la loi de proportionnalité entre le débit et la pente motrice.

Soient un tuyau ab, de longueur L, H_a et H_b les hauteurs piézométriques aux extrémités, Q_{ab} le débit, R_{ab} la résistance hydraulique, et supposons, pour plus de généralité, que le tuyau soit le siège de forces motrices dont E_{ab} est le travail par unité de temps pour le passage d'un débit unité. Si ϵ est le coefficient de viscosité du fluide, σ l'aire de la section du tuyau, ϖ le poids spécifique, C un coefficient numérique qui est fonction de la forme de la section, on a pour la résistance hydraulique:

$$R_{ab} = \frac{\epsilon L}{(\varpi\sigma^2)}$$

et, pour l'équation du régime permanent dans le tuyau:

$$Q_{ab} = \frac{H_a - H_b + E_{ab}}{R_{ab}} . \tag{4}$$

Si l'on a r tuyaux, on obtient r équations analogues. Et si l'on a n nœuds, on a $n-1$ équations indépendantes de la forme:

$$\Sigma Q = 0 \tag{5}$$

la somme étant étendue à tous les tuyaux aboutissant à un nœud.

Des équations (4) et (5), il résulte que: *si l'on imagine, dans un réseau de tuyaux avec mouvement à la Poiseuille, deux modes différents de circulation d'un liquide, les débits étant Q et Q', les forces motrices intérieures: E et E', les hauteurs piézométriques: H et H', les débits: Q_e et Q'_e aux nœuds où il y a entrée ou sortie d'eau; on obtient alors la relation:*

$$\Sigma E'Q + \Sigma_1 H'Q_e = \Sigma EQ' + \Sigma_1 HQ'_e, \tag{6}$$

les sommes Σ étant relatives à tous les tuyaux et les sommes Σ_1, à tous les nœuds où entre ou sort le liquide.

Multiplions, en effet, l'équation (4) par Q'_{ab}, ajoutons toutes les équations analogues, tenons compte de (5), nous obtenons:

$$\Sigma EQ' + \Sigma_1 HQ'_e = \Sigma RQQ'; \tag{7}$$

et, par analogie:

$$\Sigma E'Q + \Sigma_1 H'Q_e = \Sigma RQQ'. \tag{7'}$$

La comparaison des équations (7) et (7') justifie l'équation (6).

10. La méthode même qui conduit à la formule de réciprocité met bien en évidence le fait que la formule ne subsiste pas lorsque le mouvement du liquide dans les tuyaux est turbulent au lieu d'être un mouvement régulier à la Poiseuille.

On a, en effet, alors, pour des tuyaux de section circulaire de diamètre D_{ab}, K_{ab} étant le coefficient de résistance, $|Q_{ab}|$ étant la valeur absolue du débit Q_{ab}:

$$Q_{ab} = \frac{H_a - H_b + E_{ab}}{\frac{K_{ab} L_{ab}}{D^5_{ab}} |Q^{ab}|} . \tag{8}$$

Aux formules (7) et (7'), il faut alors substituer les suivantes:

$$\Sigma E'Q + \Sigma_1 H'Q_e = \Sigma \frac{KL}{D^5} |Q| QQ',$$

$$\Sigma EQ' + \Sigma_1 HQ_e' = \Sigma \frac{KL}{D^5} |Q'| QQ',$$

desquelles on ne peut visiblement pas déduire l'équation (6).

11. Ainsi, la validité du Principe de Réciprocité implique l'existence d'une relation linéaire entre les débits et les pentes motrices. Une pareille loi, la loi de Darcy-Ritter, pouvant être admise approximativement pour les milieux perméables tels que graviers, sables, on peut prévoir un Principe de Réciprocité pour les milieux perméables à l'eau, sous réserve, bien entendu, que les deux membres de l'équation exprimant le Principe de Réciprocité aient les mêmes champs de sommation, ou d'intégration. Cette dernière remarque fait comprendre qu'il sera plus facile d'énoncer le Principe de Réciprocité pour les nappes filtrantes à volume invariable: les nappes en pression, ou *artésiennes*, que pour les nappes à volume variable: les nappes à surfaces libres ou *phréatiques*. Pour ces dernières, il faudra des artifices ou des limitations particulières.

12. Commençons par les nappes artésiennes.

Considérons un milieu poreux, de forme et de grandeur quelconques, limité en partie par des surfaces imperméables, en partie par des surfaces perméables que nous supposons en communication avec un réservoir à niveau libre qui fournit ou retire de l'eau au milieu poreux.

Supposons les surfaces libres de l'eau des réservoirs assez hautes pour que le milieu perméable soit toujours sous pression: une variation des niveaux des surfaces libres créera une variation des pressions et des débits d'alimentation et d'émission, mais ne changera pas le volume occupé par le liquide en mouvement. Cette dernière circonstance se produirait, au contraire, si une partie seulement, et non la totalité du terrain perméable était occupée par le liquide mobile, ce qui serait le cas si l'eau, au lieu d'être sous pression dans le milieu poreux, y avait une surface libre.

Soit μ le coefficient de filtration du milieu, ζ la hauteur piézométrique en un point $M(x, y, z)$, quelconque, de la nappe fluide. En M, on peut envisager deux vecteurs: l'un f, que nous appellerons la *pente motrice*, de composantes suivant les axes: $\frac{\partial \zeta}{\partial x}$, $-\frac{\partial \zeta}{\partial y}$, $-\frac{\partial \zeta}{\partial z}$ et l'autre, q: le *débit*, de composants $-\mu\frac{\partial \zeta}{\partial x}$, $-\mu\frac{\partial \zeta}{\partial y}$, $-\mu\frac{\partial \zeta}{\partial z}$. Le premier vecteur est lamellaire pur et le second solénoïdal.

Envisageons deux systèmes différents de niveaux pour les surfaces par lesquelles la nappe est en contact avec l'extérieur. Il leur correspond des conditions de mouvements définies respectivement par la fonction potentielle ζ, les vecteurs f et q, et par le potentiel ζ', les vecteurs f' et q'. Désignons par $f\times f'$ le produit scalaire de f par f', et par τ le volume occupé par la masse poreuse et évaluons l'intégrale $\int_\tau \mu f\times f'd\tau$. Divisons le domaine τ en tubes élémentaires du flux du vecteur q, de sorte que:

$$d\tau = dsd\sigma,$$

où ds est un élément d'arc de l'axe du tube, et $d\sigma$ l'aire de la section normale du tube. On en déduit:

$$\int_\tau \mu f\times f'd\tau = \int_\sigma\int_s \mu f d\sigma . f'\times ds.$$

Dans le second membre, f est un scalaire et la qualité vectorielle est reportée sur l'élément ds. Or, la quantité $\mu f d\sigma = q d\sigma$ est constante le long du tube de flux du vecteur q, et, d'autre part, l'intégrale $\int_s f'\times ds$, évaluée pour une ligne de flux entière est égale à la différence $H'-H'_1$, des hauteurs piézométriques aux deux extrémités de cette ligne. Donc, on peut écrire:

$$\text{(9)} \qquad \int_\tau \mu f\times f'd\tau = \int_\sigma q(H'-H'_1)d\sigma,$$

le champ d'intégration σ étant étendu à toutes les surfaces de contact de la nappe artésienne avec l'extérieur, soit les surfaces par lesquelles l'eau est reçue, soit les surfaces par lesquelles elle est émise.

Si, par contre on étendait le champ à ces deux catégories de surfaces simultanément, on aurait une expression plus simple:

$$\text{(10)} \qquad \int_\tau \mu f\times f'd\tau = \int_\sigma qH'd\sigma,$$

mais, ici, le domaine d'intégration σ est étendu à toutes les surfaces de communication de la nappe avec l'extérieur.

En second lieu, subdivisons le volume τ en tubes correspondant au vecteur q' et répétons le même calcul. On obtient:

$$\text{(11)} \qquad \int_\tau \mu f\times f'd\tau = \int_\sigma q'Hd\sigma;$$

et, en comparant (10) et (11):

$$\text{(12)} \qquad \int_\sigma qH'd\sigma = \int_\sigma q'Hd\sigma$$

C'est là l'expression analytique du Principe de Réciprocité pour les nappes artésiennes.

Si, en particulier, tous les points d'une même surface d'alimentation ou d'émission ont la même hauteur piézometrique, les deux intégrales de la formule (12) se changent en sommes d'un nombre fini de termes

$$\Sigma QH' = \Sigma Q'H. \tag{13}$$

Dans ces sommes les signes Σ doivent comprendre toutes les surfaces d'alimentation et d'émission, Q et H étant les débits et les hauteurs piézométriques pour chacune de ces surfaces dans un des systèmes de mouvements considérés, Q', H' les débits et les hauteurs dans l'autre système.

La formule (13) peut être énoncée de la manière suivante: dans une nappe artésienne, si l'on considère deux systèmes différents de hauteurs piézométriques et de débits d'entrée et de sortie, la somme des produits des débits d'un système par les hauteurs de l'autre est égale à la somme des produits des débits du deuxième système par les hauteurs du premier.

J'ai donné au theorème de réciprocité une forme plus commode pour la pratique. En considérant un ensemble de puits très éloigné de tout autre puits de la nappe, j'ai obtenu la formule:

$$\Sigma Qh' = \Sigma Q'h, \tag{14}$$

dans laquelle les sommes doivent comprendre tous les puits, et eux seuls. Les débits dans les deux systèmes sont Q et Q', et h et h' sont les dénivellations piézométriques dans les puits, rapportées aux hauteurs piézometriques que l'on aurait dans chaque puits si son débit était nul.

13. Prenons maintenant le cas des nappes phréatiques.

La nécessité où l'on est d'envisager un champ d'intégration invariable m'a amené à me restreindre à la considération de nappes phréatiques de faibles hauteurs relativement à leurs dimensions superficielles, à surfaces libres à pente faible, et, enfin, à surfaces latérales composées de cylindres à génératrices verticales. Avec ces restrictions, on peut prendre, dans les différents cas, le même champ d'intégration, à savoir, le champ à deux dimensions formé par la projection horizontale de la nappe.

On obtient alors une formule de réciprocité rigoureuse bien qu'un peu compliquée. Elle se simplifie, toutefois, dans certains cas particuliers. Si, par exemple, on compare deux régimes permanents sans apports superficiels de pluie, et si on appelle H et H' les hauteurs piézométriques rapportées au fonds imperméable, cela sur chacune des surfaces latérales de constant avec l'extérieur, on obtient:

$$\Sigma QH'^2 = \Sigma Q'H^2. \tag{15}$$

Et si l'on envisage un autre régime avec apports superficiels de pluie, mais sans alimentation latérale, on obtient:

$$\Sigma Q'h^2 + p\left\{\sigma h_2'^2 - \int_\sigma h'^2 d\sigma\right\} = 0, \tag{16}$$

où p est l'apport superficiel de pluie par unité de surface libre, σ la projection horizontale de la nappe et h'_2 la hauteur piézométrique d'émission lorsque les débits ont les valeurs Q'.

Mais, tout récemment, en 1923, M. Marcello Lelli a donné, pour les nappes phréatiques, un énoncé du Principe de Réciprocité qui n'exige de limitations, ni de forme, ni de fonctionnement et qui est très analogue à l'énoncé relatif aux nappes artésiennes. Mais il faut prendre comme domaine d'intégration pour les intégrales de la formule (9), le volume du domaine filtrant commun aux deux régimes de mouvement envisagés. Il faut, de plus, entendre que les surfaces d'intégration comprennent, non pas seulement les surfaces σ de contact avec les réservoirs extérieurs communs aux deux régimes, mais encore les surfaces w qui séparent le domaine filtrant, commun aux deux régimes, des portions du domaine filtrant qui ne sont pas communes aux deux régimes. Plus précisément, on a:

$$\int_\sigma qH'd\sigma+\int_\omega qH'd\omega=\int_\sigma q'Hd\sigma+\int_\omega q'Hd\omega. \tag{17}$$

14. En ce qui concerne les nappes phréatiques, on a déjà déduit du Principe de Réciprocité des conséquences intéressantes pour les applications pratiques.

De la formule (16), j'ai déduit la connaissance de l'état de régime permanent pour un système filtrant de forme prismatique à génératrices horizontales de longueur indéfinie, reposant sur un fonds horizontal imperméable, avec constant apport de pluie. A cet effet, je compare ce régime au régime des filtres à fonds horizontal avec des surfaces libres cylindriques à génératrices horizontales, régime dont le débit est exprimé par une formule très simple et bien comme de Dupuit.

Et, en utilisant la formule (17), M. Lelli a montré que la formule de Dupuit et celle du débit des filtres circulaires à surface libre, qui semblent être des formules d'approximation puisqu'elles sont fondées sur une hypothèse non rigoureuse concernant la forme des surfaces de hauteur piézométrique constante, sont au contraire des formules parfaitement rigoureuses.

15. Mais c'est pour les nappes artésiennes, à cause de la plus grande simplicité de l'énoncé, que les déductions et les applications pratiques sont le plus facile.

Je me rapporterai premièrement à la formule (14) et à la considération de deux puits seulement, qui s'influencent mutuellement, mais qui sont très éloignés de tout autre puits. J'appelerai ces puits: 1 et 2.

Considérons le régime consistant uniquement en un débit unité du puits 1, puis un second régime consistant uniquement en un débit unité du puits 2.

Nous aurons:

	Puits 1		Puits 2	
	débit	dépression	débit	dépression
1^er^ régime	1	h_1	0	h_{2-1}
2^e^ "	0	h_{1-2}	1	h_2

En appliquant la formule (14) à la comparaison de ces deux régimes, l'on a:

$$h_{1-2}=h_{2-1},$$

c'est-à-dire: la dépression qu'un débit égal à l'unité sortant d'un puits exerce sur un autre est égale à la dépression qu'un débit égal à l'unité sortant du second puits exerce sur le premier.

Ce théorème correspond pour la forme au théorème classique de Maxwell pour les corps élastiques.

16. Je me propose maintenant de déduire la connaissance d'un régime tout-à-fait arbitraire de débits simultanés de deux puits de la connaissance d'un régime quelconque de fonctionnement du seul puits 1, et d'un régime quelconque de fonctionnement du seul puits 2. Nous pourrons donc, par exemple, résoudre le problème: trouver les débits de deux puits lorsqu'on connait les dépressions.

Nous avons

	Puits 1		Puits 2	
	débit	dépression	débit	dépression
1^er^ régime connu	q_1	h_1	0	h_{2-1}
2^e^ " connu	0	h_{1-2}	q_2	h_2
3^e^ " inconnu	x	$\overline{h}_1$	y	$\overline{h}_2$

De la comparaison du premier régime avec le troisième et du deuxième avec le troisième, on tire:

$$\begin{cases} q_1\overline{h}_1=xh_1+yh_{2-1}, \\ q_2\overline{h}_2=xh_{1-2}+yh_2. \end{cases}$$

On voit donc que, si l'on connait les régimes 1 et 2, on trouve les débits x, y simultanés qui correspondent aux valeurs des dépressions $\overline{h}_1$, $\overline{h}_2$.

17. Par un procédé tout à fait semblable, on obtient un théorème remarquable que l'on peut appeler le *théorème de la superposition des effets*, c'est-à-dire: *la dépression que les débits* $Q_1, Q_2, \ldots, Q_n$ *de* n *puits exercent sur un point* M *de la nappe est égale à la somme des produits des débits* $Q_1, Q_2, \ldots, Q_n$ *par les dépressions qu'un débit unité d'un puits placé en* M *exercerait sur les puits* $1, 2, \ldots, n$.

Ce théorème ramène la construction par points de la surface piézométrique sous l'action de n puits au problème beaucoup plus simple de la connaissance de la surface piézometrique sous l'action d'un seul puits.

18. J'ai eu l'occasion de faire une application immédiate du principe, lorsque je fus chargé de prévoir les débits de six puits artésiens, éloignés l'un de l'autre de cent mètres à peu près en supposant que les six puits devaient fonctionner avec des dépressions piézometriques déterminées.

Le moyens que j'avais à ma disposition pour les expériences me permettaient de pomper de l'eau dans un seul puits à la fois. Je fis six expériences l'une après l'autre en faisant fonctionner un seul puits à la fois; et j'enregistrai le débit et la dépression du puits fonctionnant et les dépressions dans les cinq autres puits. La comparaison, au moyen du Principe de Réciprocité, du régime inconnu de fonctionnement simultané des six puits dont les dépressions sont données, avec les six régimes révélés par l'expérience donna lieu à six équations du premier degré entre les six débits inconnus. En résolvant le système de ces six équations, je résolus le problème que l'on m'avait proposé.

19. Une vérification du Principe de Réciprocité, donnée par l'expérience, pour les nappes artésiennes peut être trouvée dans les recherches qui furent exécutées par une Commission d'Ingénieurs sur des puits artésiens de la Venaria pour l'Aqueduc de la ville du Turin, plusieurs années avant que le Principe ne fût énoncé. Il s'agissait de trois puits artésiens que l'on faisait fonctionner, soit isolément, soit par couples, soit tous les trois simultanément.

Les expériences furent exécutées en 1903 et donnèrent les résultats suivants:

	Puits	Débits	Dépressions
Expériences		litres par seconde	mètres
	I	. . . . .	4.17
A	II	13.60	11.27
	III	. . . . .	2.29
	I	. . . . .	3.28
B	II	. . . . .	2.03
	III	10.80	10.87
	I	. . . . .	6.28
C	II	12.20	11.43
	III	8.55	11.10
	I	7.28	10.94
D	II	9.45	11.73
	III	7.00	11.23

En calculant les sommes de la formule (14), on trouve que le Principe est vérifié avec une bonne approximation. Si l'on excepte la comparaison du régime A avec le régime B, dans laquelle on trouve une erreur du 12%, dans les comparaisons entre les autres régimes où les données expérimentales sont plus nombreuses, on trouve des erreurs très petites.

Comparaison de A avec C	erreur 4.00%
" " A " D	" 1.50%
" " B " C	" 1.82%
" " B " D	" 1.82%
" " C " D	" 3.37%

On voit donc que le principe est bien vérifié par ces données expérimentales.

20. Mais on peut se procurer d'autres vérifications expérimentales indirectes, en construisant des modèles électrolytiques des champs de filtrations. Qu'il me soit permis de rappeler la vérification dont j'ai déjà parlé, que j'ai effectuée en expérimentant sur un électrolyte, avec la collaboration de M. V. Gori. Elle a permis de contrôler l'exactitude de la loi identique, pour la forme, à la loi de Maxwell pour les corps élastiques.

A cet égard, je désire noter que des modèles électrolytiques peuvent servir, non-seulement à vérifier expérimentalement des lois déjà connues de filtration, mais aussi à découvrir des lois qui peuvent être reportés dans un champ de filtration.

Quelquefois, la recherche de la loi qui exprime le débit d'un système perméable en fonction des dimensions et de la forme du système et des coefficients de filtration peut exiger des opérations trop difficiles au point de vue analytique et au point de vue expérimental.

Dans ce cas, on peut construire un petit modèle de forme géométriquement semblable à la forme du système perméable et dans lequel au milieu poreux est substitué un milieu conducteur, par exemple un électrolyte. Alors, la relation que l'on tire en conséquence de mesures expérimentales sur l'intensité du courant, les dimensions et la forme du modèle, peut être transposée au cas de la filtration. Naturellement, il faut tenir compte du rapport de similitude.

Des modèles électrolytiques furent proposés autrefois pour l'étude des condensateurs, en s'appuyant dans ce cas sur l'identité formelle des lois des courants continus en milieu conducteur et des lois du champ de force dans un diélectrique. De pareils modèles furent notamment utilisés pour la détermination de la capacité du système terre-antenne, en télégraphie sans fil.

D'une manière analogue, des modèles électrolytiques doivent pouvoir être utilisés pour l'étude de cas plus ou moins compliqués de mouvements de liquides filtrants, en confirmant une fois de plus ce fait que Principe de Réciprocité illustre particulièrement, et qui est que des phénomènes appartenant à des domaines bien différents de la Physique peuvent souvent être représentés par une formule mathématique unique.

HYPERBOLIC-FUNCTION SERIES OF INTEGRAL NUMBERS AND THE OCCASIONS FOR THEIR OCCURRENCE IN ELECTRICAL ENGINEERING

BY PROFESSOR A. E. KENNELLY,
Harvard University, Cambridge, Massachusetts U.S.A.

Hyperbolic-Cosine Series.

Consider the infinite series of complex numbers:

(1*a*) $A\cosh(-m\theta+\alpha)\,.\,.A\cosh(-\theta+\alpha),\ A\cosh\alpha,\ A\cosh(\theta+\alpha)\,.\,.A\cosh(n\theta+\alpha)\ldots$

where θ and α are constant complex hyperbolic angles of the type

(1*b*) $$\theta=\theta_1+i\theta_2.$$

A is also a constant complex number for the series, positive, negative or zero, integral or fractional. The numbers $-m$ and n are real integers, which may be indefinitely large. We may call any number belonging to this series, such as $A\cosh(n\theta+\alpha)$ a *hyperbolic-cosine number*, and the series a *general hyperbolic-cosine series.*

Mean-to-Mid Constant Ratio of a Cosine Series.

Let any three successive terms of a hyperbolic cosine series be called a *triad*, such as

(1) $A\cosh\{(n-1)\theta+\alpha\}$ denoted by I

(2) $A\cosh\{n\theta+\alpha\}$ " " II

(3) $A\cosh\{(n+1)\theta+\alpha\}$ " " III

Terms I and III of a triad may be called the *outers*, and term II the middle term or *mid*. The triad may be rewritten as:

(4) $\mathrm{I}=A\cosh\{(n\theta+\alpha)-\theta\}=A\cosh(n\theta+\alpha)\cosh\theta-A\sinh(n\theta+\alpha)\sinh\alpha,$

(5) $\mathrm{II}=A\cosh\{n\theta+\alpha\},$

(6) $\mathrm{III}=A\cosh\{(n\theta+\alpha)+\theta\}=A\cosh(n\theta+\alpha)\cosh\theta+A\sinh(n\theta+\alpha)\sinh\alpha.$

Adding together the outers,

(7) $$\mathrm{I}+\mathrm{III}=2A\cosh(n\theta+a)\,.\cosh\theta=\mathrm{II}.2\ \cosh\theta.$$

Consequently, the sum of the outers of any triad is always equal to the mid multiplied by $2\cosh\theta$, no matter what constant complex values may attach to A and α. This quantity

(8) $$2\cosh\theta=\frac{\mathrm{I}+\mathrm{III}}{\mathrm{II}}=k,$$

the constant ratio of the outers' sum to the mid of any triad in the series, may be called the *characteristic ratio k* of the series. Expressing the same property of a general hyperbolic-cosine series in another way, the half sum of the outers $\frac{\mathrm{I}+\mathrm{III}}{2}$ is their complex arithmetical mean, and the mean-to-mid ratio of any triad is

$$\frac{\mathrm{I}+\mathrm{III}}{2\,\mathrm{II}}=\cosh\theta, \tag{9}$$

a constant complex numerical* ratio. The mean-to-mid ratio of any trio in the series is therefore half the characteristic ratio of the series.

Integral Hyperbolic-Cosine Series.

We are interested here only in hyperbolic-cosine series of *real integral* numbers, as distinguished from general series of complex numbers. It is evident from (7), that in any triad, III will always be an integer (positive, negative or zero), if I and II are integers, and if the characteristic ratio $2\cosh\theta$ is also an integer. If $2\cosh\theta$ is an integer, it will be seen that each of the following expressions are likewise integers.

$$4\sinh^2\theta,\ 4\cosh^2\theta,\ 4\sinh^2\left(\frac{\theta}{2}\right),\ 4\cosh^2\left(\frac{\theta}{2}\right). \tag{10}$$

It will be evident that there are an infinite number of hyperbolic angles θ which satisfy the relation

$$2\cosh\theta=\text{integer characteristic ratio } k \tag{11}$$

The following table gives a list of the values of θ corresponding to characteristic ratios between -6 and $+6$.

TABLE I

Brief List of angles θ corresponding to integral characteristic ratios from -6 to $+6$

k	$\cosh\theta$	θ	k	$\cosh\theta$	θ
-6	-3	$1.7627+i2\ (\pi/2)$	6	3	1.7627
-5	-2.5	$1.5668+i2\ (\pi/2)$	5	2.5	1.5668
-4	-2	$1.3170+i2\ (\pi/2)$	4	2	1.3170
-3	-1.5	$0.9624+i2\ (\pi/2)$	3	1.5	0.9624
-2	-1	$0\quad +i2\ (\pi/2)$	2	1	0
-1	-0.5	$0\quad -i1.3\dot{3}\ (\pi/2)$	1	0.5	$0+i0.6\dot{6}\ (\pi/2)$
-0	-0	$0\quad -i1\ (\pi/2)$	0	0	$0+i1\ (\pi/2)$

Central Triad and Central Term of a Series.

If in the expressions (1), (2) and (3), we put $n=0$, we obtain the *central triad* of the series; namely:

$$A\cosh(-\theta+\alpha)=A\cosh(\theta-\alpha),\quad N \tag{12}$$

$$A\cosh\alpha,\quad O \tag{13}$$

$$A\cosh(\theta+\alpha).\quad P \tag{14}$$

**The Constant Ratio of Mean-to-Mid Potential or Current at Section-Junctions of a Uniform Electric Conducting Line, Real or Artificial, in the Steady State,* by A. E. Kennelly, Proc. Nat. Acad. of Sciences, June 1923.

We may designate the terms of this central triad by the letters NOP. The middle term will then be O.

In order that all the terms of the series shall be integers, it is necessary and sufficient that in addition to having an integral characteristic ratio, the central term O shall be an integer and one of the outers, say N, shall also be an integer. In that case, by (8), P is an integer, and all the terms of the series must follow the same rule. Hence the following conditions must be satisfied:

(15) $O = A\cosh\alpha$ = an integer,

(16) $k = 2\cosh\theta$ = " " ,

(17) $2A\sinh\theta\sinh\alpha$ = " " ,

(18) $2A\cosh\theta\cosh\alpha$ = " " .

In order, therefore, to form an integral cosine series, we select a central integer O, and a characteristic ratio $k = 2\cosh\theta$. Then

(19) $N + P = kO$ an integer.

We now select any integer for N, having the same sign as O, and numerically lying between O/k and $4O - O/k$ inclusive,* and the integer P becomes determined. This fixes the central triad, and all the other terms become definite.

Thus, taking $O = 5$, and $k = 4$, for which, by Table I, $\theta = 1.317$, we have $N + P = 20$. We may now take for N any integer from 1.25 to $20 - 1.25$ inclusive, to form a corresponding series. Suppose $N = 8$ then $P = 12$. This gives

	N	O	P
(20)	8	5	12.

We may now form Q by treating P as the mid of the triad OPQ, so that $O + Q = kP = 48$, whence $Q = 43$. Again, taking N as the mid of a triad MNO, we have $O + M = kN = 32$, whence $M = 27$. We now have:

	M	N	O	P	Q
(21)	27	8	5	12	43.

Proceeding in this way, we obtain the following indefinitely long series of hyperbolic cosine numbers as far as

	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V
(22)	19388	5195	1392	373	100	27	8	5	12	43	160	597	2228	8315	31032.

To resolve this series into its hyperbolic-cosine terminology, we may notice that from (14) and (13)

$$\frac{P}{O} = \frac{A(\cosh\theta\cosh\alpha + \sinh\theta\sinh\alpha)}{A\cosh\alpha} = \cosh\theta + \sinh\theta\tanh\alpha. \quad (23)$$

In this equation, we know $P/O = 2.4$, $\cosh\theta = 2$ and $\sinh\theta = \sqrt{3}$. Hence $\tanh\alpha = 0.4/\sqrt{3} = 0.23094$, and $\alpha = 0.23513$ hyperbolic radians. Again from (13)

$$A = \frac{O}{\cosh\alpha} = \frac{5}{\cosh 0.23513} = 4.865, \quad (24)$$

*A more precise rule is expressed by formula (28).

so that the general term of the series in relation to (2) is

$$4.865 \cosh(n \times 1.317 + 0.23513) \quad (25)$$

where n is any integer from $+\infty$ to $-\infty$ including zero. With $n=0$ we have $O=4.865 \cosh 0.23513=5$.

An examination of series (22), shows that all the terms have the same sign. The minimum integer of the series always lies within or in contact with the central triad.

Table II gives all of the integral cosine series for the central term $O=5$, and for the characteristic ratio of 4, ($\theta=1.317$), between terms J and T inclusive. It will be observed that α varies between $+1.6145$ and -1.6145; while A varies between 5 and 1.914. In order that all of the terms of each series shall keep the same sign, it is necessary from (23) that

$$\frac{\frac{P}{O} - \cosh\theta}{\sinh\theta} = \tanh\alpha < 1. \quad (26)$$

Hence

$$\frac{P}{O} < (\cosh\theta + \sinh\theta), \quad (27)$$

or

$$Oe^{-\theta} < P < Oe^{\theta}. \quad (28)$$

Thus with $\theta=1.317$ in (22), and $O=5$, $e^{\theta}=3.732$, and $e^{-\theta}=0.268$, $1.34<P<18.66$. The greatest integer less than 18.66 being 18, the family of cosine series for $O=5$ and $k=4$ is limited, as in Table II, to within the integral range $Oe^{-\theta}$ and Oe^{θ}, or in this case from $P=2$ to $P=18$. Outside of these limits, the series will change sign once, and changes from a hyperbolic-cosine to a hyperbolic-sine series.

As we advance along a series from the central term, in either direction, the ratio of successive terms approaches more and more nearly to the limiting value e^{θ}.

Table II presents the 17 integral cosine series for $k=4$ and $O=5$. All of the integers in the table have the positive sign. Since, however, cosh $(\theta+\alpha)$ is essentially positive, no matter what the signs of θ and α may be, it follows that by changing the sign of A, all the signs of the integers in the table will become negative.

Geometry of the Series.

If we plot the integers in the series of Table II, as ordinates, against the number of the terms as abscissae, we locate points on 17 catenary curves. The vertex of the catenary appears as a plotted point only once, *i.e.*, in series number 9, with $\alpha=0$ and $A=5$. In other words, the catenaries have all their axes parallel to the axis of ordinates through the central terms; but only in the case of number 9 does the axis of the curve coincide with the axis of ordinates.

TABLE II

Integral Hyperbolic Cosine Series for $k=4$, amd $O=5$ ($\theta=1.317$)

	$A\cosh(-5\theta+\alpha)$	$A\cosh(-4\theta+\alpha)$	$A\cosh(-3\theta+\alpha)$	$A\cosh(-2\theta+\alpha)$	$A\cosh(-\theta+\alpha)$	$A\cosh\alpha$	$A\cosh(\theta+\alpha)$	$A\cosh(2\theta+\alpha)$	$A\cosh(3\theta+\alpha)$	$A\cosh(4\theta+\alpha)$	$A\cosh(5\theta+\alpha)$		
No.	J	K	L	M	N	O	P	Q	R	S	T	α	A
1	138	37	10	3	2	5	18	67	250	933	3482	1.614	1.914
2	347	93	25	7	3	5	17	63	235	877	3273	1.122	2.944
3	556	149	40	11	4	5	16	59	220	821	3064	0.8534	3.606
4	765	205	55	15	5	5	15	55	205	765	2855	0.6585	4.083
5	974	261	70	19	6	5	14	51	190	709	2646	0.4996	4.435
6	1183	317	85	23	7	5	13	47	175	653	2437	0.3614	4.690
7	1392	373	100	27	8	5	12	43	160	597	2228	0.2351	4.865
8	1601	429	115	31	9	5	11	39	145	541	2019	0.1160	4.967
9	1810	485	130	35	10	5	10	35	130	485	1810	0	5
10	2019	541	145	39	11	5	9	31	115	429	1601	−0.1160	4.967
11	2228	597	160	43	12	5	8	27	100	373	1392	−0.2351	4.865
12	2437	653	175	47	13	5	7	23	85	317	1183	−0.3614	4.690
13	2646	709	190	51	14	5	6	19	70	261	974	−0.4996	4.435
14	2855	765	205	55	15	5	5	15	55	205	765	−0.6585	4.083
15	3064	821	220	59	16	5	4	11	40	149	556	−0.8534	3.606
16	3273	877	235	63	17	5	3	7	25	93	347	−1.122	2.944
17	3482	933	250	67	18	5	2	3	10	37	138	−1.615	1.914
Diff.	209	56	15	4	1	0	−1	−4	−15	−56	−209		

Properties of the Integral Cosine Series.

(1) An examination of Table II shows that the differences between the terms in successive series, as presented on the lowest line, are constant. These differences form a hyperbolic-function integral series of the same characteristic ratio, ($k=4$), passing through zero, and changing sign, thus constituting a sine series. These are general properties of all hyperbolic-cosine integral series. Moreover, it follows that the sums of series 8 and 10, 7 and 11, 6 and 12, etc., are constant, and equal to twice the central series number 9, or to

(28*a*) ...3620 970 260 70 20 10 20 70 260 970 3620...

a cosine series of $k=4$.

(2) If in any series of Table II, say number 13, we take out every alternate or second term, commencing say with O, we form a new series thus:

	K	M	O	Q	S
(29)	...709	51	5	19	261...

We find that this is also an integral cosine series, with $k=14=2\cosh\theta_2$, from which $\theta_2=2.634=2\theta$.

Similarly, if we take out every third term of any cosine series, or in the case considered:

$$
\begin{array}{lccccc}
 & I & L & O & R & V \\
(30) & \ldots 9095 & 175 & 5 & 85 & 4415\ldots
\end{array}
$$

this series is again an integral cosine series, of characteristic ratio $52 = 2\cosh\theta_3$, where $\theta_3 = 3.951 = 3\theta$.

Proceeding in this way, it will be evident that if we select the successive nth terms from an integral cosine series, we produce a new integral cosine series, in which

$$k_n = 2 \cosh n\theta.$$

(3) Moreover, $k, k_2, k_3, \ldots, k_n$ form evidently another integral cosine series, the law of their formation being $2\cosh\theta$, $2\cosh 2\theta$, $2\cosh 3\theta$, $\ldots$, $2\cosh n\theta$.

Consequently, any infinite main integral cosine series contains an indefinitely large number of integral cosine sub-series, formed by taking out the successive nth terms of the main series.

(4) If in any integral cosine series, starting at any term say O, we form the sums of successive equidistant pairs of terms, these sums form another integral hyperbolic function series. Thus taking series 17 on the lowest line of Table II, the double of O is 10, the sum of N and P is 20, of M and Q is 70, of L and R is 260; we thus obtain the series

(31) $\ldots 10, 20, 70, 260, 970, 3620\ldots$

This is an integral cosine series, the characteristic ratio k of which is 4, the same as that of the original series.

Again, taking say the central term doubled and the sums of the seconds, fourths, etc., on each side $(O+O)$, $(M+Q)$, $(K+S)$, etc., we have the series

(32) $\ldots 10 \quad 70 \quad 970 \quad 10510\ldots$

which is again an integral hyperbolic series, of characteristic ratio $k_2 = 14$, but which changes sign when continued towards the left, so that it is not a cosine series, but a sine series.

Again, taking say the central term doubled, and the sums of the thirds, sixths, etc., on each side $(O+O)$, $(L+R)$, $(I+V)$, etc., we find the series

(33) $\ldots 10, 260, 13510\ldots$

another hyperbolic sine series of characteristic ratio $k_3 = 52$.

Taking the successive characteristic ratios produced by this process, k_1, k_2, k_3, $\ldots$, k_n, we have the series

(34) $\ldots 4, 14, 52, 194\ldots$

an integral cosine series of $k = 4$.

(5) The sum of any two integral cosine series of the same characteristic ratio is an integral cosine series of that ratio.

Thus the sum of series 6 and 11 say, in Table II is as follows:

	J	K	L	M	N	O	P	Q	R	S	T
No. 6	1183	317	85	23	7	5	13	47	175	653	2437
No. 11	2228	597	160	43	12	5	8	27	100	373	1392
(35) Sum	3411	914	245	66	19	10	21	74	275	1026	3829

The summation series has $k=4$.

It is not necessary, however, for the central terms of the two added series to be brought together. The added series may be displaced thus:

No. 6	1183	317	85	23	7	5	13	47	175	653	2437	....	
No. 11	....		2228	597	160	43	12	5	8	27	100	373	1392
(36) Sum	....		2313	620	167	48	25	52	183	680	2537	....	

The summation series is again a cosine series of $k=4$.

Not only may any two series be added together to form a cosine series after being relatively displaced; but one or both may be reversed about any given term as fixed. Thus, we may write number 6 backwards and then add No. 11 displaced:

No. 6	2437	653	175	47	13	5	7	23	85	317	1183	..
No. 11	....	2228	597	160	43	12	5	8	27	100	373	1392
(37) Sum	....	2881	772	207	56	17	12	31	112	417	1556	....

The summation series is again a cosine series of $k=4$.

Again, a summation cosine series may be added to an original series, to produce a cosine series. Thus, adding the last sum to No. 2 of Table III:

	J	K	L	M	N	O	P	Q	R	S	T
No. 2	347	93	25	7	3	5	17	63	235	877	3273
Last sum	2881	772	207	56	17	12	31	112	417	1556	....
(38) Sum	3228	865	232	63	20	17	48	175	652	2433	....

The summation series is again a cosine series of $k=4$.

From repeated application of this theorem, it follows that the sum of n integral cosine series of the same k, displaced or reversed in any manner with respect to each other, is another integral cosine series of that k. For example, waiving displacements or reversals, the sum of the 17 series in Table II gives:

	M	N	O	P	Q
(39)	...595	170	85	170	595...

for which $k=4$.

(6) The difference of any two integral cosine series of the same characteristic ratio k is an integral hyperbolic-function series (in general, a sine series) of that k.

In Table II, an example of this theorem is found in the series of differences between successive pairs of series there presented. Taking say series 13, and subtracting from it series 3 displaced leftwards, two steps, we obtain:

			J	K	L	M	N	O	P	Q	R	S	T
	No. 13	...	...	2646	709	190	51	14	5	6	19	70	261	974
	No. 3	556	149	40	11	4	5	16	59	220	821	3064	....
(40)	Diff.	...	...	2606	698	186	46	−2	−54	−214	−802	−2994	....

This series changes sign, and is therefore an integral sine series of $k=4$.

By repeated application of this theorem, it follows that the sums and differences of any number of integral cosine series of constant k, with or without relative displacements or reversals, leads to an integral hyperbolic-function series of the same k. This proposition indeed follows from proposition (5) preceding, when the sign of A is reversible. For a further addition-subtraction property, see (79), (80) and (81).

(7) If all the terms of an integral cosine series are multiplied by a given integer, the resulting series is likewise an integral cosine series of the same k. This follows also from proposition (5) above. Thus multiplying series 6 of Table II by 10, we obtain:

	J	K	L	M	N	O	P	Q	R	S	T
(41)	11830	3170	850	230	70	50	130	470	1750	6530	24370

a new cosine series of $k=4$.

Symmetrical and dissymmetrical cosine integral series.

An integral cosine series of the general type presented in (1) to (6) leads to a dissymmetrical series, because in the central triad, the two outers are unequal. If however $\alpha=0$, or $\alpha=\frac{\theta}{2}$, the series becomes symmetrical, like that of number 9, Table II, where the outers N and P of the central triad are equal. A symmetrical series repeats itself on each side of the central term. A dissymmetrical series does not. The terms of a symmetrical cosine series are:

(42) $$\ldots+A\cosh(-3\theta)+A\cosh(-2\theta)+A\cosh(-\theta)+A+A\cosh\theta$$
$$+A\cosh(2\theta)+A\cosh(3\theta)\ldots$$

In a symmetrical cosine series, A may be any integer and $\cosh\theta$ an integer. The characteristic ratio $k=2\cosh\theta$ must then be integral. Consequently, for any given value of θ, for which $\cosh\theta$ is an integer, there must be an infinite number of integral cosine series; namely two for each integral multiple of A (one for $-nA$ and the other for $+nA$). Propositions (1) to (7) apply to both symmetrical and dissymmetrical series.

Hyperbolic-Sine Series

We consider the infinite series of complex numbers:

(42a) $$A\sinh(-m\theta+\alpha),\ \ldots A\sinh(-2\theta+\alpha),\ A\sinh(-\theta+\alpha),\ A\sinh\alpha,$$
$$A\sinh(\theta+\alpha),\ A\sinh(2\theta+\alpha),\ \ldots,\ A\sinh(n\theta+\alpha),\ \ldots.$$

where θ and α are constant complex hyperbolic angles. A is also a constant complex number for the series, positive, negative or zero, integral or fractional. The numbers $-m$ and n are real integers, which may be indefinitely large. We may call any number belonging to this series, such as $A\sinh(n\theta+\alpha)$, a *hyperbolic-sine number* and the series a general *hyperbolic-sine series.*

Mean-to-Mid Constant Ratio of a Sine Series

Let us take any triad of a hyperbolic sine series.

(43) $$A\sinh\{(n-1)\theta+\alpha\}=A\sinh\{(n\theta+\alpha)-\theta\}, \qquad I$$

(44) $$A\sinh\{n\theta+\alpha\} \quad =A\sinh\{n\theta+\alpha\}, \qquad II$$

(45) $$A\sinh\{(n+1)\theta+\alpha\}=A\sinh\{(n\theta+\alpha)+\theta\}. \qquad III$$

These may be rewritten as:

(46) $$A\{\sinh(n\theta+\alpha)\cosh\theta-\cosh(n\theta+\alpha)\sinh\theta\}, \qquad I$$

(47) $$A\{\sinh(n\theta+\alpha)\}, \qquad II$$

(48) $$A\{\sinh(n\theta+\alpha)\cosh\theta+\cosh(n\theta+\alpha)\sinh\theta\}. \qquad III$$

Adding together the outers,

(49) $$I+III=2A\{\sinh(n\theta+\alpha)\cosh\theta\}=II\ 2\cosh\theta=II.k.$$

Consequently, the complex sum of the outers of any triad is equal to the mid multiplied by $k=2\cosh\theta$, the characteristic constant of the series. Expressing the same property in another way, the complex or planevector mean of the outers is

(50) $$\frac{I+III}{2}=A\sinh(n\theta+\alpha)\cosh\theta,$$

and the ratio of this mean to the mid is

(51) $$\frac{I+III}{2II}=\cosh\theta.$$

These are the same ratios as pertain to any complex cosine series, by (8) and (9).

Integral Hyperbolic-Sine Series

As we are here only interested in the integral sine series, we can set aside the complex cases, which give rise to integral numbers only in exceptional cases. Confining then our attention to real values for A, θ and α, we require that k should be an integer as in (11).

Calling as before the initial triad of the series, N, O, P

(52) $$N=A\sinh(-\theta+\alpha)=-A\sinh(\theta-\alpha),$$

(53) $$O=A\sinh\alpha,$$

(54) $$P=A\sinh(\theta+\alpha).$$

It is necessary that O should be an integer (positive, negative or zero). If then N be assigned any integral value, P will necessarily become an integer; viz.:

(55) $$P=kO-N.$$

There will be either a sine or cosine integral series for each integral value of N selected. The ratio of P to O for a sine series is

(56) $$\frac{P}{O}=\frac{A(\sinh\theta\cosh\alpha+\cosh\theta\sinh\alpha)}{A\sinh\alpha}=\cosh\theta+\sinh\theta\coth\alpha,$$

so that

(57) $$\coth\alpha=\frac{\frac{P}{O}-\cosh\theta}{\sinh\theta}$$

with α real, $\coth\alpha$ must be greater than unity, and therefore

(58) $$\frac{P}{O}>(\cosh\theta+\sinh\theta) \quad \text{or} \quad \frac{P}{O}>e^{\theta},$$

that is

(59) $$Oe^{-\theta}>P>Oe^{\theta}.$$

Thus a sine series will be produced whenever P lies outside of the range $Oe^{-\theta}$ to Oe^{θ}. Otherwise, a cosine series will be produced, by (28).

Table III presents the J to T terms of the first dozen sine series based on $O=3$ with $k=3$, or $\theta=\cosh^{-1}(3/2)=0.9624$. The value of $e^{0.9624}$ being 2.617, and $e^{-0.9624}$ being 0.382, P must be outside the limits 7.851 and 1.146; *i.e.*, outside of the integral range 7 to 2. If either N or P falls below 2, a sine series will be produced, but if neither N nor P falls below 2, a cosine series will be formed. Thus series numbers 7 to 12 inclusive, in Table III, are cosine series and do not change sign; whereas all the other series are sine series and the terms change sign. There is no limit to the number of sine series beyond the confines of the table, based upon $O=3$ alone; but those above are seen to be the respective reversed counterparts of those below. Thus series number 2 is the same as number 17 reversed. Other systems of sine series could be based on $O=3$, using other integral values of k, as indicated briefly in Table I. Again, other systems could be based on different values of O.

Properties of the Integral Sine Series.

The properties of integral sine series are analogous to those of integral cosine series, which have been already considered. Consequently the sinusoidal properties may be considered more briefly in what follows.

(1) As shown in Table III, the differences between the terms of the same value of n in successive series, are constant, and are presented in the lowest line. These differences also form a sine series of the same characteristic ratio ($k=3$). Summing up what was pointed out in this respect for cosine series, we may say that the tabular differences between corresponding terms of hyperbolic-function series of a given k are constant and form a sine series of the same k. Moreover, the sums of 9 and 10, 8 and 11, 7 and 12, 6 and 13, etc., are constant.

(2) If in any series of Table III, say number 17, we take out every alternate or second term, commencing say with O, we form a new series, thus:

	K	M	O	Q	S
(60)	228	33	3	-12	-87.

TABLE III

Table III of Integral Sine and Cosine Series for O = 3 and $k=2\cosh\theta=3$, or $\theta=0.96242$

	$A\sinh(-5\theta+\alpha)$	$A\sinh(-4\theta+\alpha)$	$A\sinh(-3\theta+\alpha)$	$A\sinh(-2\theta+\alpha)$	$A\sinh(-\theta+\alpha)$	$A\sinh\alpha$	$A\sinh(\theta+\alpha)$	$A\sinh(2\theta+\alpha)$	$A\sinh(3\theta+\alpha)$	$A\sinh(4\theta+\alpha)$	$A\sinh(5\theta+\alpha)$		
	J	K	L	M	N	O	P	Q	R	S	T	α	A
1	−283	−108	−41	−15	−4	3	13	36	95	249	652	0.4172	6.986
2	−228	− 87	−33	−12	−3	3	12	33	87	228	597	0.4812	6.0
3	−173	− 66	−25	− 9	−2	3	11	30	79	207	542	0.5709	4.98
4	−118	− 45	−17	− 6	−1	3	10	27	71	186	487	0.7087	3.898
5	− 63	− 24	− 9	− 3	0	3	9	24	63	165	432	0.9624	2.683
6	− 8	− 3	− 1	0	1	3	8	21	55	144	377	1.925	0.8943
7	47	18	7	3	2	3	7	18	47	123	322		
8	102	39	15	6	3	3	6	15	39	102	267		
9	157	60	23	9	4	3	5	12	31	81	212	Cosine	
10	212	81	31	12	5	3	4	9	23	60	157	Series	
11	267	102	39	15	6	3	3	6	15	39	102		
12	322	123	47	18	7	3	2	3	7	18	47		
13	377	144	55	21	8	3	1	0	− 1	− 3	− 8	−1.925	−0.8943
14	432	165	63	24	9	3	0	− 3	− 9	− 24	− 63	−0.9624	−2.6833
15	487	186	71	27	10	3	−1	− 6	−17	− 45	−118	−0.7087	−3.898
16	542	207	79	30	11	3	−2	− 9	−25	− 66	−173	−0.5709	−4.98
17	597	228	87	33	12	3	−3	−12	−33	− 87	−228	−0.4812	−6
18	652	249	95	36	13	3	−4	−15	−41	−108	−283	−0.4172	−6.986
Diff.	− 55	− 21	− 8	− 3	−1	0	1	3	8	21	55		

We find that this is also an integral sine series with $k_2=7$ from which $\cosh\theta_2=3.5$, or $\theta_2=1.92485=2\theta$.

Again, taking every third term, or in the case considered,

$$\begin{array}{lccccc} & J & L & O & R & U \\ (61) & 1563 & 87 & 3 & -33 & -597, \end{array}$$

this series is again an integral sinh series of $k_3=18$, or $\theta_3=\cosh^{-1}(18/2)=2.88727=3\theta$.

Proceeding in this way, it will be found, and can also be readily demonstrated, that the successive nth terms of an integral sinh series whose angle is θ form another sinh series whose angle is $n\theta$.

(3) Moreover k_1, k_3, k_5, ... form another hyperbolic-function series of the same k as the original series. In the case considered it is:

(62) $\quad 3 \quad 2 \quad 3 \quad 7 \quad 18 \quad 47 \quad 123,$

a cosine series, No. 12 in Table III.

(4) If in any integral sine series, starting at any term say O, we form the sums of successive equidistant pairs of terms, these sums form another hyperbolic-function series. Thus, taking series 13 in Table III, the double of O is 6, the sum of N and P is 9, of M and Q 21, of L and R 54. We thus arrive at the cosine series

(63) $\ldots 6 \quad 9 \quad 21 \quad 54 \quad 141 \quad 369 \ldots$

whose k_2 is 3, the same as that of the original series.

Again, following the procedure in Series 13, Table III, indicated in connection with (32), we find the cosine series

(64) $\ldots 6, 21, 141, 966 \ldots$

whose k_3 is 7 and $\theta_3 = 1.92485$.

The series of $k_2, k_3, k_4, \ldots$ formed in this way, is the cosine series

(65) $\ldots 3 \quad 7 \quad 18 \quad 47 \ldots$

whose k is 3.

(5) As in connection with (35) to (38), the sum of any sine series with another sine series of the same k, displaced or not, is another hyperbolic-function series of that k. Thus in Table III, adding series 3 to series 16 displaced rightwards two points, we have

No. 3	−173	−66	−25	−9	−2	3	11	30	79	207	542	....	.....
No. 16	.....	.....	542	207	79	30	11	3	−2	−9	−25	−66	−173
(66) Sum	.....	.....	517	198	77	33	22	33	77	198	517	....	.....

a cosine series of $k=3$.

Similarly the sum of a sine series and a cosine series of like k, either being displaced or reversed, produces another hyperbolic-function series of like k. Thus adding in Table III series 9 to series 15 displaced one point leftwards, we obtain

No. 9	...	157	60	23	9	4	3	5	12	31	81	212
No. 15	487	186	71	27	10	3	−1	−6	−17	−45	−118	...
(67) Sum	...	343	131	50	19	7	2	−1	− 5	−14	− 37	...

a sine series of $k=3$.

Repeated application of this theorem leads to the conclusion that the sum of any number of integral hyperbolic-function series including displacements, all having the same k, is another integral hyperbolic-function series of that k.

(6) Since we can always reverse the sign of A in any hyperbolic-function series, and thus repeat the same with every sign reversed, it follows from the last proposition (5), that the sums and differences of any number of hyperbolic-function integral series of the same k, including displacements, is likewise a hyperbolic-function series of like k. Thus the sum of all the 18 series in Table III gives the series

	J	K	L	M	N	O	P	Q	R	S	T
(68)	3321	1269	486	189	81	54	81	189	486	1269	3321,

a cosine series of $k=3$.

(7) As in connection with (41), repeated application of the addition theorem (5) leads to the conclusion that if any integral sine or cosine series be multiplied by any integer, the resulting series will be likewise a sine or cosine series respectively, of the same k.

Symmetrical and Dissymmetrical Sine Integral Series

An integral sine series of the general type presented in (43) to (48), leads to dissymmetrical series; because in the central triad the two outers are not the same number, with opposite signs. If, however, $\alpha=0$, or $\alpha=\pm\frac{n\theta}{2}$, where n is any integer (see Table III, series 2, 5, 6, 13, 14 and 17), the series becomes symmetrical. A symmetrical series repeats itself with opposite signs on each side of the central zero term, or on each side of the change of sign. In the former case, the terms of a symmetrical sine series are:

$$(69) \qquad \ldots A\sinh(-3\theta)+A\sinh(-2\theta)+A\sinh(-\theta)+0+A\sinh\theta+A\sinh(2\theta)+A\sinh(3\theta)+\ldots$$

the term A sinh θ must be either a positive or negative integer, and the characteristic ratio $k=2$ cosh θ must likewise be an integer. Consequently, for any such value of θ, there must be an infinite number of integral sine series; namely two for each integral multiple of A (one for $-nA$ and the other for $+nA$, which may be regarded however as the mutual reversals one of the other). Propositions (1) to (7) apply, however, to both symmetrical and dissymmetrical series.

Limiting case of $\theta=0$.

It may be noted that the series of successive integers

$$(70) \qquad \ldots -6 \quad -5 \quad -4 \quad -3 \quad -2 \quad -1 \quad 0 \quad 1 \quad 2 \quad 3 \quad 4 \quad 5 \quad 6\ldots$$

is a sine series of characteristic ratio $k=2$, or cosh $\theta=1$, for which $\theta=0$. The substitution of $\theta=0$ in (69) leads to the series

$$(71) \qquad \ldots 0 \quad 0 \quad 0 \quad 0 \quad 0 \ldots$$

but if we make A an indefinitely large number and $\theta=1/A$, we obtain series (70) as the limiting value $\theta=0$ is approached.

Exponential numbers.

If we add together the general sine series (42*a*) and the general cosine series (1*a*) we obtain:

$$\ldots A\sinh(-2\theta+\alpha)+A\sinh(-\theta+\alpha)+A\sinh\alpha+A\sinh(\theta+\alpha)+A\sinh(2\theta+\alpha)+\ldots$$
$$\ldots A\cosh(-2\theta+\alpha)+A\cosh(-\theta+\alpha)+A\cosh\alpha+A\cosh(\theta+\alpha)+A\cosh(2\theta+\alpha)+\ldots$$

$$(72) \qquad \ldots Ae^{(-2\theta+\alpha)}+Ae^{(-\theta+\alpha)}+Ae^{+\alpha}+Ae^{(\theta+\alpha)}+Ae^{(2\theta+\alpha)}+\ldots$$

which can be written in the following form if $Ae^{\alpha}=B$

(73) $$\ldots Be^{-2\theta}+Be^{-\theta}+B+Be^{\theta}+Be^{2\theta}+\ldots$$

This series may be called a series of exponential numbers. In order that (73) may be an integral series, we require that B should be an integer and also e^{θ}. The characteristic ratio $k=\dfrac{e^{2\theta}+1}{e^{\theta}}=e^{\theta}+e^{-\theta}=2\cosh\theta$; but this ratio does not need to be an integer. The series will run to infinity on the right hand side, for any positive value of θ; but must terminate on the left hand side with the integer nearest above unity. Dividing then (73) by B, we have the exponential series

(74) $$1+e^{\theta}+e^{2\theta}+e^{3\theta}+\ldots$$

The ratio of each term to its predecessor is e^{θ}; or each term is the geometrical mean of its predecessor and successor, as well as being their arithmetical mean divided by $\cosh\theta$.

Table IV gives the first five integral exponential series for $B=1$. Each series is capable of producing an infinite number of similar series by taking B through successive integral values from $B=2$ to $B=\infty$.

TABLE IV

Integral Exponential Series for $B=1$

	e^{0}	e^{θ}	$e^{2\theta}$	$e^{3\theta}$	$e^{4\theta}$	$e^{5\theta}$				
No.	O	P	Q	R	S	T	e^{θ}	θ	$\cosh\theta$	k
1	1	1	1	1	1	1	1	0	1	2
2	1	2	4	8	16	32	2	0.6932	1.25	2.5
3	1	3	9	27	81	243	3	1.0985	$1.6\dot{6}$	$3.3\dot{3}$
4	1	4	16	64	256	1024	4	1.3865	2.125	4.25
5	1	5	25	125	625	3125	5	1.6095	2.6	5.2

TABLE V

Table V contains the first five series corresponding to No. 3 in Table IV

No.	O	P	Q	R	S	T	e^{θ}	θ	$\cosh\theta$	k
1	1	3	9	27	81	243	3	1.0985	1.667	3.333
2	2	6	18	54	162	486	3	"	"	"
3	3	9	27	81	243	729	3	"	"	"
4	4	12	36	108	324	972	3	"	"	"
5	5	15	45	135	405	1215	3	"	"	"
Diff.	1	3	9	27	81	243	"	"	"	"

Properties of the Integral Exponential Series.

(1) Table V shows that the differences between the terms of successive series of constant k are constant. These differences form an exponential series of the same k.

(2) Taking alternate terms in any series, say No. 4, we find

(75) $$4 \quad 36 \quad 324 \quad 2916\ldots$$

an exponential series of $k_2=9.11$, or $\theta_2=2.197=2\theta$. Similarly, taking every third term, we obtain

$$4 \quad 108 \quad 2916 \quad 70732\ldots$$

an exponential series of

(76) $$k_3 = 27.037, \text{ or } \theta_3 = 3.2955 = 3\theta.$$

Similarly, taking every nth term we find a series of $\theta_n = n\theta$.

(3) Moreover, $k_1, k_2, k_3, \ldots, k_n$ form another exponential series of the same k as the original series.

(4) The sum or difference of any two integral exponential series of the same k, with or without mutual displacements, is a new series of that k. Thus in Table IV adding 3 to 5 displaced rightwards one point

	No. 3	3	9	27	81	243	729	....
	No. 5	..	5	15	45	135	405	1215
(77)	Sum	..	14	42	126	378	1134	....
(78)	Difference	..	4	12	36	108	324	....

Each of these two series has the constant $k = 3.\dot{3}$. Repeated application of this theorem leads to the conclusion that the sums or differences of any number of exponential series of given k, with or without displacements, lead to a new exponential series of that k.

It will be observed that the properties of sine, cosine, and exponential integral series are analogous. All three types may be included under the title hyperbolic-function series.

Additional Properties of Hyperbolic-Function Series

If instead of adding or subtracting two different hyperbolic-function integral series of the same k, as in (35) to (40), (63) to (67) and (77) to (78), we confine ourselves to one and the same series with mutual displacements, we arrive at another form of the addition-subtraction theorem, which is a corollary to that already enunciated. Thus, taking series 13 of Table III and adding it to itself with a displacement of one step rightwards, we have:

		377	144	55	21	8	3	1	0	−1	−3	−8
		...	377	144	55	21	8	3	1	0	−1	−3
(79)	Sum	...	521	199	76	29	11	4	1	−1	−4	−11
(80)	Diff.	...	−233	−89	−34	−13	−5	−2	−1	1	2	5

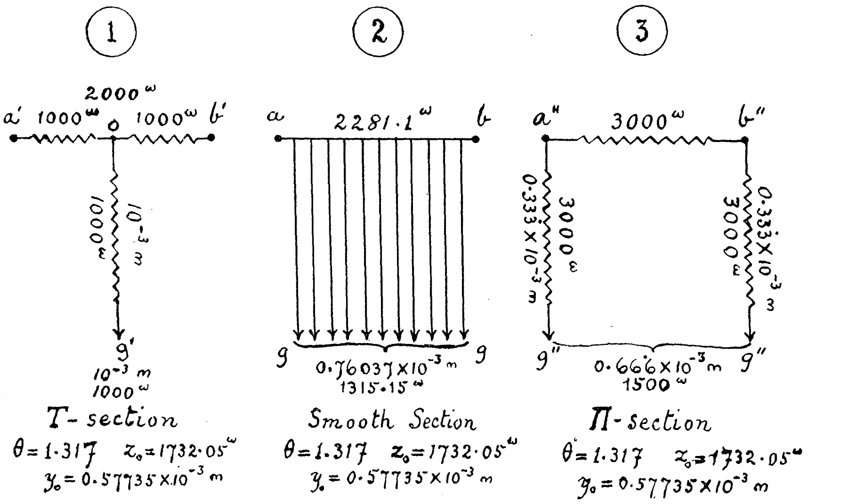
Each of these resulting series is a sine series of $k = 3$, and this can be shown to be a general proposition, by the addition theorem. Consequently, if we add to each term of a series the adjacent term on the same side, the new series will be a hyperbolic-function series of the same k. By the repeated application of this corollary, it will be evident that if we add to each term of any hyperbolic-function series, the mth, nth and pth terms on one side of it, subtracting also say the rth term from it, the resulting term will be a hyperbolic-function series. Thus if to each term of Series 3 in Table II, we add the next and also the next but one, right hand terms and subtract the left-hand term, we virtually add the following four series:

		.....	−173	−66	−25	−9	−2	3	11	30	79	207	542
		−173	−66	−25	−9	−2	+3	11	30	79	207	542	...
	−173	−66	−25	−9	−2	3	11	30	79	207	542	...	...
				173	66	25	9	2	−3	−11	−30	−79	−207
(81)	.....	.....	.....	73	30	17	21	46	117	305	798	....	.....

This series is a cosine series of $k=3$.

Conditions Under which Hyperbolic-Function Series May Present Themselves in Electrical Engineering

It is shown in text-books* relating to the hyperbolic-function properties of a uniform conducting line, real or artificial, operated in the steady state by continuous or alternating currents, that a symmetrical T-section of resistances, as in Fig. 1, or a smooth line of uniformly distributed resistance and leakance, as in Fig. 2, or a symmetrical Π-section of resistance as in Fig. 3, subtends a certain

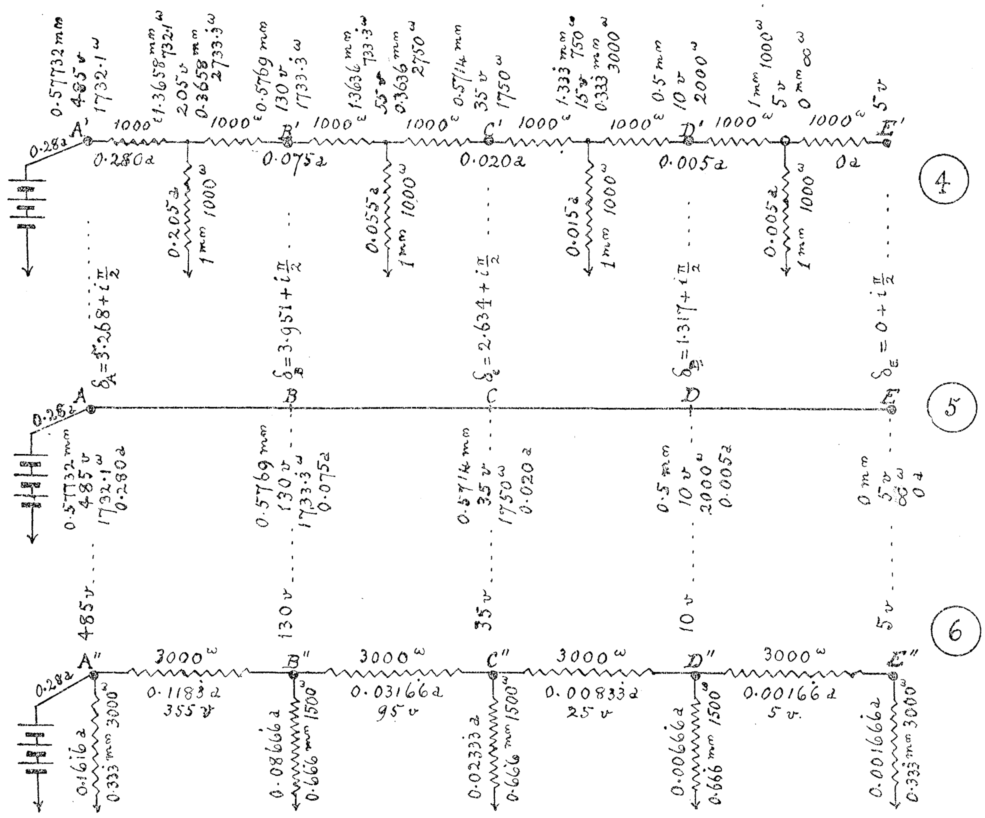


Figs. 1, 2, 3. T-Section, Smooth Section and Π-Section, all externally equivalent, each subtending an angle $\theta = 1.317$ hyperbolic radians.

real hyperbolic angle θ. If a line *ABCDE* of four sections, of either of these types, be connected as shown in Figs. 4, 5 and 6, with the distant end freed at *E*, and a steady e.m.f. of 485 volts be impressed at *A*, the potentials of the points *A*, *B*, *C*, *D*, *E* will be directly proportional to the sines of the position angles of these points, or in this case, to 5 cosh 4θ, 5 cosh 3θ, 5 cosh 2θ, 5 cosh θ and 5 cosh 0, where $\theta = 1.317$ hyp. radians. These potentials will be 485, 130, 35, 10 and 5 volts respectively, which form the cosine series No. 9 of Table II. This shows

**The application of Hyperbolic Functions to Electrical Engineering Problems,* by A. E. Kennelly, The University of London Press, 3rd Edition 1924, and McGraw-Hill Book Co., N.Y. *Artificial Electrical Lines,* by A. E. Kennelly, McGraw-Hill Book Co., N.Y., 1917, *Les applications Elémentaires des Fonctions Hyperboliques à la Science de l'Ingénieur Electricien,* by A. E. Kennelly, Gauthier-Villars, Paris, 1922.

that the series No. 9, Table II can be realized by the potentials at terminals and section-junctions of any one of the three types of line indicated in Figs. 1 to 6. By extending the number of line sections indefinitely, the numbers in the



Figs. 4, 5, 6. T, Smooth and Π Lines Voltaged at A and freed at E.

cosine series might be likewise extended, as far as might be practically feasible. Moreover, the voltage drops in these sections follow the series

(82) 355 95 25 5 volts,

which is evidently a sine series of $k=4$. The currents flowing along the line at terminals and junctions, when expressed in milliamperes are

(83) 280 75 20 5 0,

see Fig. 4, and this is also a sine series of $k=4$. The architrave currents in Fig, 6, when expressed in milliamperes and multiplied by 3, become identical with (82).

The single lines of Figs. 4, 5 and 6, freed at the distant end E, only present parts of the sine or cosine series in the cases above referred to; but by taking two such lines in series, and applying voltages to both outer ends, complete integral series of any desired numbers of terms may be obtained. Thus Fig. 7 shows the central five sections of such a Π-line with 5 volts at E, 10 volts at D and D', and 35 volts at C and C'. Extending this line in both directions, the series No. 9 of Table II might readily be realized, at least as far as it is carried in that Table.

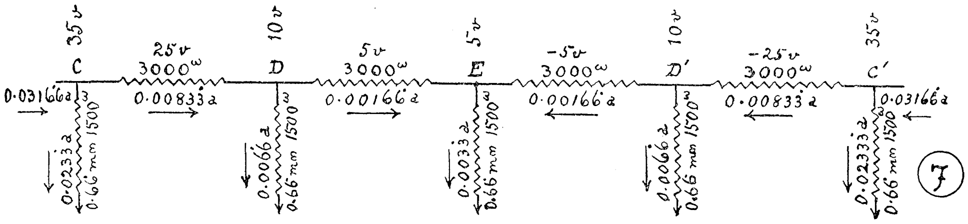
If now we apply 39 volts at C', and 31 volts at C, as in Fig. 8, we obtain the potential series

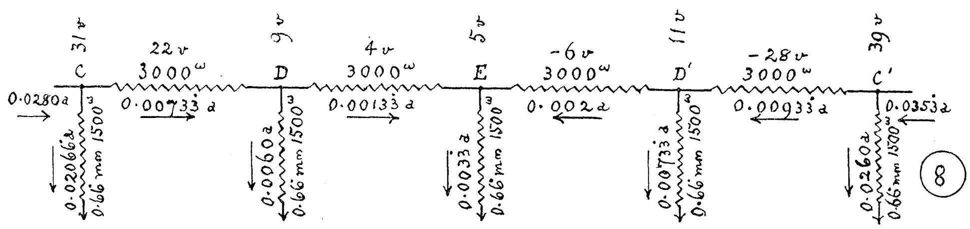
(84) 31 9 5 11 39

which corresponds to series 8 of Table II, while the successive architrave voltage drops are

(85) 22 4 -6 -28 volts,

a sine series of $k=4$.





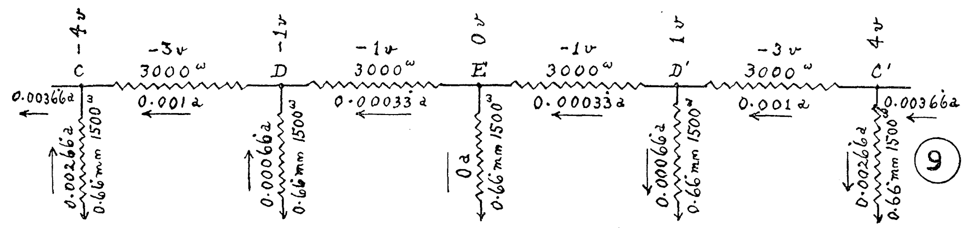


Fig. 7.—Cosine Series of Junction Potentials, according to Line 9, Table II.

Fig. 8.—Cosine Series of Junction Potentials, according to Line 8, Table II.

Fig. 9.—Sine Series of Junction Potentials, according to differences between Figs. 7 and 8.

By comparing the distributions of Figs. 7 and 8, it will be seen that they differ by the distribution of Fig. 9, where 4 volts is applied at C' and -4 volts at C. The junction potentials in Fig. 9 conform to the difference series at the foot of Table II.

Proceeding in the manner indicated in Figs. 1-9, it may be shown that any hyperbolic-function series of integral numbers may be realized by the potentials or the currents at the junctions of suitable uniform sections of real or artificial line, suitably energized at the final outer terminals. The line section T, smooth, or Π, must be chosen conformably with the θ of the series and must subtend the angle θ, the position angles at section junctions must then be made to correspond to $(-n\theta+\alpha)$, . ., $(-\theta+\alpha)$, α, $(\theta+\alpha)$, . . , $(n\theta+\alpha)$, by suitably adjusting the terminal currents or voltages. The integral magnitudes of either currents or voltages will then be adjusted in accordance with the A of the series. In fact, the existence of these hyperbolic-function series of integral numbers was first stumbled upon* by actual measurements of such potentials and currents at the junctions of artificial lines in the laboratory.

Corresponding to the mathematical property of the superposition of hyperbolic-function series rests the physical property of the superposition of electric potential and current distributions in a conducting network.

LIST OF SYMBOLS EMPLOYED

- A A constant, in general complex, but specifically an integer, positive, negative, or zero.
- a Abbreviation for amperes.
- α A constant hyperbolic angle, in general complex, but specifically real.
- B A constant integer.
- $ABCDE$ Terminals and junction points of a uniform-section conducting line.
- δ Position angle of a point on a conducting line (Hyp. radians).
- $e=2.71828...$ Napierian base.
- θ Hyperbolic angle in general complex, but specifically real, also the hyperbolic angle subtended by a section of real or artificial conducting line (Hyp. radians).
- $\theta_1\,\theta_2$ Components of a complex hyperbolic angle θ.
- $\theta_1, \theta_2, \theta_3$ Real hyperbolic angles.
- $i=\sqrt{-1}$.
- $k=2\cosh\theta$, a constant, in general complex, but specifically real. The Characteristic ratio of a series.
- $k_1k_2k_3$ Set of characteristic ratios.
- m℧ Abbreviation for millimhos.
- m_1n_1 Real integers specifying terms in a series.
- v Abbreviation for volts.
- $\pi=3.14159....$
- y_0 Surge admittance of line (mhos).
- z_0 Surge impedance of line (ohms).
- ∞ Infinity.
- ω Abbreviation for ohms.
- ℧ Abbreviation for mhos.

**Artificial Electric Lines*, by A. E. K., p. 314.

A NEW FORMULA FOR USE IN CALCULATING REPULSION OF COAXIAL COILS

By Dr. H. B. Dwight,
Canadian Westinghouse Company, Hamilton, Canada.

A practical problem in electrical engineering is the calculation of the mechanical force in pounds exerted between air-cored reactance coils, which are installed in power stations to limit the value of short-circuit currents. It is usually necessary to economize space in the power stations, and the coils are often placed so close together that the force between them amounts to several thousand pounds at times of short-circuit, when large currents of the order of 20,000 amperes flow through them. It is desirable to calculate the magnitude of the force in order that the coil supports may be designed to have sufficient strength. Since commercial reactance coils have considerable radial thickness, it is desirable to use formulas which take the thickness into account.

If s is the distance in centimetres between two air-cored coils, the mechanical force in dynes between them when one absampere is flowing in each, is equal to the differential with respect to s of the mutual inductance of the coils, using absolute electromagnetic units, the currents being in-phase alternating currents or direct currents.

When the coils are side by side with parallel axes, one formula can be used for all positions in usual practice*. This formula is derived from expressions involving zonal harmonics, and is expressed in the form of an algebraical series.

When the coils are placed end to end along the same axis, the best accuracy is not obtained by employing a single formula, but it is found more suitable to make use of the differential of the well-known expression for the mutual inductance of two coaxial coils of equal diameter and of lengths g and h†,‡, namely,

$$(1) \qquad \frac{d}{ds} M_{gh} = \frac{d}{ds} \tfrac{1}{2} \left[L_{g+h+t} + L_t - L_{g+t} - L_{h+t} \right],$$

where L_t is the self-inductance of an assumed coil of length t of similar section to the two coils and filling the space between them, where L_{g+h+t} is the self-inductance of the complete coil of length $g+h+t$, and so on.

**Some New Formulas for Reactance Coils,* by H. B. Dwight, Trans. Amer. Inst. Elect. Eng., 1919, p. 1675, equation (4).

†*Formulas and Tables for the Calculation of Mutual and Self-Induction,* by E. B. Rosa and F. W. Grover, Scientific Paper No. 169 of the Bureau of Standards, Washington, D.C., 1911, equation 51.

‡*Repulsion and Mutual Inductance of Reactance Coils with the Same Axis,* by H. B. Dwight, The Electric Journal, May, 1918, page 166, equation (1).

The differentials of various self-inductance formulas can be used for the four terms of expression (1), depending on the relative lengths of the portions of coil considered. Such formulas are generally derived from expressions in elliptic integrals, and can sometimes be expressed as simple algebraical series. However, the choice of formulas is considerably restricted because there are few formulas which take into consideration the radial thickness, c, of the coil and which can be readily differentiated. The value of expression (1) is a comparatively small difference between larger quantities, each of which must therefore be calculated with some precision.

Differential expressions for use in formula (1) were published by the writer in his second paper referred to in the footnote on the preceding page. In this paper, where l is the length of the portion of coil considered, equation (3) holds for cases where l is greater than d, equation (4) for cases where l is less than d and greater than c, and equation (5) for cases where l is less than c.

In practical calculations, the poorest precision is in the range $\frac{l}{d}=1$ to $\frac{l}{d}=1.3$. The following formula is presented herewith to replace equation (3) in the paper cited, since it is more rapidly convergent and gives greater precision with fewer terms, especially in the range just described. Although it appears longer, that is merely because more terms have been given. The formula in question is:

$$
\begin{aligned}
(2)\qquad Q_l = \tfrac{1}{2}\pi^2 dlm^2 \Big[& 2q+\frac{3}{4}q^5-\frac{5}{8}q^7+\frac{105}{64}q^9-\frac{189}{64}q^{11} \\
& +\frac{11\times 315}{512}q^{13}-\frac{13\times 297}{256}q^{15}+\frac{45\times 13013}{16384}q^{17}-\dots\dots\dots \\
& -\frac{2}{3}\frac{c}{l}+\frac{1}{3}\frac{c^2}{dl}+\frac{c^2}{d^2}\Big\{-\frac{2}{3}q^3+\frac{5}{2}q^5-\frac{20}{3}q^7+\frac{665}{32}q^9-\frac{1953}{32}q^{11} \\
& +\frac{11\times 2135}{128}q^{13}-\frac{143\times 1961}{512}q^{15}+\frac{195\times 68{,}915}{8192}q^{17}-\dots\Big\} \\
& +\frac{c^4}{d^4}\Big\{-\frac{1}{9}q^3+\frac{17}{15}q^5-\frac{265}{24}q^7+\frac{7\times 1265}{144}q^9-\frac{9\times 38{,}857}{1152}q^{11} \\
& +\frac{11\times 3913}{32}q^{13}-\frac{3003\times 9551}{5120}q^{15}+\frac{2145\times 10{,}625}{1024}q^{17}-\dots\Big\} \\
& +\frac{c^6}{d^6}\Big\{\frac{1}{10}q^5-\frac{75}{28}q^7+\frac{7\times 1117}{168}q^9-\frac{9\times 1183}{24}q^{11}+\frac{231\times 3641}{256}q^{13} \\
& -\frac{143\times 367{,}621}{2560}q^{15}+\frac{2145\times 109{,}353}{2048}q^{17}-\dots\Big\}+\dots\Big]
\end{aligned}
$$

where

$$
(3)\qquad q^2=\frac{d^2}{d^2+4l^2}.
$$

In the above formula, the dimensions may all be in inches, or they may all be in centimetres. The quantity m is not the number of turns in a coil, but it is the number of turns per inch of coil when all the dimensions are in inches, and it is the number of turns per centimeter of coil when all the dimensions are in centimetres.

This formula is obtained by differentiating formulas (11) and (13) of the first reference on the first page of this paper, which formulas together constitute a very rapidly convergent self-inductance formula for coils whose length is equal to or greater than their diameter, and, as may be observed, terms involving the radial thickness, c, of the coil are included. In effecting the derivation it may be noted that

$$\frac{d}{dl}(q^n) = \frac{-4lnq^{n+2}}{d^2} \quad . \tag{4}$$

Since differentiating these power series makes their convergence less rapid, there is even more need for a series of rapid convergence in calculating mechanical force than there is in calculating self-inductance.

As in equation (6) of the third reference in this paper, the average force in pounds is

$$F = \frac{I_1 I_2}{4.45 \times 10^7} \left(Q_{g+t} + Q_{h+t} - Q_{g+h+t} - Q_t\right) \tag{5}$$

where I_1 and I_2 are in amperes.

Formula (5), as described in the reference just cited, gives the average force or the steady push exerted between coils for single-phase or three-phase currents. The maximum momentary value of force for a single-phase short-circuit, which is of considerable importance, is equal to twice the average force. The maximum momentary value of force for a three-phase short-circuit, which sometimes occurs, may be obtained by using the single-phase formulas combined with a process of drawing curves of instantaneous values of currents in the phases.

In making calculations when the coils are far enough apart so that t is greater than d, it is worth while evaluating first the expressions of the form

$$\left(l_1 q_1^n + l_2 q_2^n - l_3 q_3^n - l_4 q_4^n\right) \tag{6}$$

because formula (2) of this paper is used for all four terms of (5).

Example: Find the force between two like coils placed end to end, in which a single-phase short-circuit current of 11,700 effective amperes is flowing:

Mean diameter $= d = 22.5$ inches.
Radial thickness $= c = 13.5$ inches.
Axial length of each coil $= g = h = 8.5$ inches.
Air space between coils $= t = 31.5$ inches.
Turns per inch of coil $= m = 11.8$.

The calculated average force is $8500 + 540 + 80 - 10 = 9110$ pounds.

This example is an actual case taken from practical engineering. The extra convergence of the formula presented in this paper is necessary in order to calculate this example.

THE STEAM TURBINE

By the Honourable Sir Charles A. Parsons,
Chairman of the Marine Steam Turbine Company, Wallsend-on-Tyne, England.

The object of the present paper is to correlate certain well known principles of thermodynamics with the practical progress of the steam turbine and with important advances that are being made at the present time. These advances promise considerable improvements in the efficiency of fuel utilization for the production of motive power.

The story of the steam engine and, later, of the steam turbine, is largely the story of progress along the path pointed out by the genius of the earlier pioneers—Savory, Newcomen, Smeaton, the famous Watt, Wolff, Trevethick and others who, in the words of Carnot, were the veritable creators of the steam engine. Theoretical knowledge and mechanical progress have advanced along that path, side by side, and to-day the engineer has at his command a comprehensive knowledge of the properties of steam and indeed of thermodynamics, and almost unlimited mechanical resources for the execution of his designs.

Present day knowledge of both the theory and the technique of steam engines has justified Watt in the principles which he laid down, which were that the greater the steam pressure and range of expansion, the greater will be the work obtained from a given weight of steam, and that the cylinder should be kept as hot as the steam which enters it. The main development of the steam engine since his time has consisted in an extension of the range of expansion and in an ever increasing degree of compounding. The separation of the condenser from the cylinder by Watt was followed by the division of the expansion into stages in separate cylinders, finally into four stages in the highest development of the reciprocating engine. The steam turbine carries this subdivision still further, the number of stages included in a reaction turbine of high efficiency being commonly sixty or more. In the compound reciprocating engine the primary object of compounding is to reduce the temperature range of each individual cylinder, and so to diminish the condensation and loss which occur when hot steam is admitted into a cylinder the walls of which have been cooled to the temperature of the steam just discharged from it. In a steam turbine on the other hand, it is interesting to note that the turbine ideally fulfils Watts principle of keeping the cylinder as hot as the steam that enters it, because the turbine at any part is only exposed to the steady temperature of the steam that is passing through it.

The extensive compounding which is now adopted in all types of steam turbines designed for high efficiency is of vital importance to their performance.

There appears to be considerable evidence from the results obtained from actual turbines that the efficiency of conversion of the available energy of the steam into kinetic energy is greater at low steam velocities than at high steam velocities. Since the high velocity ratio necessary for efficiency can be attained in two ways, either by reducing the steam velocity or by increasing the blade velocity, the most efficient type would appear to be that in which both steam velocity and blade velocity are kept moderate, by adopting a large number of pressure stages.

The consideration of the complete turbine as a series of simple turbines, in each of which the pressure drop is small, is consistent with a simple mathematical treatment for the determination of the energy available from an expanding gas. By the fundamental principles of hydraulics, if p be the pressure and ρ the density of an incompressible fluid, and if we consider V as the velocity generated at each stage by the conversion of pressure energy into velocity energy, we have simply $V^2/2g = -\Delta p/\rho$, which is true also of an expansible fluid when the pressure drop is so small that ρ is sensibly constant. Thus the total kinetic energy so generated per pound of steam in a large number of stages is $-\sum\Delta p/\rho$ or $-\sum v\Delta p$, if v is the specific volume. We thus see that the available energy of an expanding gas is $-vdp$, for a small pressure drop, and $-\int vdp$, between wider limits of pressure, without any reference to thermodynamics. By the principles of thermodynamics however this can be identified with the energy equivalent of the drop of total heat.

The principle of compounding finds a simple mathematical expression if we consider the case of a turbine consisting of a large number of simple turbines with a constant velocity ratio. For just as $\sum V^2/2g$ for such a series is equal to $-\int vdp$, so $\sum u^2/2g$ where u is the blade speed, must be proportional to this value of the available energy. The value of $\sum u^2$ or of $\sum d^2R^2$ where d is the mean diameter of the blade annulus, and R the revolutions per minute, is thus a suitable measure of the degree of compounding, and has been adopted as such from the time of the earliest compound turbine.

The high efficiency of which the steam turbine is capable, the large pressure and temperature range which it can utilize without mechanical difficulty, and the fact that it can be built for very large power output from one unit, have made it the most suitable prime mover for electric generating stations.

Even with extensive compounding for the sake of efficiency, the turbine is essentially a high speed engine, and the alternating current dynamo to which it is coupled has consequently required great modification into a high speed type capable of acting efficiently as the consort of the turbine. Fortunately the relation between the output and speed of an alternator follows nearly the same law (the inverse square) as the steam turbine, and it has always been found possible to keep pace with the development of the turbine prime mover in providing a suitable alternator for it to drive. The combination of turbine and alternator thus rapidly progressed in size to meet the demand for large units, and in all large power stations using fuel, steam turbines now provide the motive power in sizes ranging up to 60,000 K.W. Some of these stations now have a total capacity of supply of 500,000 K.W. feeding into wide-spread distribution systems and covering large areas.

In marine work there has been a similarly rapid progress both in economy and capacity, the marine turbine having advanced in output during 20 years development from the 2000 S.H.P. for the "Turbinia" to the 150,000 H.P. of H.M.S. "Hood".

The first marine turbines were directly coupled to the propeller shafting. Under these conditions on account of the low revolutions, the efficiency was determined by the maximum value of K or $\sum d^2R^2$ that could be provided on a given weight, and in a given space. The value of K necessary for maximum efficiency would have led to sizes of turbine and weights that would have been prohibitive. The association of the turbine with a low speed propeller put the marine turbine at a disadvantage so long as it remained directly coupled, and its early development was therefore confined to the propulsion of high speed vessels such as warships, liners and channel steamers, in which it soon established its superiority over the reciprocating engine. All marine turbines built in recent years are connected to their propeller through gearing, and this limitation has consequently been removed. The turbine has now been applied for the propulsion of all classes of sea-going vessels and considerable improvement in economy has been effected by the increased efficiencies of both turbines and propeller when these are allowed each to rotate at its most efficient speed.

The high speed turbines can utilize higher expansion ratios and have a higher efficiency of conversion, and the general result is that steam consumptions have been reduced to less than one half of that of the early direct coupled turbines.

Whilst mechanical gearing was introduced in the first place with the object of making the turbine applicable to low speed vessels, it was soon found to be of value even in higher speed vessels, and the direct coupled turbine may now be said to be completely superseded by the geared turbine in all classes.

Mechanical gearing has also assisted in the extension of the field of application of turbines on land. Continuous current generators which have been made in sizes up to 3000 K.W. at moderate speeds can now be driven by efficient high speed turbines, as also can low frequency alternators and alternators of small or moderate output. The geared turbine has also been used to a considerable extent for the driving of mills, such as textile, paper and jute mills.

It is estimated that the total output of steam turbines on land and sea has now reached a total of over 120,000,000 H.P.

In considering the steam turbine alone, the efficiency is expressed as the proportion of the available energy of the steam which is converted into useful work, and in large turbines this proportion is now in the neighbourhood of 85%. The ultimate measure of the efficiency of the heat engine, of which the heating apparatus as well as the turbine must be considered a part, is the ratio of the useful output to the heat supplied by the fuel. It is necessary therefore to include in the assessment the efficiency of heat transmission of the boiler, the proportion of the heat transmitted to the steam that becomes available for conversion into work, as well as the efficiency of this conversion in the turbine.

The development of the turbine has therefore brought us to the point where it becomes essential to examine closely all practical means of increasing the efficiency of the thermodynamic steam cycle which alone determines the proportion of the heat supplied to the steam that is available for conversion.

It is known from the laws of thermodynamics that the thermal efficiency of a heat engine depends upon its temperature range, in other words, that in accordance with Carnot's thesis, the temperature at which heat is supplied and that at which it is withdrawn should be as widely separated as possible.

There are two cycles of operation well known in the thermodynamic theory of heat engines. The first is the ideal reversible cycle of Carnot, consisting of two isothermals traversed by two adiabatics. The heat received and heat rejected are respectively proportional to the absolute temperatures and the efficiency of such a cycle therefore $= \frac{T_1 - T_2}{T_1}$. A serious drawback to this cycle is the large proportion of negative work involved in the adiabatic compression.

The other cycle is that which is most closely followed in all practical steam engines, and is known as the Rankine steam cycle. The heat received is received at constant pressure, and during the process of evaporation this is isothermal; but for any superheat subsequent to evaporation the temperature rises. The heat rejected is also rejected at constant pressure and in general at constant temperature, viz., the temperature of condensation. The expansion from the high temperature to the lower is adiabatic, and effected in an engine separate from both boiler and condenser. The cycle is completed by the return of the condensed fluid to the boiler by the feed pump.

It will be readily seen that for saturated steam, the Rankine cycle has this advantage over the Carnot cycle, that the negative work is reduced to the almost negligible amount necessary to return the feed water to the boiler. The practical superiority of the Rankine steam cycle and the ability of the steam turbine to work on this cycle is the explanation of the success of steam as a working fluid, in spite of the comparatively low temperature at which heat is received. The sole defect of the Rankine cycle for saturated steam is the necessity to reheat the feed water to the boiler temperature, since this heat is added at temperatures below the maximum.

It is clear that this defect could be overcome if the heating of the feed could be accomplished by transfer of heat from the steam at corresponding stages of equal temperatures, or in other words, by a regenerative process. With the addition of such a process the cycle would be thermodynamically reversible and under such conditions the efficiency of the Rankine cycle for saturated steam would be brought up to that of the Carnot cycle. In practice a close approximation to this regenerative process can be obtained by the employment of a sufficient number of feed water heaters in cascade supplied with steam tapped off from suitable stages of the turbine*. The steam which thus

*Feed heating in a single stage by partly expanded steam is a well-known expedient. It was proposed by James Weir in 1876 and by Normand in 1889; and feed heating in progressive stages was proposed by Ferranti in 1906.

transfers its heat to the feed heater is first of all made to do some work by expansion in the turbine down to the temperature at which it is required for withdrawal to the corresponding heaters, and since a certain amount of heat is required for the feed in any case, the work obtained from this tapped off steam is obtained merely at the expense of additional heat equal to the work done by it, in other words this additional heat is utilized at nearly 100% efficiency. Expressed in another way, the utilization of some of the heat of the steam to preheat the feed water reduces the amount of heat that has to be rejected to the condenser.

A large number of such stages of feed heating introduces some complications in the pipe work involved, but even with two or three stages very considerable heat economy can be effected. Generally speaking, economizers are used for the final heating of the feed water, but there is a tendency to develop the cascade system further so as to carry the steam heating of the feed water to the highest possible temperature. The economizer being thus replaced the residual heat of the flue gases of the boiler plant can be most advantageously utilized for regenerative pre-heating of the incoming air for the furnace.

With perfect regenerative feed heating the theoretical efficiency of the Rankine cycle for saturated steam can thus be brought up to that of the ideal Carnot cycle, viz., $\frac{T_1 - T_2}{T_1}$ which is the maximum thermodynamically possible between the temperature limits T_1 and T_2. Further increase of efficiency can only be sought in an extension of these limits.

Progress in the manufacture of materials capable of withstanding high temperatures now makes available for the steam turbine a maximum temperature of about 750° F. Unfortunately increase of pressure has not as yet kept pace with increase of temperature to this value, and it is not possible to continue to employ the saturated steam cycle receiving all its heat at this higher temperature. Since however in a superheated steam cycle the heat that is absorbed during the process of evaporation, viz., the latent heat of the steam, is a large proportion of the total heat transmitted by the boiler, some improvement in efficiency is obtained by increasing the boiler pressure, which by increasing the temperature of evaporation, increases the mean temperature of heat reception of the whole cycle. Thus, if we assume that by the use of feed heaters in cascade as just mentioned the water is fed into the boiler at the boiling temperature, with a boiler pressure of 250 lbs. per square inch, and a maximum temperature of 750° F. after superheating, the mean temperature of heat reception is about 430° F. If however the boiler pressure is increased to say 2000 lbs. per square inch with the same upper limit of temperature as before, owing to the fact that the latent heat is now received at a higher evaporation temperature, the mean temperature of reception is considerably increased, viz., to 680° F.

When adopting a high boiler pressure such as that just mentioned, it appears advisable to have a small auxiliary high pressure turbine having several stages of small diameter running at high revolutions and exhausting into the turbines at what is now ordinary boiler pressure. Such high pressure turbines are actually being built, one in Great Britain for 1,500 lbs. per square inch and one in the United States for 1250 lbs. per square inch.

Even with such pressures a maximum temperature of 750° F. still allows further heat to be added thereby superheating the steam, and it is not therefore possible in this cycle quite to realize the full efficiency corresponding to a maximum temperature of 750°. But a further improvement would result if after superheat to the maximum temperature, further heat was added so as to maintain isothermal expansion at this maximum temperature throughout the initial stages of the turbine, carrying it to such a point that subsequent adiabatic expansion in the turbine to the condenser pressure would leave the steam just saturated at that pressure. Such isothermal expansion of gaseous steam is hardly practicable in a separated engine, but as a practical approximation to it, there is being adopted in some installations at the present time a process of reheating the steam, which, after a certain amount of expansion in the turbine is extracted from the turbine, raised again to a high temperature in the reheater and led back to the turbine to continue its expansion. Such additional heat is added at a temperature higher than what would otherwise have been the mean temperature of heat reception for the working fluid.

Reheating has the additional and important advantage of keeping the steam longer in a superheated condition, and therefore diminishing the range where moisture, and consequent loss by water resistance, occur.

With regard to the lower temperature limit, it is evident that any reduction of the final temperature T_2 will increase the value of $\frac{T_1 - T_2}{T_1}$ for the theoretical maximum efficiency. This for a steam engine means a higher condenser vacuum, and the utilization of high vacuum becomes a most important factor in the design of the turbine, since it involves the necessity of providing ample blade area and passage way for the great volume of low pressure steam. The volumetric expansion of the steam in its passage through a turbine is provided for partly by increase in the area of the blade passages and partly by increase in velocity. With the high peripheral speeds adopted in most low pressure blading at the present time, whilst the steam speed may be such as to give a perfectly satisfactory velocity ratio, it may entail a loss of kinetic energy of considerable amount in the steam leaving the terminal blades.

A simple solution for the exhaust area problem in large machines is to provide a separate low speed turbine of large diameter for the final stages of the expansion arranged in close proximity to the condenser.

Application of all the foregoing principles is being made in a turbine plant of 50,000 K.W. capacity which has been built at Newcastle-on-Tyne for the new Crawford Avenue Power Station, Chicago. This plant has been described and illustrated in the contemporary press to which I may refer those desirous of studying the technical details. It is perhaps interesting to add that a thermal efficiency from fuel to electricity is anticipated about equal to that of the best internal combustion engine.

An appendix is added giving the estimated consumption of fuel oil per brake horse power burnt under the boilers at various boiler pressures up to 2000 lbs. per square inch.

Applying the same principles to marine work, it is interesting to enquire what results can be obtained using high temperature and boiler pressure, stage feed heating and air preheating but without intermediate re-heat*.

In such an arrangement it is desirable in order to obtain good efficiency in the high pressure range to adopt a separate fast running high pressure turbine. The three turbine arrangement, consisting of high pressure, intermediate and low pressure turbines geared to the propeller shaft, is specially applicable under such conditions. Assuming such a design to be adopted for an installation of about 5000 S.H.P., with boiler pressure 500 lbs., temperature 700° F., feed heating to 350° F., and effective air preheating to give a boiler efficiency of about 84%, a careful estimate based on known performance shows that owing to the increased available energy of the steam, a fuel consumption of about .57 lb. per S.H.P. would be obtained for turbines only, and .69 lb. per S.H.P. including auxiliary machinery.

When the difference in price between fuel oil and Diesel oil is taken into consideration, these being in about the ratio of 3 to 4, and also allowance made for the cost of lubricating oil for the Diesel Engine in excess of that for the turbine, it will be seen that these figures are equivalent to about .38 lb. of Diesel oil per S.H.P. for turbines only, and .475 lb. Diesel oil per S.H.P. for all purposes, including auxiliary machinery, and therefore promise an economy of fuel superior to that obtained with Diesel engines; whilst the first cost of such an installation will be little if any greater than that of a geared turbine installation of the usual design with low pressures and moderate superheat, and considerably less than the cost of an oil engine.

By practical application of the principles discussed above to the expansion of steam in a turbine, overall thermal efficiencies can be realized which are not inferior to those of internal combustion engines.

It has been further suggested to employ a binary vapour process utilizing another fluid with a lower vapour pressure such as mercury for the higher range of temperature. This would enable the whole of the heat to be received at the maximum temperature without employment of high pressures and theoretically a slight gain would result. Within the limits of temperature, however, which present day materials make practicable, it would appear doubtful whether the losses incurred through the double transmission of heat in such a process would not outweigh the theoretical gain, and moreover there are obvious objections to the use of such a fluid as mercury.

*The adoption of reheat after partial expansion would involve carrying a steam pipe of large capacity twice through the bulkhead separating engine room from boiler room, and this though not impossible is open to objection.

TABLE SHOWING THE OVERALL THERMAL EFFICIENCIES WHICH IT IS ESTIMATED COULD BE REALIZED WITH INCREASED BOILER PRESSURES UP TO 2,000 LBS. PER SQUARE INCH
(BASED ON THE CALLENDAR TABLES AND FORMULAE FOR THE PROPERTIES OF STEAM)

Column No.	1	2	3	4	5	6	7	8	9	10	11	12	13
Case No.	S.V.P. lbs. Gauge	S.V.T. °F.	°F. Initial Super-heat	Assumed Re-heat Pressure lbs. Abs.	Re-heat Temp. ° F.	Restored Super-heat °F.	Exhaust Vacuum Ins. Hg. Bar. 30.0″	Stage Feed Water Heating from 65° F. up to (° F.)	Thermal Efficiency (from Steam to Electricity)	Assumed Boiler Plant Efficiency (including all Aux.)	Overall Thermal Efficiency (from Fuel to Electricity)	Per Cent. Reduction in Fuel Con-sumption	Equiva-lent lbs. Oil B.H.P. Hr.
I	250	750°	344°	65	700°	402°	29.25″	360°	31.6	84.0	26.52		0.501
II	500	750°	281°	100	700°	372°	29.25″	420°	33.5	83.5	27.97	5.45	0.475
III	1,000	750°	204°	150	700°	343°	29.25″	510°	35.2	83.0	29.20	10.08	0.455
IV	1,500	750°	153°	250	700°	299°	29.25″	550°	36.6	82.5	30.20	13.82	0.440
V	2,000	750°	114°	400	700°	254.5°	29.25″	600°	37.5	82.0	30.75	15.90	0.432

NOTES:—S.V.P. = Stop Valve Pressure.
S.V.T. = Stop Valve Temperature.
Feed Water Heated to within about 50° F. of Boiler Temperature by Means of Steam Withdrawn from Turbine.
Blading at Several Points.
Boilers Equipped with Economizers or Air Pre-Heaters or both, so as to Recover Heat from the Flue Gases and Maintain High Thermal Efficiency.
Dynamo Efficiency Taken as 96.5 Per Cent. at Full Rated Output.
Gross Calorific Value of Fuel Oil Taken as 18,500 $\frac{\text{BThU}}{\text{lb.}}$.

THE USE OF EXPONENTIALS IN THE ANALYSIS OF MACHINE MOTIONS

By Professor Patrick Cormack,
College of Science for Ireland, Dublin, Irish Free State.

1. The purpose of this paper is to illustrate the convenience of exponential forms in the study of machine motions. The illustrative examples here taken are the motions of the piston and connecting rod of both ordinary and offset-cylinder engines. Reference, however, may be made to a paper by the author on *The Harmonic Analysis of Motion transmitted by Hooke's Joint*—Philosophical Magazine, July 1922. The harmonic analysis of the ordinary engine piston motion was first given by J. H. Macalpine (Engineering, 1897), the harmonic coefficients being given as infinite power series in ρ, where ρ is the ratio of crank to connecting rod. As here developed the coefficients are given as power series in h, where h is the tangent of half the angle whose sine is ρ. The great advantage of this for computation lies in the fact that the harmonic coefficients are now given by infinite power series which converge much more rapidly than those previously found.

2. Expanding $(1+h^2x)^n$ and $(1+h^2x^{-1})^n$ by the Binomial Theorem and forming the product of the expansions we obtain

$$(1+h^2x)^n(1+h^2x^{-1})^n=a_0+a_1(x+x^{-1})+a_2(x^2+x^{-2})+a_3(x^3+x^{-3})+\ldots, \tag{1}$$

where

$$\begin{aligned} a_0&=c_0^2+c_1^2h^4+c_2^2h^8+\ldots,\\ a_1&=c_0c_1h^2+c_1c_2h^6+c_2c_3h^{10}+\ldots,\\ a_2&=c_0c_2h^4+c_1c_3h^8+c_2c_4h^{12}+\ldots,\\ a_3&=c_0c_3h^6+c_1c_4h^{10}+c_2c_5h^{14}+\ldots,\\ &\ldots\ldots\ldots\ldots\ldots\ldots \end{aligned} \tag{2}$$

and where $c_0, c_1, c_2, \ldots$, are the Binomial Coefficients $1, n, \frac{1}{2}n(n-1), \ldots\ldots$ A relation among the coefficients $a_1, a_2, \ldots$ may be established by equating the coefficients of corresponding powers of x in the expansion of $(1+h^2x)\,(1+h^2x^{-1})$ times the differential coefficient of (1), and in the expansion of $n(h^2-h^2x^{-2})$ times (1). The relation found is

$$a_{r+1}=-\frac{(h^2+h^{-2})ra_r+(r-n-1)a_{r-1}}{n+r+1}. \tag{3}$$

3. In the ordinary crank and connecting rod engine (Fig. 1) we have

$$x=r\{\cos\theta+\rho^{-1}(1-\rho^2\sin^2\theta)^{\frac{1}{2}}\}$$

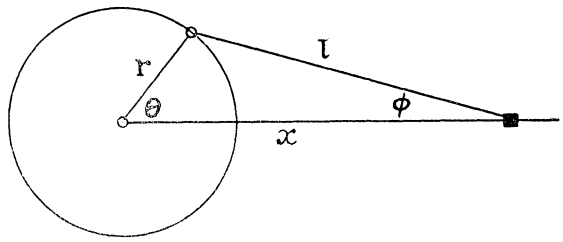


Fig. 1

where $\rho=r/l$. Also

$$(1+h^2)^{-2}(1+h^2e^{2i\theta})(1+h^2e^{-2i\theta})=1-\frac{4h^2}{(1+h^2)^2}\sin^2\theta=1-\rho^2\sin^2\theta,$$

if $\rho=2h/(1+h^2)$, or in other words if $h=\tan\frac{1}{2}\alpha$ where $\sin\alpha=\rho$. With this value for h we have therefore

$$\begin{aligned}\rho^{-1}(1-\rho^2\sin^2\theta)^{\frac{1}{2}}&=\tfrac{1}{2}h^{-1}(1+h^2e^{2i\theta})^{\frac{1}{2}}(1+h^2e^{-2i\theta})^{\frac{1}{2}}\\&=\tfrac{1}{2}h^{-1}\{a_0+a_1(e^{2i\theta}+e^{-2i\theta})+a_2(e^{4i\theta}+e^{-4i\theta})+a_3(e^{6i\theta}+e^{-6i\theta})+\ldots\}\\&=h^{-1}(\tfrac{1}{2}a_0+a_1\cos 2\theta+a_2\cos 4\theta+a_3\cos 6\theta+\ldots),\end{aligned}$$

the expansion being obtained by writing $e^{2i\theta}$ for x in (1). The coefficients a_0, a_1, a_2,...are determined from (2) by putting $n=\frac{1}{2}$ in the formulae for the binomial coefficients c_0, c_1, c_2, ... Thus

$$\begin{aligned}a_0&=1+\frac{1}{4}h^4+\frac{1}{64}h^8+\frac{1}{256}h^{12}+\ldots,\\a_1&=\frac{1}{2}h^2-\frac{1}{16}h^6-\frac{1}{128}h^{10}-\ldots,\\a_2&=-\frac{1}{8}h^4+\frac{1}{32}h^8+\frac{5}{1024}h^{12}+\ldots,\\a_3&=\frac{1}{16}h^6-\frac{5}{256}h^{10}-\ldots,\\&\ldots\ldots\ldots\ldots\ldots\ldots\end{aligned}\tag{4}$$

Hence we have the piston displacement x given by

$$x=r\,(\cos\theta+p_0+p_1\cos 2\theta+p_2\cos 4\theta+p_3\cos 6\theta+\ldots),$$

where

(5)
$$p_0=\frac{1}{2}h^{-1}+\frac{1}{8}h^3+\frac{1}{128}h^7+\ldots,$$
$$p_1=\frac{1}{2}h-\frac{1}{16}h^5-\frac{1}{128}h^9-\ldots,$$
$$p_2=-\frac{1}{8}h^3+\frac{1}{32}h^7+\frac{5}{1024}h^{11}+\ldots,$$
$$p_3=\frac{1}{16}h^5-\frac{5}{256}h^9-\ldots.$$
$$\ldots\quad\ldots$$

From (3) the relation among the coefficients $p_1, p_2, \ldots$ is

$$p_{r+1}=\frac{2(h^2+h^{-2})rp_r+(2r-3)p_{r-1}}{2r+3}.$$

The series for the harmonic coefficients $p_0, p_1, p_2, \ldots$ converge very rapidly. In fact for the usual values of the ratio of crank to connecting rod, the first term gives the harmonic coefficient correct to five places of decimals.

The piston velocity may be obtained by differentiating term by term the series for x, while a second differentiation with respect to t gives the piston acceleration. In the case of a rotating cylinder engine, when the cylinder makes an angle θ with the (fixed) crank, the piston displacement is x along the centre line of the cylinder, and it is therefore given by the vector

$$z=xe^{i\theta}.$$

The piston velocity is, by differentiation

$$\dot{z}=\dot{x}e^{i\theta}+xie^{i\theta}\dot{\theta}$$

and the piston acceleration is

$$\ddot{z}=e^{i\theta}(\ddot{x}+2i\dot{\theta}\dot{x}-x\dot{\theta}^2+ix\ddot{\theta}).$$

We may note that the piston acceleration may be analysed as compounded of its acceleration relative to the cylinder, the acceleration of the cylinder, and the acceleration given by twice the product of the angular velocity of the cylinder into the velocity of the piston relative to the cylinder, the direction of this last component being that of the relative velocity turned through a right angle in the direction of the angular velocity.

4. To find the angular velocity of the connecting rod, we have from Figure 1

$$r\sin\theta=l\sin\phi;$$

hence

$$r\cos\theta=l\cos\phi\frac{d\phi}{d\theta},$$

therefore

$$\frac{d\phi}{d\theta}=r\cos\theta/l\cos\phi=\rho\cos\theta(1-\rho^2\sin^2\theta)^{-\frac{1}{2}}.$$

The expansion of this can now be written down from the work of the previous section. We have

$$\begin{aligned}\frac{d\phi}{d\theta}&=2h\cos\theta\{a_0+a_1(e^{2i\theta}+e^{-2i\theta})+a_2(e^{4i\theta}+e^{-4i\theta})+\dots\}\\&=h(e^{i\theta}+e^{-i\theta})\{a_0+a_1(e^{2i\theta}+e^{-2i\theta})+a_2(e^{4i\theta}+e^{-4i\theta})+\dots\}\\&=h\{(a_0+a_1)(e^{i\theta}+e^{-i\theta})+(a_1+a_2)(e^{3i\theta}+e^{-3i\theta})+(a_2+a_3)(e^{5i\theta}+e^{-5i\theta})+\dots\}.\end{aligned}$$

To determine $a_0, a_1, a_2, \dots$ we use $n=-\frac{1}{2}$ in calculating the binomial coefficients $c_0, c_1, c_2, \dots$ which occur in table (2). We find

$$(6)\qquad \begin{aligned}a_0&=1+\frac{1}{4}h^4+\frac{9}{64}h^8+\dots,\\a_1&=-\frac{1}{2}h^2-\frac{3}{16}h^6-\frac{15}{128}h^{10}-\dots,\\a_2&=\frac{3}{8}h^4+\frac{5}{32}h^8+\frac{105}{1024}h^{12}+\dots,\\a_3&=-\frac{5}{16}h^6-\frac{35}{256}h^{10}-\dots,\\&\dots\dots\dots\dots\dots\dots\end{aligned}$$

Hence we have

$$\frac{d\phi}{d\theta}=q_1\cos\theta+q_3\cos3\theta+q_5\cos5\theta+\dots,$$

where

$$\begin{aligned}q_1&=2h-h^3+\frac{1}{2}h^5-\frac{3}{8}h^7+\dots,\\q_3&=-h^3+\frac{3}{4}h^5-\frac{3}{8}h^7+\frac{5}{16}h^9-\dots,\\q_5&=\frac{3}{4}h^5-\frac{5}{8}h^7+\frac{5}{16}h^9-\dots,\\&\dots\dots\dots\dots\dots\dots\end{aligned}$$

5. A series for the square of the angular velocity of the rod is readily found:

$$\left(\frac{d\phi}{d\theta}\right)^2=\rho^2\cos^2\theta(1-\rho^2\sin^2\theta)^{-1}=\frac{\rho^2\cos^2\theta\,(1+h^2)^2}{(1+h^2e^{2i\theta})(1+h^2e^{-2i\theta})}=\frac{h^2(e^{2i\theta}+e^{-2i\theta}+2)}{(1+h^2e^{2i\theta})(1+h^2e^{-2i\theta})}$$

$$=\frac{h^2-1}{h^2+1}\left\{\frac{1}{1+h^2e^{2i\theta}}+\frac{1}{1+h^2e^{-2i\theta}}\right\}+\frac{2}{h^2+1}$$

$$=\frac{2}{h^2+1}+\frac{h^2-1}{h^2+1}\,(1-h^2e^{2i\theta}+h^4e^{4i\theta}-\ldots+1-h^2e^{-2i\theta}+h^4e^{-4i\theta}-\ldots)$$

$$=\frac{2}{h^2+1}+\frac{2h^2-2}{h^2+1}(1-h^2\cos 2\theta+h^4\cos 4\theta-h^6\cos 6\theta+\ldots).$$

6. For the offset engine (Fig. 2) we have

$$l\sin\phi=r\cos\theta-E,$$

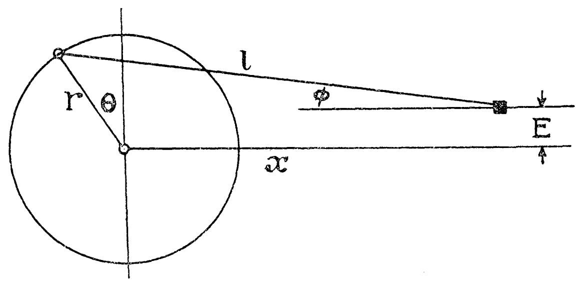


Fig. 2

therefore

$$\sin\phi=\rho\cos\theta-e,$$

where $\rho=r/l$, and $e=E/l$.

Also from the figure we have

$$x=l\cos\phi-r\sin\theta=r(\rho^{-1}\cos\phi-\sin\theta).$$

Now $\cos\phi=(1-\sin^2\phi)^{\frac{1}{2}}$, hence

$$\cos\phi=\left\{1-(\rho\cos\theta-e)^2\right\}^{\frac{1}{2}}=(1+\rho\cos\theta-e)^{\frac{1}{2}}(1-\rho\cos\theta+e)^{\frac{1}{2}}$$

$$=(1-e+\rho-2\rho\sin^2\tfrac{1}{2}\theta)^{\frac{1}{2}}(1+e-\rho+2\rho\sin^2\tfrac{1}{2}\theta)^{\frac{1}{2}}.$$

Take $u^2=2\rho/(1-e+\rho)$ and $v^2=2\rho/(1+e-\rho)$.

We have then

$$\rho^{-1}\cos\phi=2(uv)^{-1}(1-u^2\sin^2\tfrac{1}{2}\theta)^{\frac{1}{2}}(1+v^2\sin^2\tfrac{1}{2}\theta)^{\frac{1}{2}}$$

$$=2(uv)^{-1}(1+h^2)^{-1}(1+h^2e^{i\theta})^{\frac{1}{2}}(1+h^2e^{-i\theta})^{\frac{1}{2}}(1-h_1^2)^{-1}(1-h_1^2e^{i\theta})^{\frac{1}{2}}(1-h_1^2e^{-i\theta})^{\frac{1}{2}}$$

where

$$u=2h/(1+h^2),\text{ that is, } h=\tan\tfrac{1}{2}\alpha \text{ where } \sin\alpha=u$$

and where

$$v=2h_1/(1-h_1^2),\text{that is, } h_1=\tan\tfrac{1}{2}\beta \text{ where } \tan\beta=v.$$

Therefore

$$\rho^{-1}\cos\phi=(2hh_1)^{-1}(1+h^2e^{i\theta})^{\frac{1}{2}}(1+h^2e^{-i\theta})^{\frac{1}{2}}(1-h_1^2e^{i\theta})^{\frac{1}{2}}(1-h_1^2e^{-i\theta})^{\frac{1}{2}}$$

$$=(2hh_1)^{-1}[\{a_0+a_1(e^{i\theta}+e^{-i\theta})+a_2(e^{2i\theta}+e^{-2i\theta})+\dots\}$$
$$\{a_0'+a_1'(e^{i\theta}+e^{-i\theta})+a_2'(e^{2i\theta}+e^{-2i\theta})+\dots\}]$$

$$=(2hh_1)^{-1}\{a_0a_0'+a_1a_1'+a_2a_2'+\dots$$
$$+(a_0a_1'+a_1a_0'+a_1a_2'+a_2a_1'+a_2a_3'+a_3a_2'+\dots)(e^{i\theta}+e^{-i\theta})$$
$$+(a_0a_2'+a_2a_0'+a_1a_3'+a_3a_1'+a_2a_4'+a_4a_2'+\dots)(e^{2i\theta}+e^{-2i\theta})$$
$$+(a_0a_3'+a_3a_0'+a_1a_4'+a_4a_1'+a_2a_5'+a_5a_2'+\dots)(e^{3i\theta}+e^{-3i\theta})+\dots\}$$

$$=(2hh_1)^{-1}(b_0+2b_1\cos\theta+2b_2\cos 2\theta+2b_3\cos 3\theta+\dots),$$

in which b_n is written for $\Sigma a_r a_s'$ where $r-s=\pm n$. The quantities $a_0, a_1, \dots$ are given by table (4) and the same table also gives $a_0', a_1', \dots$, if h_1 be read for h. Hence

$$x=r\{-\sin\theta+(2hh_1)^{-1}(b_0+2b_1\cos\theta+2b_2\cos 2\theta+2b_3\cos 3\theta+\dots\}.$$

7. We shall conclude by giving the development for the angular velocity of the rod in the offset engine. We have

$$r\cos\theta=l\sin\phi+E,$$

hence

$$r\sin\theta=l\cos\phi\frac{d\phi}{d\theta},$$

therefore

$$\frac{d\phi}{d\theta}=\rho\sin\theta/\cos\phi.$$

From section 6 we have

$$\rho/\cos\phi=2hh_1(1+h^2e^{i\theta})^{-\frac{1}{2}}(1+h^2e^{-i\theta})^{-\frac{1}{2}}(1-h_1^2e^{i\theta})^{-\frac{1}{2}}(1-h_1^2e^{-i\theta})^{-\frac{1}{2}}$$
$$=2hh_1\{a_0+a_1(e^{i\theta}+e^{-i\theta})+a_2(e^{2i\theta}+e^{-2i\theta})+\dots\}\{a_0'+a_1'(e^{i\theta}+e^{-i\theta})+\dots\}.$$

Hence

$$\rho\sin\theta/\cos\phi=\frac{2hh_1}{2i}(e^{i\theta}-e^{-i\theta})\{b_0+b_1(e^{i\theta}+e^{-i\theta})+b_2(e^{2i\theta}+e^{-2i\theta})+\dots\};$$

therefore

$$\frac{d\phi}{d\theta}=2hh_1\{(b_0-b_2)\sin\theta+(b_1-b_3)\sin 2\theta+(b_2-b_4)\sin 3\theta+\dots\},$$

in which b_n is written for $\Sigma a_r a_s'$ where $r-s=\pm n$. Table 6 gives the quantities $a_0, a_1, a_2, \dots$ and by reading h_1 for h this table also gives the quantities $a_0', a_1', \dots$

SUR LES PERTES PAR FROTTEMENTS DANS LES MOTEURS A EXPLOSIONS

PAR M. ANDRÉ PLANIOL,
Docteur ès Sciences, Paris, France.

Les travaux déjà effectués sur les moteurs à explosions en vue de la détermination de leur rendement organique montrent que les pertes par frottement dépendent, pour une machine donnée, de la charge de cette machine, ou puissance développée par elle, de sa vitesse de rotation, de la température de ses différentes articulations et particulièrement de celle de la paroi du cylindre.

Toutefois, bien que ces points soient reconnus depuis longtemps, l'accord est loin d'être fait, surtout en ce qui concerne l'influence de la charge, puisque le sens même de la variation n'est pas unanimement reconnu par les différents auteurs qui se sont occupés de cette question. Quant aux autres causes de variation des pertes par frottement, les valeurs numériques trouvées pour leur influence sont toujours très discordantes.

Les recherches exposées ci-dessous ont été entreprises dans le but de contribuer à éclaircir cette question par une étude systématique.

I. VARIATION DES PERTES AVEC LA PUISSANCE DÉVELOPPÉE

Bien que certains auteurs aient mis en doute l'augmentation des pertes lorsque la charge de la machine s'accroit, la plupart a cependant admis cet accroissement comme probable théoriquement, sinon comme démontré par l'expérience.

Celle-ci donne, en effet, si l'on ne prend pas des précautions exceptionnelles en vue d'augmenter sa précision, des résultats extrêmement variables et incertains expliquant les divergences d'opinion des expérimentateurs.

De plus, la variation des pertes qui accompagne généralement celle de la puissance ne doit pas être considérée, en bonne logique, comme en résultant directement ainsi qu'on l'admet implicitement, en général.

Une étude complète du phénomène montre, en effet, que la cause réelle de cette variation des pertes tient à la variation des réactions moyennes exercées entre les différentes pièces mobiles de la machine qui sont soumises directement ou non aux effets moteurs.

On peut en effet concevoir certains régimes de fonctionnement pour lesquels une augmentation de puissance de la machine, correspondant à une diminution des réactions moyennes, amènerait une diminution des résistances passives.

Il est donc logique de rechercher une grandeur variant avec la puissance et liée le plus étroitement possible aux causes directes des pertes par frottement.

Des travaux faits avant la guerre par M. G. Lumet au Laboratoire de Mécanique Physique et Expérimentale de la Faculté des sciences à Paris, avaient montré la voie à suivre en introduisant la notion de pression moyenne en diagramme développé.

Il nous a suffi de compléter cette notion en y faisant intervenir celle des forces d'inertie des pièces à mouvement alternatif pour nous trouver en possession d'une variable éminemment propre à l'étude des pertes par frottement, parceque touchant directement aux forces de liaison, causes de ces pertes.

Enfin, l'application de certaines corrections fréquemment négligées, a permis de rendre les résultats entièrement cohérents.

Considérons, en effet, la valeur, pour chaque position de piston, de la pression relative p_g du gaz contenu dans le cylindre du moteur et construisons en le diagramme en portant en abscisses les courses successives du piston pendant un cycle, comptées toujours positivement sans tenir compte des changements de sens, et en ordonnées les pressions p_g.

Si nous calculons, pour la vitesse de rotation considérée, la valeur de la composante F_i des forces d'inertie des pièces à mouvement alternatif dirigée suivant l'axe du cylindre, le quotient de cette force par la section droite S du cylindre a les dimensions d'une pression, et nous l'appellerons pression fictive d'inertie $p_i = \dfrac{F_i}{S}$.

Reportant alors sur le diagramme précédemment tracé les valeurs successives de p_i pour les différentes positions de piston, on obtient une courbe représentative des pressions fictives d'inertie dont l'ordonnée peut, en tout point, être ajoutée à celle de la courbe p_g chacune des deux valeurs étant prise avec son signe.

La pression résultante $p = p_i + p_g$ représente l'effort transmis aux liaisons pour chaque unité de surface du piston; c'est donc elle qui détermine la grandeur des réactions exercées entre les différentes pièces de moteur.

La valeur moyenne de cette pression est, par suite, éminemment propre à la représentation des pertes par frottements qu'elle conditionne directement.

Ainsi donc, en prenant pour variable indépendante le volume v engendré par le piston en un point quelconque de sa course, l'expression

$$P = \frac{1}{v}\int_u^{v+u} |p|\, dv$$

donnera, pour chaque course du piston la valeur de la pression moyenne, et pour le cycle entier la pression moyenne au diagramme développé sera:

$$\pi = \frac{1}{n}\sum_1^n P,$$

n étant le nombre des temps du cycle étudié.

On est évidemment amené à prendre pour le calcul de P la valeur absolue de la pression résultante p puisque les frottements sont toujours résistants, quel que soit le sens de l'effort qui les engendre.

Les essais que nous avons effectués en vue de vérifier la valeur des considérations ci-dessus ont été faits au Laboratoire de Mécanique Physique et Expérimentale de la Faculté des Sciences sur un moteur Winterthür monocylindrique fonctionnant au gaz de ville: Alésage 290^{mm}, Course 430^{mm}, Puissance normale 30^{cv} à 200 t/m, Compression volumétrique 7,01, Cycle à 4 temps.

Faisant emploi de la variable π ci-dessus définie, et exprimant les pertes par frottement sous la forme du couple résistant sur l'arbre moteur qui leur est équivalent, on trouve entre ce couple résistant de frottement C_f et π une relation de forme linéaire qui est pour le moteur expérimental en question:

$$C_f = 16{,}9 + 3{,}9\pi.$$

Dans le but de généraliser les résultats et de permettre la comparaison entre différentes machines, on peut avantageusement faire intervenir la notion de couple moteur indiqué maximum qu'est susceptible de fournir la machine; la forme de la relation obtenue

$$\frac{C_f}{C_{im}} = \alpha + \beta\pi$$

reste la même, mais les coefficients α et β présentent alors le même degré de généralité que le rendement organique généralement utilisé dans l'étude de ces questions et qui, lui aussi, représente le quotient de deux couples.

Les résultats observés sur le moteur Winterthür se mettent alors sous la forme

$$\frac{C_f}{C_{im}} = 0{,}125 + 0{,}0285\pi.$$

La variable π employée présente de plus l'avantage de conserver une valeur définie lorsque la machine étudiée fonctionne à vide, entraînée par un moteur extérieur par exemple, ou pendant qu'elle s'arrête d'elle-même, alors que dans ces conditions, la variable puissance ne pourrait plus être employée à la représentation des pertes puisqu'elle ne présente plus de signification.

Il est d'ailleurs facile de transformer l'expression des résistances en fonction de la pression moyenne au diagramme développé en une expression en fonction de la puissance en déterminant la relation existant entre ces deux quantités lorsque la puissance est définie.

Méthodes employées pour ces mesures

1° *Indicateur.*

L'indicateur de Watt dont l'emploi donne déjà rarement des mesures précises sur les machines à vapeur nécessite pour fonctionner convenablement sur les moteurs à explosions des précautions toutes spéciales.

En effet, les conditions d'emploi y sont beaucoup plus difficiles, en raison, de la vitesse de rotation plus grande, de la variation de pression plus rapide, et du fait que le cycle comprend, en général, 4 temps au lieu de 2.

Il importe de soigner tout particulièrement la commande du mouvement de l'indicateur et, plus encore peut-être, la communication de cet appareil avec le cylindre.

Nous avons, en effet, constaté expérimentalement des différences de 9 à 10% entre les puissances indiquées par deux appareils enregistrant simultanément le même cycle ou le même groupe de cycles sur deux orifices différents de communication avec le cylindre, suivant la forme des canaux faisant suite à ces orifices.

Seuls les canaux présentant des résistances à l'écoulement très réduites donnent des résultats exacts et nous avons éliminé toutes les autres formes pour nos recherches.

Enfin, les mesures à l'indicateur sont entachées d'erreurs accidentelles qu'il a été possible d'étudier en quelque sorte statistiquement, grâce au grand nombre de diagrammes relevés et qu'on peut répartir entre les catégories suivantes:

(*a*) Erreur due au planimètre: valeur probable relative 0,33%,

(*b*) Erreur due à l'indicateur lui-même: valeur probable relative 0,21%,

(*c*) Erreur due aux ressorts d'indicateur: valeur probable relative 0,78%,

les valeurs ci-dessus étant relatives à une mesure unique.

Mais une détermination se faisant au moyen de plusieurs mesures, on est amené à fixer la valeur de l'erreur relative probable sur un essai à 0,5% environ dans les conditions pratiques où ont été faites nos expériences.

Cette précision a permis de déterminer dans des conditions de sécurité satisfaisantes les relations précédemment exposées.

2° *Méthode cinétique.*

En lançant la machine à une vitesse supérieure à sa vitesse de régime, et en supprimant brusquement l'action motrice la vitesse de rotation diminue sous l'action des frottements, la connaissance de cette vitesse en fonction du temps permet donc de les calculer.

Soient en effet I le moment d'inertie total de la machine, ω la vitesse angulaire à l'instant t, C le couple résistant total à cet instant, on a évidemment:

$$C = I\frac{d\omega}{dt}.$$

Mais il y a lieu de remarquer qu'ainsi appliquée, la méthode est incomplète. En effet, lorsque le moteur s'arrête de cette façon, il continue à aspirer, comprimer, détendre et évacuer de l'air, et le travail correspondant à ce fonctionnement n'est généralement pas nul.

Il y a donc lieu de tenir compte du couple correspondant, généralement résistant, et c'est ce que nous avons fait en relevant les diagrammes à différentes vitesses de rotation pour les régimes de marche étudiés.

Le couple correspondant aux résistances de transvasement de l'air dans le cylindre ainsi mesuré étant C_t la valeur du couple résistant dû aux seuls frottements mécaniques est

$$C_f = C - C_t.$$

3° *Méthode électrique.*

En entraînant le moteur au moyen d'une machine électrique convenablement étalonnée, on obtient une nouvelle mesure de C et par suite de C_t.

Les résultats fournis par les trois méthodes sont absolument comparables et aucune discordance systématique n'apparait entre eux.

Les mesures faites ainsi étant complètement indépendantes, on peut considérer les résultats obtenus comme présentant une sécurité réelle.

II. Variation des pertes avec la vitesse de rotation

Les trois méthodes précédentes permettent de mesurer les résistances aux différentes vitesses et par suite d'établir la relation existant entre les pertes et la vitesse de rotation.

Elles ont été appliquées aux régimes de marches suivants:

(*a*) Marche sans explosions avec admission ouverte en grand et compression maxima (Mesures électriques et cinétiques).

(*b*) Marche à vide avec explosions (Mesures à l'indicateur).

(*c*) Marche sans explosions avec admission fermée et compression minima (Mesures électriques et cinétiques).

Les courbes expérimentales obtenues sont si voisines de lignes droites que l'on est fondé, en considération des erreurs expérimentales inévitables, à les représenter par des équations linéaires dont les expressions sont respectivement, aux régimes ci-dessus:

$$\left.\begin{array}{l} C_A = 7{,}1 + 0{,}080n \\ C_B = 6{,}1 + 0{,}072n \\ C_C = 4{,}5 + 0{,}073n \end{array}\right\} \quad (C \text{ en kilogrammètres, } n \text{ en } t/m).$$

L'accroissement des résistances dans les limités de vitesse étudiées, soit de 0 à 240 t/m est très supérieur aux augmentations généralement admises, puisqu'ici les couples résistants sont en moyenne 4 fois plus forts qu'à une vitesse voisine de zéro.

Enfin, à petite comme à grande vitesse, la forme linéaire de la relation $C_f = \phi(\pi)$ se retrouve avec la même précision.

III. Variation des pertes avec la température du cylindre et la viscosité de l'huile

Les mesures faites pour l'étude de l'effet de la charge sur les pertes ayant montré la très grande influence de la température de la chemise d'eau sur les pertes et la nécessité de définir cette température avec beaucoup de précision, il est apparu nécessaire de construire un échangeur de chaleur à grande masse thermique. Cet appareil permet de faire circuler à travers le moteur, et en circuit fermé, un courant d'eau de grand débit dont la température, très peu différente entre l'entrée et la sortie, peut être élevée ou abaissée à volonté.

Nous avons alors mesuré par voie électrique et par voie cinétique les résistances du moteur pour différentes températures moyennes de l'eau de circulation 9°, 24°, 44°,5, 65°, 96° et pour des vitesses variées.

A toutes les allures, il se manifeste un accroissement considérable des résistances passives quand la température s'abaisse de 96° à 9°: les pertes par frottement s'accroissent alors dans l'ensemble, d'environ 1 à 4. Mais si l'on rapporte la variation observée, due seulement aux variations de résistance du piston, à la valeur initiale de celle-ci, l'accroissement correspondant à ce refroidissement de 87° est alors de 1 à 6.

Cette augmentation considérable est due à l'élévation de la viscosité de l'huile servant au graissage du cylindre lorsque sa température s'abaisse.

Pour mettre ce fait en évidence, nous avons étudié deux huiles présentant des viscosités aussi différentes que possible.

La première est l'huile de ricin (Castrol R) de viscosité considérable, dite huile R.

La seconde est l'huile minérale (Gargoyle Arctic) de viscosité très faible, dite huile A.

Ces deux huiles ont été tout d'abord étudiées au viscosimètre absolu Baume-Vigneron, et leurs courbes de viscosité tracées en fonction de la température: on constate alors que les deux courbes sont presque identiques et que la courbe R se déduit de la courbe A par une élévation de température de 30° environ, toutes deux s'élevant très vite vers les basses températures.

D'autre part, ces deux huiles ont été soumises à des essais systématiques sur le moteur en vue de déterminer, à différentes températures les couples résistants, lorsque le cylindre est graissé avec l'une ou l'autre de ces huiles, le graissage des autres parties du moteur restant le même.

On trouve alors comme représentation des pertes par frottement en fonction de la température deux courbes analogues aux courbes du coefficient de viscosité, c'est-à-dire pouvant se déduire l'une de l'autre par une translation parallèle à l'axe des températures, l'amplitude de ce décalage étant de 30° environ comme dans le cas des courbes de viscosité.

Ce fait remarquable montre que les résistances de frottement du piston dépendent seulement du coefficient de viscosité de l'huile employée au graissage et sont indépendantes de la nature de cette huile puisqu'il suffit de donner, par une variation de température convenable, le même coefficient de viscosité à deux huiles différentes, pour que le piston présente des résistances de frottement identiques.

L'expérience montre de plus que les variations de ces résistances en fonction de la température sont moins rapides que celles de la viscosité. Ce fait tient à l'existence de résistances extérieures au piston, d'une part, et, d'autre part, au fait que les résistances du piston lui-même ne proviennent pas exclusivement de frottements visqueux, mais aussi de frottements solides dont l'action vient troubler l'influence de la viscosité du lubrifiant.

Il en résulte qu'il n'est pas intéressant de diminuer à l'extrême cette viscosité, la résistance totale ne diminuant plus que très lentement pour les viscosités minimes: les lubrifiants très fluides présentent d'ailleurs de graves inconvénients au point de vue de la sécurité, car ils risquent d'être expulsés d'entre les surfaces frottantes et d'amener par suite le grippage de la machine.

IV. Répartition des pertes entre les différents organes du moteur

En vue d'étudier cette question, on a mesuré séparément les pertes des groupes d'organes suivants du moteur expérimental:

A. Arbre vilebrequin et volant.

B. Les mêmes que A, plus la distribution complète.

C. Les mêmes que B, plus la bielle et le piston dépourvu de segments.

D. Les mêmes que C, plus les segments du piston.

Les couples résistants, mesurés aux différentes vitesses de rotation réalisables, peuvent être exprimés en fonction de cette vitesse par les relations linéaires suivantes et cela avec une grande précision, le nombre des mesures faites étant considérable et les écarts faibles:

$$\left.\begin{array}{l} C_A = 0{,}5 \ \ +0{,}023n \\ C_B = 0{,}85+0{,}032n \\ C_C = 1{,}3 \ \ +0{,}056n \\ C_D = 2{,}05+0{,}079n \end{array}\right\} C \text{ en kilogrammètres, } n \text{ en tours minute.}$$

Les différences entre les couples successifs donnent la valeur des couples résistants des groupes d'organes ajoutés successivement, ainsi qu'on l'a dit, et qui sont:

$$\begin{array}{rll} C_A = & 0{,}5 \ \ +0{,}023n & \text{Volant et vilebrequin,} \\ C_B - C_A = & 0{,}35+0{,}009n & \text{Distribution,} \\ C_C - C_B = & 0{,}45+0{,}024n & \text{Bielle et piston,} \\ C_D - C_C = & 0{,}75+0{,}023n & \text{Segments du piston.} \end{array}$$

D'autre part, la résistance de frottement du piston dans le cylindre a été mesurée directement, dans un but de vérification, pour des vitesses de glissement de $0^{m/s},2$ à $1^{m/s}$, et les résultats obtenus ainsi ont été comparés au moyen d'une transformation convenable, avec les mesures précédentes: la concordance entre ces deux méthodes entièrement différentes a été bonne puisque l'écart maximum des résultats atteint à peine 5%.

On peut également tirer parti de ces mesures pour calculer le coefficient de frottement dans certaines des articulations principales du moteur, ce qui donne les résultats suivants:

Articulation	Charge par cm²	Vitesse de rotation	Vitesse de glissement	Coefficient de frottement
Paliers de l'arbre....	$3^{k/cm^2}, 7$	$200^{t/m}$	$1^{m/s},20$	$f<0{,}032$
"	"	20	0,12	$f=0{,}0076$
Piston seul.........	0,036	200	2,87	$f=0{,}27$
"	"	20	0,29	$f=0{,}046$
Segments du piston..	0,40	200	2,87	$f=0{,}12$
"	"	20	0,29	$f=0{,}028$

La valeur de $f=0{,}032$ obtenue pour les paliers à l'allure de $200^{t/m}$ est supérieure à la valeur exacte car la résistance de l'air au mouvement du volant, qu'il n'a pas été possible de mesurer séparément, est loin d'être négligeable à cette allure, tandis qu'elle l'est sensiblement à $20^{t/m}$.

On voit ainsi que l'on retrouve dans toutes les parties du moteur l'influence considérable de la vitesse de rotation sur les résistances par frottement, lesquelles sont en moyenne 4 ou 5 fois plus grandes dans les articulations étudiées à l'allure de $200^{t/m}$ qu'elles ne le sont à $20^{t/m}$.

Les formules généralement utilisées dans des cas semblables ne donnent que des accroissements beaucoup plus faible des frottements en fonction de la vitesse.

V. Comparaison des résultats obtenus sur différents moteurs

En vue de vérifier sur d'autres machines la généralité des conclusions tirées de l'étude du moteur expérimental Winterthür, nous avons entrepris des essais, forcément moins complets que ceux qu'il est possible de faire au Laboratoire, sur différents moteurs industriels définis ci-dessous:

(*a*) Mot. S.M.I.M. $60^{c.v.}$ à $200^{t/m}$ Monocylindrique 4 temps, gaz pauvre, Alés. $380^{mm.}$, Course $560^{mm.}$.

(*b*) Mot. S.M.I.M. $35^{c.v.}$ à $220^{t/m}$ Monocylindrique 4 temps, gaz pauvre, Alés. $320^{mm.}$, Course $380^{mm.}$.

(*c*) Mot. P.S. $3^{c.v.}$ à $4500^{t/m}$ Monocylindrique 2 temps, essence, Alés. $54^{mm.}$, Course $54^{mm.}$.

Nous comparerons les résultats obtenus avec ceux du moteur expérimental dont les caractéristiques ont été données et que nous appellerons *W*.

1° Variation des pertes avec la pression moyenne au diagramme développé

Les résistances des différents moteurs étudiés à ce point de vue, soient (*a*) et (*b*), peuvent être représentées, elles aussi, par des expressions linéaires:

$$(a) \qquad \frac{C_f}{C_{im}} = 0{,}135 + 0{,}0295\pi,$$

$$(b) \qquad \frac{C_f}{C_{im}} = 0{,}076 + 0{,}050\pi,$$

alors que le moteur Winterthür avait donné :

$$(W) \qquad \frac{C_f}{C_{im}} = 0,125 + 0,0285\pi.$$

Ces trois expressions représentent dans la région de fonctionnement des moteurs trois droites voisines; leur moyenne fournit une droite :

$$\frac{C_f}{C_{im}} = 0,112 + 0,036\pi,$$

qui peut être considérée, en tenant compte des erreurs expérimentales, comme donnant une réprésentation convenable des couples résistants des trois moteurs considérés, et montrent que les phénomènes observés dans le moteur expérimental se conservent sensiblement pour d'autres machines.

2° VARIATION DES PERTES AVEC LA VITESSE DE ROTATION

Les moteurs étudiés présentant des vitesses de rotation très différentes, il serait illogique de prendre comme variable cette vitesse n elle-même.

L'emploi du quotient $\frac{n}{N}$ de la vitesse de rotation par sa valeur de régime maximum N est donc tout indiqué.

Dans ces conditions, les résultats obtenus peuvent être mis sous la forme suivante :

$$\text{Moteurs } (W) \qquad \frac{C_f}{C_{im}} = 0,0525 + 0,118\,\frac{n}{N},$$

$$(a) \qquad = 0,051 \;\; + 0,151\,\frac{n}{N},$$

$$(b) \qquad = 0,035 \;\; + 0,115\,\frac{n}{N},$$

$$(c) \qquad = 0,095 \;\; + 0,088\,\frac{n}{N}.$$

On voit que, malgré les différences considérables existant entre les types de moteurs étudiés, les résultats obtenus restent cependant comparables entre eux.

Il est intéressant de constater ainsi que l'influence de la vitesse sur les pertes par frottement est sensiblement la même pour un moteur à essence à deux temps, de 3 chevaux pesant 3 kilos par cheval, monté sur billes et galets et muni d'un piston d'aluminium, que pour des gros moteurs à gaz à 4 temps montés sur paliers lisses et munis de pistons de fonte, et dont le poids moyen dépasse une centaine de kilos par cheval.

Enfin, nous ferons remarquer en terminant que l'intérêt de l'étude des pertes par frottement dans les machines motrices apparait à la réflexion plus grand qu'on ne serait tenté de le penser tout d'abord.

De fait, outre leur influence directe sur le rendement thermique utile d'un moteur déterminé $\eta_u = \frac{A.P_u}{Q}$, ces pertes jouent, dans l'évolution des machines motrices, un rôle très important en déterminant, au moins partiellement, les limites du progrès du rendement utile compatible avec le mode de construction à une époque donnée.

Le rendement utile η_u est en effet égal au produit du rendement thermique η_i par le rendement mécanique du moteur η_m, lequel est déterminé par les pertes par frottement.

Or, on sait qu'en règle générale, une amélioration du rendement thermique indiqué nécessite une augmentation du rapport des pressions maxima et moyenne du cycle utilisé, ce qui a pour effet d'amener dans une machine donnée une diminution correspondante du rendement organique. Cette diminution peut abaisser considérablement, et même annihiler, le gain réalisé sur le rendement thermique indiqué.

Pour pouvoir tirer parti pratiquement dudit gain, il est donc nécessaire, dans la plupart des cas, de réaliser en même temps, un accroissement du rendement organique de la machine, soit, par conséquent, une diminution de ses pertes par frottement: leur étude tire de ce fait un intérêt considérable qui suffirait, à lui seul, à justifier les efforts nécessaires pour la mener à bien.

BERECHNUNG DER AIRYSCHEN SPANNUNGSFUNKTION FÜR RECHTECKIGE SCHEIBEN

Von A. Frey Samsioe,
Zivilingenieur, Stockholm, Schweden.

Fouriersche Reihen haben in der mathematischen Physik häufig Anwendung gefunden; unter anderem sind solche Reihen, die nach ungeraden Vielfachen des Argumentes fortschreiten, angewendet worden, um die in einer längs ihren Rändern eingespannten rechteckigen *Platte* bei gewissen Belastungen auftretenden Spannungen und Formänderungen zu ermitteln*. In der vorliegenden Abhandlung wird der Nachweis erbracht, dass mit Hilfe einer solchen Reihenentwickelung auch die airysche Spannungsfunktion für eine rechteckige *Scheibe* berechnet werden kann.

In einer Scheibe mit der Dicke 1 können bekanntlich unter Voraussetzung eines ebenen Formänderungszustandes oder eines ebenen Spannungszustandes die Spannungen σ_ξ, σ_η und $\tau_{\xi\eta}$ als Ableitungen einer Funktion F, der airyschen Spannungsfunktion, so ausgedrückt werden, dass

$$\sigma_\xi = \frac{\partial^2 F}{\partial \eta^2}, \quad \sigma_\eta = \frac{\partial^2 F}{\partial \xi^2}, \quad \tau_{\xi\eta} = -\frac{\partial^2 F}{\partial \xi \partial \eta}.$$

Die Funktion F soll dann der Differentialgleichung

$$\frac{\partial^4 F}{\partial \xi^4} + 2\frac{\partial^4 F}{\partial \xi^2 \partial \eta^2} + \frac{\partial^4 F}{\partial \eta^4} = 0$$

und gewissen Randbedingungen, die von der äusseren Belastung abhängen, genügen.

Die Randbedingungen können geschrieben werden†:

$$\left(\frac{\partial F}{\partial \xi}\right)_s = -\int_0^s p_\eta ds + C_1, \quad \left(\frac{\partial F}{\partial \eta}\right)_s = \int_0^s p_\xi ds + C_2,$$

wo s den Abstand des betreffenden Punktes, dem Rande entlang von einem gewissen Ursprung aus gemessen, bezeichnet, und p_η und p_ξ die rechtwinkligen Komponenten der am Randelement ds angreifenden, auf die Längeneinheit bezogenen, äusseren Kräfte bezeichnen, bezw. C_1 und C_2 Konstanten sind, über welche beliebig verfügt werden kann.

*Vgl. u.a. Arpad Nadai: *Die Formänderungen und die Spannungen von rechteckigen elastischen Platten*, Berlin, 1915, und V. Lewe, *Die strenge Lösung des Pilzdeckenproblems*, Berlin, 1922.

†Vgl. z. B. A. und L. Föppl, *Zwang und Drang*, Berlin, 1920.

Es soll hier nur der Fall behandelt werden, dass die Scheibe mit den Rändern $\xi = \pm a$, $\eta = \pm b$ nur dem Rande $\eta = \pm b$ entlang mit einer senkrecht zum Rande stehenden Last p_η, die ausserdem, wie Abb. 1 zeigt, bezüglich den ξ-und η-Achsen

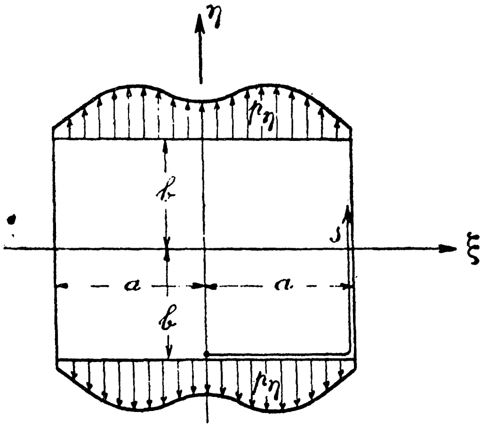


Abb. 1

symmetrisch sein soll, belastet ist. Wählen wir als Ursprung der Länge s, dem Rande entlang gemessen, den Punkt ($\xi = 0$, $\eta = -b$) und setzen $C_1 = C_2 = 0$, so erhalten wir

$$\left(\frac{\partial F}{\partial \xi}\right)_s = -\int_0^s p_\eta ds, \quad \left(\frac{\partial F}{\partial \eta}\right)_s = 0.$$

Setzen wir $p = p_\eta$ für $\eta = +b$, ergibt sich für den Rand vom Punkte $(0, -b)$ zum Punkte $(a, -b)$

$$\frac{\partial F}{\partial \xi} = \int_0^\xi p d\xi.$$

Setzen wir ferner $P = \int_0^a p d\xi$, so ergibt sich für den Rand vom Punkte $(a, -b)$ zum Punkte (a, b)

$$\frac{\partial F}{\partial \xi} = P,$$

für den Rand vom Punkte (a, b) zum Punkte $(-a, b)$

$$\frac{\partial F}{\partial \xi} = P + \int_a^\xi p d\xi = \int_0^\xi p d\xi,$$

für den Rand vom Punkte $(-a, b)$ zum Punkte $(-a, -b)$

$$\frac{\partial F}{\partial \xi} = -P,$$

und schliesslich für das übriggebliebene Randstück vom Punkte $(-a,-b)$ zum Punkte $(0,-b)$

$$\frac{\partial F}{\partial \xi}=-P+\int_{-a}^{\xi} p d \xi=\int_{0}^{\xi} p d \xi .$$

Dem ganzen Rande entlang soll

$$\frac{\partial F}{\partial \eta}=0$$

sein.

Durch diese Randbedingungen ist die Funktion F bis auf eine Konstante bestimmt. Wählen wir diese Konstante so, dass

1. für $\xi= \pm a$, $F=0$ sein soll,

so wird

2. für $\eta= \pm b$, $F=\int_{0}^{\xi} d \xi \int_{0}^{\xi} p d \xi-\int_{0}^{a} d \xi \int_{0}^{\xi} p d \xi .$

Weiter soll

3. $$\begin{cases}\text{für } \xi=a, & \dfrac{\partial F}{\partial \xi}=P, \\ \text{für } \xi=-a, & \dfrac{\partial F}{\partial \xi}=-P,\end{cases}$$

und

4. $$\text{für } \eta= \pm b, \quad \frac{\partial F}{\partial \eta}=0$$

sein.

Führen wir die lineare Substitution

(1) $$x=\frac{\pi}{2 a} \xi, \quad y=\frac{\pi}{2 b} \eta, \quad p(\xi)=\left(\frac{\pi}{2 a}\right)^{2} . \rho(x)$$

ein, und setzen $Q=\int_{0}^{\frac{\pi}{2}} \rho . d x$, so gehen die vorstehenden Randbedingungen in folgende über:

1. $$x= \pm \frac{\pi}{2}, \quad F=0,$$

2. $$y= \pm \frac{\pi}{2}, \quad F=\int_{0}^{x} d x \int_{0}^{x} \rho d x-\int_{0}^{\frac{\pi}{2}} d x \int_{0}^{x} \rho d x,$$

3. $$\begin{cases}x=\dfrac{\pi}{2}, & \dfrac{\partial F}{\partial x}=Q, \\ x=-\dfrac{\pi}{2}, & \dfrac{\partial F}{\partial x}=-Q,\end{cases}$$

4. $$y= \pm \frac{\pi}{2}, \quad \frac{\partial F}{d y}=0,$$

wo nun F als Funktion von x and y aufzufassen ist.

Die Differentialgleichung geht dann über in

(2) $$b^4 \frac{\partial^4 F}{\partial x^4} + 2a^2b^2 \frac{\partial^4 F}{\partial x^2 \partial y^2} + a^4 \frac{\partial^4 F}{\partial y^4} = 0.$$

Die Belastung ρ im Intervalle $-\frac{\pi}{2}$ bis $+\frac{\pi}{2}$ entwickeln wir in eine fouriersche Reihe, die nach ungeraden Vielfachen des Argumentes fortschreiten soll, indem wir eine gerade Funktion $f(x)$ so definieren, dass

$$0 \leq x < \frac{\pi}{2}, \quad f(x) = \rho(x),$$

$$\frac{\pi}{2} < x < \pi, \quad f(x) = -\rho(\pi - x)$$

und diese in eine fouriersche cosinus-Reihe im Intervalle 0 bis π entwickeln.

Wir erhalten dann

$$f(x) = \tfrac{1}{2} b_0 + \sum_n b_n \cos nx,$$

wo $b_0 = 0$ und

$$\begin{cases} b_n = 0 & \text{für gerade } n, \\ b_n = \dfrac{4}{\pi} \displaystyle\int_0^{\frac{\pi}{2}} \rho \cos nx \, dx & \text{für ungerade } n. \end{cases}$$

Demnach kann geschrieben werden

(3) $$0 \leq x < \frac{\pi}{2}, \quad \rho = \sum_{n=1,3,5,\ldots}^{\infty} b_n \cos nx,$$

wo b_n oben definiert ist und die Summe über alle ungerade Zahlen n zu erstrecken ist.

Weiter ergibt sich

$$\int_0^x \rho dx = \sum_n b_n \cdot \frac{1}{n} \cdot \sin nx,$$

$$\int_0^x dx \int_0^x \rho dx - \int_0^{\frac{\pi}{2}} dx \int_0^x \rho dx = -\sum_n b_n \cdot \frac{1}{n^2} \cos nx,$$

(4) $$Q = \int_0^{\frac{\pi}{2}} \rho dx = \sum_n b_n \frac{1}{n} \left(-1\right)^{\frac{n-1}{2}}.$$

Nun nehmen wir an, dass die airysche Spannungsfunktion in der Form

(5) $$F = \sum_n \psi_n(y) \cos nx + \sum_n \phi_n(x) \cos ny$$

geschrieben werden kann, wo die Summen über alle ungerade Zahlen erstreckt sind, wo die Glieder *jedes für sich* der Differentialgleichung (2) genügen sollen wo die durch die erste Summe dargestellte Reihe im Intervalle $0 \leq x < \frac{\pi}{2}$, $0 \leq y < \frac{\pi}{2}$ eine stetige Funktion darstellt, die für $x = \frac{\pi}{2} - 0$ verschwindet, und wo die durch die zweite Summe dargestellte Reihe in demselben Intervalle eine stetige Funktion darstellt, die für $y = \frac{\pi}{2} - 0$ verschwindet.

Zufolge der Symmetrie der Belastung sind offenbar $\psi_n(y)$ und $\phi_n(x)$ gerade Funktionen von y bezw. x und demnach

$$\psi_n(y) = \psi_n(-y), \ \phi_n(x) = \phi_n(-x).$$

Die Randbedingungen geben daher für das Intervall

$$0 \leq x < \frac{\pi}{2}, \ 0 \leq y < \frac{\pi}{2},$$

folgende vier Gleichungen:

1. $$\sum_n \phi_n\left(\frac{\pi}{2} - 0\right) \cos ny = 0,$$

2. $$\sum_n \psi_n\left(\frac{\pi}{2} - 0\right) \cos nx = -\sum_n b_n \frac{1}{n^2} \cos nx,$$

3. $$-\sum_n \psi_n(y) n(-1)^{\frac{n-1}{2}} + \sum_n \phi_n'\left(\frac{\pi}{2} - 0\right) \cos ny = Q,$$

4. $$\sum_n \psi_n'\left(\frac{\pi}{2} - 0\right) \cos nx - \sum_n \phi_n(x) n(-1)^{\frac{n-1}{2}} = 0.$$

Die Randbedingungen (1) und (2) geben, weil eine gerade Funktion im Intervalle 0 bis $\frac{\pi}{2}$ nur in einer einzigen Weise in eine fouriersche Reihe, die nach ungeraden Vielfachen des Argumentes fortschreitet, entwickelt werden kann,

(6) $$\phi_n\left(\frac{\pi}{2} - 0\right) = 0,$$

(7) $$\psi_n\left(\frac{\pi}{2} - 0\right) = -\frac{1}{n^2} b_n.$$

Sollen die Glieder $\psi_n(y) \cos nx$ jedes für sich der Differentialgleichung (2) genügen, so ergibt sich zur Bestimmung von $\psi_n(y)$ folgende Differentialgleichung

$$b^4 n^4 \psi - 2a^2 b^2 n^2 \frac{d^2\psi}{dy^2} + a^4 \frac{d^4\psi}{dy^4} = 0,$$

die folgendes vollständige Integral hat:

$$\psi_n(y) = Ae^{n\kappa y} + Be^{-n\kappa y} + Cye^{n\kappa y} + Dye^{-n\kappa y},$$

wo A, B, C und D willkürliche Konstanten sind und wo $\kappa = \frac{b}{a}$.

Nun soll jedoch

$$\psi_n(y)=\psi_n(-y),$$

und demnach muss

$$A=B,\ C=-D$$

sein.

Wir können deshalb schreiben

$$\psi_n(y)=K\cosh n\kappa y+Ly\sinh n\kappa y,$$

wo nun K und L willkürliche Konstanten sind und wo mit cosh und sinh die hyperbolischen Funktionen

$$\cosh\zeta=\tfrac{1}{2}(e^{\zeta}+e^{-\zeta}),\ \sinh\zeta=\tfrac{1}{2}(e^{\zeta}-e^{-\zeta})$$

bezeichnet sind.

In derselben Weise kann gezeigt werden, dass $\phi_n(x)$ in der Form

$$\phi_n(x)=M\cosh n\lambda x+Nx\sinh n\lambda x$$

geschrieben werden kann, wo $\lambda=\dfrac{a}{b}$.

Zufolge Gl. (6) soll

$$\phi_n\left(\frac{\pi}{2}-0\right)=0$$

sein. Demnach ergibt sich

$$N=-M\frac{2}{\pi}\frac{\cosh n\lambda\frac{\pi}{2}}{\sinh n\lambda\frac{\pi}{2}},$$

und wir erhalten

$$\phi_n(x)=M\left\{\cosh n\lambda x-\frac{2x}{\pi}\frac{\cosh n\lambda\frac{\pi}{2}}{\sinh n\lambda\frac{\pi}{2}}\sinh n\lambda x\right\}.$$

Wählen wir als willkürliche Konstante

$$a_n=M\frac{\cosh^2 n\lambda\frac{\pi}{2}}{\sinh n\lambda\frac{\pi}{2}},$$

so ergibt sich

$$\phi_n(x)=a_n\left\{\frac{\sinh n\lambda\frac{\ }{2}}{\cosh^2 n\lambda\frac{\pi}{2}}\cosh n\lambda x-\frac{2}{\pi}\frac{1}{\cosh n\lambda\frac{\pi}{2}}x\sinh n\lambda x\right\},\quad 0\leq x<\frac{\pi}{2}.$$

In derselben Weise kann gezeigt werden, dass $\psi_n(y)$ geschrieben werden kann

$$\psi_n(y)=c_n\left\{\frac{\sinh n\kappa\frac{\pi}{2}}{\cosh^2 n\kappa\frac{\pi}{2}}\cosh n\kappa y-\frac{2}{\pi}\frac{1}{\cosh n\kappa\frac{\pi}{2}}\,y\sinh n\kappa y\right\}$$

$$-\frac{1}{n^2}b_n\frac{1}{\cosh n\kappa\frac{\pi}{2}}\cosh n\kappa y,\quad 0\leqq y<\frac{\pi}{2},$$

wo c_n eine willkürliche Konstante bezeichnet.

Wir erhalten weiter

$$\phi_n'(x)=a_n\left\{\frac{\sinh n\lambda\frac{\pi}{2}}{\cosh^2 n\lambda\frac{\pi}{2}}n\lambda\sinh n\lambda x-\frac{2}{\pi}\frac{1}{\cosh n\lambda\frac{\pi}{2}}\sinh n\lambda x\right.$$

$$\left.-\frac{2}{\pi}\frac{1}{\cosh n\lambda\frac{\pi}{2}}n\lambda x\cosh n\lambda x\right\},\quad 0\leqq x<\frac{\pi}{2},$$

$$\psi_n'(y)=c_n\left\{\frac{\sinh n\kappa\frac{\pi}{2}}{\cosh^2 n\kappa\frac{\pi}{2}}n\kappa\sinh n\kappa y-\frac{2}{\pi}\frac{1}{\cosh n\kappa\frac{\pi}{2}}\sinh n\kappa y,\right.$$

$$\left.-\frac{2}{\pi}\frac{1}{\cosh n\kappa\frac{\pi}{2}}n\kappa y\cosh n\kappa y\right\}-\frac{\kappa}{n}b_n\frac{1}{\cosh n\kappa\frac{\pi}{2}}\sinh n\kappa y,\quad 0\leqq y<\frac{\pi}{2},$$

und demnach

$$\phi_n'\left(\frac{\pi}{2}-0\right)=-a_n.\frac{2}{\pi}.\frac{n\lambda\pi+\sinh n\lambda\pi}{1+\cosh n\lambda\pi},$$

$$\psi_n'\left(\frac{\pi}{2}-0\right)=-c_n.\frac{2}{\pi}.\frac{n\kappa\pi+\sinh n\kappa\pi}{1+\cosh n\kappa\pi}-\frac{\kappa}{n}.b_n.\frac{\sinh n\kappa\frac{\pi}{2}}{\cosh n\kappa\frac{\pi}{2}}.$$

Nun sollen die willkürlichen Konstanten a_n und c_n so bestimmt werden, dass den Randbedingungen 3 und 4 genügt wird. Zu diesem Zweck entwickeln wir $\phi_n(x)$ und $\psi_n(y)$ in gerade fouriersche Reihen, die nach ungeraden Vielfachen des Argumentes fortschreiten, indem wir eine Funktion $f(x)$ so definieren, dass

$$f(x) = \phi_n(x), \qquad 0 \leq x < \frac{\pi}{2},$$

$$f(x) = -\phi_n(\pi - x), \qquad \frac{\pi}{2} < x < \pi,$$

und diese in eine fouriersche cosinus-Reihe im Intervalle 0 bis π entwickeln.

Wir erhalten dann

$$\phi_n(x) = a_n n\lambda \left(\frac{4}{\pi}\right)^2 \sum_m (-1)^{\frac{m-1}{2}} \frac{m}{(m^2+n^2\lambda^2)^2} \cos mx, \quad 0 \leq x < \frac{\pi}{2},$$

wo m alle ungerade Zahlen von 1 bis ∞ durchläuft.

In ähnlicher Weise ergibt sich

$$\psi_n(y) = c_n n\kappa \left(\frac{4}{\pi}\right)^2 \sum_m (-1)^{\frac{m-1}{2}} \frac{m}{(m^2+n^2\kappa^2)^2} \cos my$$

$$-b_n \frac{1}{n^2} \frac{4}{\pi} \sum_m (-1)^{\frac{m-1}{2}} \frac{m}{m^2+n^2\kappa^2} \cos my, \quad 0 \leq y < \frac{\pi}{2}.$$

Die Randbedingung 4 gibt dann

$$\sum_n \psi_n'\left(\frac{\pi}{2} - 0\right) \cos nx - \sum_n a_n (-1)^{\frac{n-1}{2}} n^2\lambda \left(\frac{4}{\pi}\right)^2 \sum_m (-1)^{\frac{m-1}{2}} \frac{m}{(m^2+n^2\lambda^2)^2} \cos mx = 0,$$

wo also n und m alle ungerade Zahlen durchlaufen. Führen wir den Faktor $a_n(-1)^{\frac{n-1}{2}} n^2\lambda \left(\frac{4}{\pi}\right)^2$ unter das innere Summenzeichen ein, so ergibt sich

$$\sum_n \psi_n'\left(\frac{\pi}{2} - 0\right) \cos nx - \sum_n \sum_m (-1)^{\frac{m-1}{2}} (-1)^{\frac{n-1}{2}} \left(\frac{4}{\pi}\right)^2 \lambda \frac{a_n . n^2 m}{(m^2+n^2\lambda^2)^2} \cos mx = 0.$$

Wenn wir in der Doppelsumme m gegen n und umgekehrt austauschen, erhalten wir

$$\sum_n \psi_n'\left(\frac{\pi}{2} - 0\right) \cos nx - \sum_m \sum_n (-1)^{\frac{n-1}{2}} (-1)^{\frac{m-1}{2}} \left(\frac{4}{\pi}\right)^2 \lambda \frac{a_m m^2 n}{(n^2+m^2\lambda^2)^2} \cos nx = 0.$$

Kehren wir die Summierungsfolge um, so haben wir

$$\sum_n \psi_n'\left(\frac{\pi}{2} - 0\right) \cos nx - \sum_n (-1)^{\frac{n-1}{2}} \left(\frac{4}{\pi}\right)^2 n\lambda \cos nx \sum_m (-1)^{\frac{m-1}{2}} \frac{a_m m^2}{(m^2\lambda^2+n^2)^2} = 0.$$

Die beiden Summen, die sich nach n erstrecken, repräsentieren fouriersche Reihen, die nach ungeraden Vielfachen des Argumentes fortschreiten. Demnach müssen sie einander Glied nach Glied gleich sein, und folglich

$$\psi_n'\left(\frac{\pi}{2}-0\right)-(-1)^{\frac{n-1}{2}}\left(\frac{4}{\pi}\right)^2 n\lambda\sum_m(-1)^{\frac{m-1}{2}}\frac{a_m\cdot m^2}{(m^2\lambda^2+n^2)^2}=0.$$

Setzen wir den Wert von $\psi_n'\left(\frac{\pi}{2}-0\right)$ ein, so ergibt sich

$$-c_n\cdot\frac{2}{\pi}\cdot\frac{n\kappa\pi+\sinh n\kappa\pi}{1+\cosh n\kappa\pi}-b_n\cdot\frac{\kappa}{n}\cdot\frac{\sinh n\kappa\frac{\pi}{2}}{\cosh n\kappa\frac{\pi}{2}}$$

$$-(-1)^{\frac{n-1}{2}}\left(\frac{4}{\pi}\right)^2 n\lambda\sum_m(-1)^{\frac{m-1}{2}}\frac{a_m\cdot m^2}{(m^2\lambda^2+n^2)^2}=0,$$

und demnach

(8)
$$c_n+b_n\cdot\frac{\kappa}{n}\cdot\frac{\pi}{2}\cdot\frac{\sinh n\kappa\pi}{n\kappa\pi+\sinh n\kappa\pi}$$
$$+(-1)^{\frac{n-1}{2}}\cdot\frac{8}{\pi}\cdot\frac{1+\cosh n\kappa\pi}{n\kappa\pi+\sinh n\kappa\pi}\,n\lambda\sum_m(-1)^{\frac{m-1}{2}}\frac{a_m\cdot m^2}{(m^2\lambda^2+n^2)^2}=0.$$

Die Randbedingung 3 gibt

$$-\sum_n(-1)^{\frac{n-1}{2}}c_n n^2\kappa\left(\frac{4}{\pi}\right)^2\sum_m(-1)^{\frac{m-1}{2}}\frac{m}{(m^2+n^2\kappa^2)^2}\cos my$$

$$-\sum_n(-1)^{\frac{n-1}{2}}(-b_n)\frac{1}{n}\cdot\frac{4}{\pi}\sum_m(-1)^{\frac{m-1}{2}}\frac{m}{m^2+n^2\kappa^2}\cos my$$

$$+\sum_n\phi_n'\left(\frac{\pi}{2}-0\right)\cos ny=Q.$$

Entwickeln wir Q im Intervalle $-\frac{\pi}{2}$ bis $+\frac{\pi}{2}$ in eine fouriersche Reihe, die nach ungeraden Vielfachen des Argumentes fortschreitet, erhalten wir

$$Q=Q\frac{4}{\pi}\sum_n(-1)^{\frac{n-1}{2}}\frac{1}{n}\cos ny,\quad 0\leq y<\frac{\pi}{2},$$

wo n alle ungerade Zahlen durchläuft. Setzen wir diesen Wert in die obige Gleichung ein, ergibt sich nach Umtransformation

$$-\sum_n\sum_m(-1)^{\frac{n-1}{2}}(-1)^{\frac{m-1}{2}}\left(\frac{4}{\pi}\right)^2\kappa\frac{c_n n^2 m}{(m^2+n^2\kappa^2)^2}\cos my$$

$$+\sum_n\sum_m(-1)^{\frac{n-1}{2}}(-1)^{\frac{m-1}{2}}\frac{4}{\pi}\frac{b_n\cdot m}{n(m^2+n^2\kappa^2)}\cos my$$

$$+\sum_n\phi_n'\left(\frac{\pi}{2}-0\right)\cos ny-\sum_n Q\frac{4}{\pi}(-1)^{\frac{n-1}{2}}\frac{1}{n}\cos ny=0.$$

Wenn wir in den beiden Doppelsummen die Summierungsfolge umkehren und in den einfachen Summen n gegen m austauschen, ergibt sich

$$-\sum_m (-1)^{\frac{m-1}{2}} \left(\frac{4}{\pi}\right)^2 m\kappa \cos my \,.\, \sum_n (-1)^{\frac{n-1}{2}} \frac{c_n n^2}{(m^2+n^2\kappa^2)^2}$$

$$+\sum_m (-1)^{\frac{m-1}{2}} \frac{4}{\pi} m \cos my \sum_n (-1)^{\frac{n-1}{2}} \frac{b_n}{n(m^2+n^2\kappa^2)}$$

$$+\sum_m \phi_m' \left(\frac{\pi}{2}-0\right) \cos my - \sum_m Q \frac{4}{\pi} (-1)^{\frac{m-1}{2}} \frac{1}{m} \cos my = 0.$$

Sämtliche Summen, die sich nach m erstrecken, repräsentieren fouriersche Reihen, die nach ungeraden Vielfachen des Argumentes fortschreiten, und demnach muss

$$-(-1)^{\frac{m-1}{2}} \left(\frac{4}{\pi}\right)^2 m\kappa \sum_n (-1)^{\frac{n-1}{2}} \frac{c_n \,.\, n^2}{(m^2+n^2\kappa^2)^2}$$

$$+(-1)^{\frac{m-1}{2}} \frac{4}{\pi} m \sum_n (-1)^{\frac{n-1}{2}} \frac{b_n}{n(m^2+n^2\kappa^2)}$$

$$+\phi_m\left(\frac{\pi}{2}-0\right) - Q \frac{4}{\pi} (-1)^{\frac{m-1}{2}} \frac{1}{m} = 0$$

sein.

Setzen wir den Wert von $\phi_m\left(\frac{\pi}{2}-0\right)$ sowie von Q nach Gl. (4) ein und multiplizieren mit $-(-1)^{\frac{m-1}{2}} \left(\frac{\pi}{4}\right)^2 \frac{1}{m\kappa}$, ergibt sich, wenn wir berücksichtigen dass $\frac{1}{\kappa} = \lambda$ ist,

$$+\sum_n (1)^{\frac{n-1}{2}} \frac{C_n n^2}{(m^2+n^2\kappa^2)^2} - \frac{\pi}{4} \lambda \sum_n (-1)^{\frac{n-1}{2}} \frac{b_n}{n(m^2+n^2\kappa^2)}$$

$$+(-1)^{\frac{m-1}{2}} \left(\frac{\pi}{4}\right)^2 \frac{\lambda}{m} a_m \frac{2}{\pi} \frac{m\lambda\pi + \sinh m\lambda\pi}{1+\cosh m\lambda\pi}$$

$$+\frac{\pi}{4} \frac{\lambda}{m^2} \sum_n (-1)^{\frac{n-1}{2}} \frac{b_n}{n} = 0. \tag{9}$$

Nach Gl. (8) ist jedoch

$$c_n + b_n \frac{\kappa}{n} \frac{\pi}{2} \frac{\sinh n\kappa\pi}{n\kappa\pi + \sinh n\kappa\pi}$$

$$+(-1)^{\frac{n-1}{2}} \frac{8}{\pi} \frac{1+\cosh n\kappa\pi}{n\kappa\pi + \sinh n\kappa\pi} n\lambda \sum_m (-1)^{\frac{m-1}{2}} \frac{a_m \,.\, m^2}{(m^2\lambda^2+n^2)^2} = 0,$$

und demnach

$$\sum_n (-1)^{\frac{n-1}{2}} \frac{c_n n^2}{(m^2+n^2\kappa^2)^2} + \sum_n (-1)^{\frac{n-1}{2}} \frac{\pi}{2}\kappa b_n \frac{\sinh n\kappa\pi}{n\kappa\pi+\sinh n\kappa\pi} \frac{n}{(m^2+n^2\kappa^2)^2}$$

$$+ \sum_n \frac{8}{\pi}\lambda \frac{1+\cosh n\kappa\pi}{n\kappa\pi+\sinh n\kappa\pi} \cdot \frac{n^3}{(m^2+n^2\kappa^2)^2} \sum_\mu (-1)^{\frac{\mu-1}{2}} \frac{a_\mu \cdot \mu^2}{(\mu^2\lambda^2+n^2)^2} = 0,$$

wo μ alle ungerade Zahlen durchläuft.

Ziehen wir diese Gleichung von Gl. (9), ergibt sich

$$(-1)^{\frac{m-1}{2}} \frac{\pi}{8} \frac{\lambda}{m} a_m \frac{m\lambda\pi+\sinh m\lambda\pi}{1+\cosh m\lambda\pi} + \frac{\pi}{4} \frac{\lambda}{m^2} \sum_n (-1)^{\frac{n-1}{2}} \frac{b_n}{n}$$

$$-\frac{\pi}{4}\lambda \sum_n (-1)^{\frac{n-1}{2}} \frac{b_n}{n(m^2+n^2\kappa^2)} - \sum_n (-1)^{\frac{n-1}{2}} b_n \frac{\pi}{2} \kappa \frac{\sinh n\kappa\pi}{n\kappa\pi+\sinh n\kappa\pi} \frac{n}{(m^2+n^2\kappa^2)^2}$$

$$-\sum_n \left(\frac{4}{\pi}\right)^2 \lambda \frac{1+\cosh n\kappa\pi}{n\kappa\pi+\sinh n\kappa\pi} \cdot \frac{\pi}{2} \cdot \frac{n^3}{(m^2+n^2\kappa^2)^2} \sum_\mu (-1)^{\frac{\mu-1}{2}} \frac{a_\mu \cdot \mu^2}{(\mu^2\lambda^2+n^2)^2} = 0,$$

oder nach Säuberung

$$(10)\quad a_m(-1)^{\frac{m-1}{2}} \frac{m\lambda\pi+\sinh m\lambda\pi}{1+\cosh m\lambda\pi}$$

$$-\left(\frac{4}{\pi}\right)^2 . 4m \sum_{n=1,3,5..}^{\infty} \frac{1+\cosh n\kappa\pi}{n\kappa\pi+\sinh n\kappa\pi} \frac{n^3}{(m^2+n^2\kappa^2)^2} . \sum_{\mu=1,3,5..}^{\infty} (-1)^{\frac{\mu-1}{2}} \frac{a_\mu \cdot \mu^2}{(\mu^2\lambda^2+n^2)^2}$$

$$= -\sum_{n=1,3,5..}^{\infty} (-1)^{\frac{n-1}{2}} \frac{2n\kappa^2 b_n}{m(m^2+n^2\kappa^2)} + \sum_{n=1,3,5..}^{\infty} (-1)^{\frac{n-1}{2}} b_n \kappa^2 \frac{\sinh n\kappa\pi}{n\kappa\pi+\sinh n\kappa\pi} \cdot \frac{4mn}{(m^2+n^2\kappa^2)^2}.$$

Dieser Ausdruck repräsentiert für jedes m eine Gleichung zwischen den unbekannten Konstanten a_1, a_3, a_5, ... und wir haben demnach, um diese zu berechnen, eine ausreichende Anzahl von Gleichungen von folgendem Aussehen erhalten:

$$\begin{aligned}
-(\alpha_{11}-\beta_1)a_1+\alpha_{13}a_3-\alpha_{15}a_5+\alpha_{17}a_7 + \ldots &= \gamma_1,\\
-\alpha_{31}a_1+(\alpha_{33}-\beta_3)a_3-\alpha_{35}a_5+\alpha_{37}a_7 + \ldots &= \gamma_3,\\
-\alpha_{51}a_1+\alpha_{53}a_3-(\alpha_{55}-\beta_5)a_5+\alpha_{57}a_7 + \ldots &= \gamma_5,\\
-\alpha_{71}a_1+\alpha_{73}a_3-\alpha_{75}a_5+(\alpha_{77}-\beta_7)a_7 + \ldots &= \gamma_7,\\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots &,
\end{aligned}$$

wo

$$\alpha_{m\mu} = \left(\frac{4}{\pi}\right)^2 4m\mu^2 \sum_{n=1,3,5..}^{\infty} \frac{1+\cosh n\kappa\pi}{n\kappa\pi+\sinh n\kappa\pi} \cdot \frac{n^3}{(m^2+n^2\kappa^2)^2(\mu^2\lambda^2+n^2)^2},$$

$$\beta_m = \frac{m\lambda\pi+\sinh m\lambda\pi}{1+\cosh m\lambda\pi},$$

$$\gamma_m = -\sum_{n=1,3,5..}^{\infty} (-1)^{\frac{n-1}{2}} \frac{2n\kappa^2 b_n}{m(m^2+n^2\kappa^2)} + \sum_{n=1,3,5..}^{\infty} (-1)^{\frac{n-1}{2}} \kappa^2 b_n \frac{\sinh n\kappa\pi}{n\kappa\pi+\sinh n\kappa\pi} \frac{4mn}{(m^2+n^2\kappa^2)^2}.$$

Die Beiwerte $\alpha_{m\mu}$ und β_m können sofort berechnet werden, weil sie von der Belastung nicht abhängen, und die Beiwerte γ_m, wenn die Belastung p bekannt ist.

Die Auflösung dieses Gleichungssystems kann vollzogen werden, wenn man berücksichtigt, dass die Summe der unbekannten Glieder rechts von der Diagonale und das Diagonalglied grösser ist als die Summe der Glieder links von der Diagonale. Bezeichnen wir mit $a_{m.0}$, $a_{m,1}$, $a_{m,2}$, $a_{m,3}$, ... auf einander folgende Annäherungswerte von a_m, können wir schreiben

$$1. \quad \begin{cases} a_{1,0} = -\dfrac{\gamma_1}{\alpha_{11}-\beta_1}, \\ a_{3,0} = \dfrac{\gamma_3 + \alpha_{31}a_{1,0}}{\alpha_{33}-\beta_3}, \end{cases}$$

$$2. \quad \begin{cases} a_{1,1} = -\dfrac{\gamma_1 - \alpha_{13}a_{3,0}}{\alpha_{11}-\beta_1}, \\ a_{3,1} = \dfrac{\gamma_3 + \alpha_{31}a_{1,1}}{\alpha_{33}-\beta_3}, \\ a_{5,0} = -\dfrac{\gamma_5 - \alpha_{51}a_{1,2} - \alpha_{53}a_{3,1}}{\alpha_{55}-\beta_5}, \end{cases}$$

$$3. \quad \begin{cases} a_{1,2} = -\dfrac{\gamma_1 - \alpha_{13}a_{3,1} + \alpha_{15}a_{5,0}}{\alpha_{11}-\beta_1}, \\ a_{3,2} = \dfrac{\gamma_3 + \alpha_{31}a_{1,2} + \alpha_{35}a_{5,0}}{\alpha_{33}-\beta_3}, \\ a_{5,1} = -\dfrac{\gamma_5 + \alpha_{51}a_{1,2} - \alpha_{53}a_{3,2}}{\alpha_{55}-\beta_5}, \\ a_{7,0} = \dfrac{\gamma_7 + \alpha_{71}a_{1,2} - \alpha_{73}a_{3,2} + \alpha_{75}a_{5,1}}{\alpha_{77}-\beta_7}, \end{cases}$$

u.s.w.

Es zeigt sich, dass die auf einander folgenden Annäherungswerten

$$a_{m,0},\ a_{m,1},\ a_{m,2},\ a_{m,3}, \ldots$$

gegen einen bestimmten Wert a_m konvergieren.

Nachdem die Konstanten a_m berechnet sind, lassen sich die unbekannten Konstanten c_n zweckmässig aus Gl. (8) berechnen.

Hiermit sind die Funktionen $\phi_n(x)$ und $\psi_n(y)$ bestimmt und somit auch die airysche Spannungsfunktion F nach Gl. (5). Durch die linearen Substitutionen (1) erhält man dann F als Funktion von ξ und η, und damit ist das aufgestellte Problem vollständig gelöst.

Um näher zu studieren, wie rasch die Annäherungswerte $a_{m,0}$, $a_{m,1}$, $a_{m,2}$ u.s.w. konvergieren, sind die Beiwerte $\alpha_{m\mu}$ und β_m für $m \leqq 7$ und $\mu \leqq 7$ bei der quadratischen Scheibe, d.h. für $\kappa = \lambda = 1$, berechnet worden.

Es ergibt sich für $\alpha_{m\mu}$ folgende Tafel:

$\mu =$	1	3	5	7
$m=1$	0.36720,	0.18673,	0.10826,	0.07225,
$m=3$	0.06225,	0.08530,	0.07901,	0.06698,
$m=5$	0.02165,	0.04746,	0.05355,	0.05139,
$m=7$	0.01032,	0.02871,	0.03671,	0.03850.

und für β_m folgende Werte:

$$\beta_1 = 1.16642,\ \beta_3 = 1.00136,\ \beta_5 = 1.00000,\ \beta_7 = 1.00000.$$

Für nachstehende Belastungen $p = \cos nx$ ergeben sich für $n \leqq 7$ die in nachstehender Tafel enthaltenen Werte von γ_m:

$n =$	1	3	5	7
γ_1	$= -0.213856$,	$+0.480183$,	-0.355029,	$+0.268800$,
γ_3	$= +0.027670$,	$+0.000169$,	-0.046136,	$+0.055490$,
γ_5	$= +0.007874$,	-0.016530,	-0.000000,	$+0.012272$,
γ_7	$= +0.003091$,	-0.010154,	$+0.006261$,	$+0.000000$.

Beispeilsweise erhält man für die Belastung $p = \cos x$ nachstehende Annäherungswerte von a_m:

$$1.\ \begin{cases} a_{1,0} = -0.267, \\ a_{3,0} = -0.0120, \end{cases}$$

$$2.\ \begin{cases} a_{1,1} = -0.265, \\ a_{3,1} = -0.01214, \\ a_{5,0} = +0.00286, \end{cases}$$

$$3.\ \begin{cases} a_{1,2} = -0.2643, \\ a_{3,2} = -0.01252, \\ a_{5,1} = +0.002901, \\ a_{7,0} = -0.000861, \end{cases}$$

$$4.\ \begin{cases} a_{1,3} = -0.26411, \\ a_{3,3} = -0.012570, \\ a_{5,2} = +0.002955, \\ a_{7,1} = -0.000889. \end{cases}$$

Um die Berechnung eines gewissen Belastungsfalles zu erleichtern, sind in nachstehender Tafel die Konstanten a_m für die einfachen Belastungsfälle $p = \cos nx$ berechnet, wo n die Werte 1, 3, 5 und 7 gegeben sind.

Es ergibt sich für

n =	1	3	5	7
a_1	−0.2641,	+0.6099,	−0.4657,	+0.3607,
a_3	−0.0126,	−0.0414,	+0.0834,	−0.0878,
a_5	+0.0030,	−0.0016,	−0.0149,	+0.0260,
a_7	−0.0009,	+0.0028,	+0.0015,	−0.0075.

Es wäre wünschenswert, dass die vorerwähnte Zifferrechnung teils mit grösserer Schärfe, teils für Konstanten a_m mit m grösser als 7, teils auch für Werte von n grösser als 7 ausgeführt werden würde. Leider hat der Verfasser zufolge der zeitraubenden Natur der Rechnungen keine Gelegenheit dazu gehabt (bei der Berechnung der Beiwerte $\alpha_{m\mu}$ sind, um die vorstehende Schärfe zu erzielen, 25 bis 40 Glieder und ein geschätztes Restglied mitgenommen worden).

Um die Methode zu prüfen, ist ein Beispiel durchgerechnet worden. Es sei eine quadratische Scheibe mit $a=b=\frac{\pi}{2}$ durch eine Belastung

$$\rho=-[0.4355 \cos x+0.3676 \cos 3x+0.2508 \cos 5x+0.1169 \cos 7x]^*,$$

die aus Abb. 2 näher hervorgeht, belastet. Die aus den vorstehenden Formeln

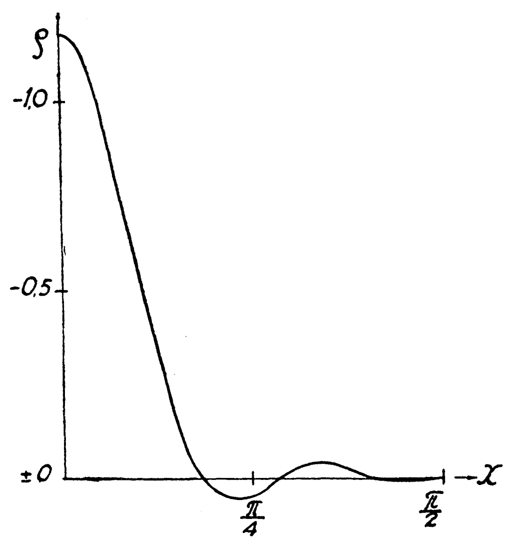


Abb. 2. Die Belastung ρ.

***Diese sind die ersten Glieder, die man erhält, wenn man die Belastung**

$$0<x<\frac{\pi}{9},\ \rho=-1;\ \frac{\pi}{9}<x<\frac{\pi}{2},\ \rho=0$$

in eine fouriersche Reihe, die nach ungeraden Vielfachen des Argumentes fortschreitet, entwickelt.

berechneten Spannungen sind in Abb. 3, 4 und 5 für das zwischen den positiven

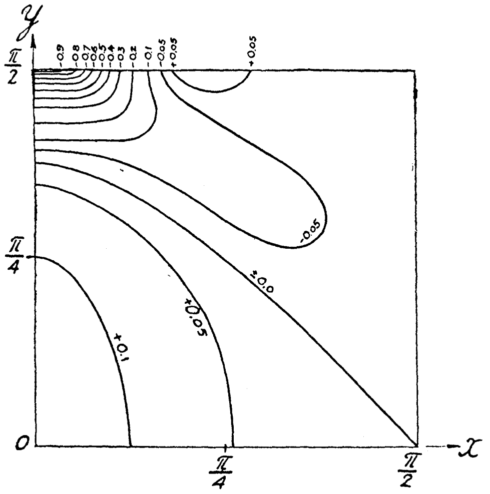


Abb. 3. Die Spannung σ_x.

x- und y-Achsen gelegene Scheibenviertel graphisch dargestellt. In Abb. 6

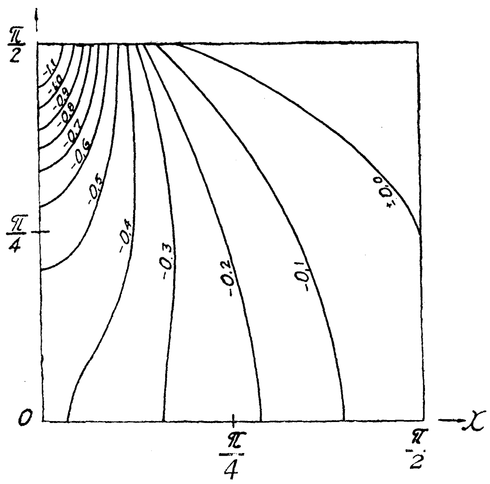


Abb. 4. Die Spannung σ_y.

sind für einige Punkte die Hauptspannungen sowie ihre Richtung mit zeichnerischer Genauigkeit dargestellt.

Zusammenfassung.

Es wird gezeigt, dass die Annahme, die airysche Spannungsfunktion könne für eine rechteckige Scheibe mit den Rändern $\xi = \pm a$, $\eta = \pm b$, welche den Rändern

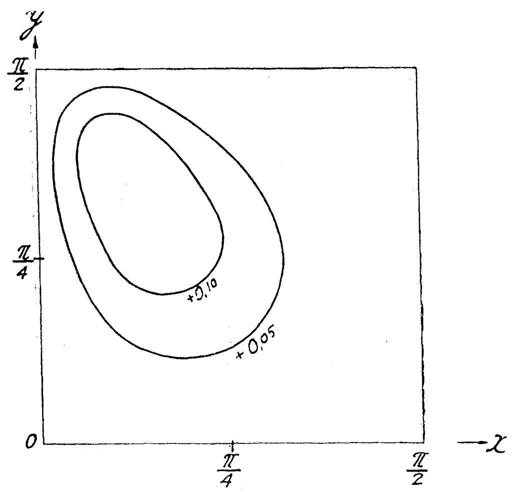


Abb. 5. Die Spannung τ_{xy}.

$\eta = \pm b$ entlang mit einer senkrecht zum Rande stehenden und bezüglich der η-Achse symmetrischen Last belastet sei, in der Form

$$F = \sum_n \psi_n \left(\frac{n}{2b} \eta \right) \cos n \frac{\pi}{2a} \xi + \sum_n \phi_n \left(\frac{\pi}{2a} \xi \right) \cos n \frac{\pi}{2b} \eta,$$

wo die Summen sich über alle ungerade Zahlen n erstrecken, dargestellt werden, zu einem Gleichungssystem führt, welches aufgelöst werden kann, so dass die

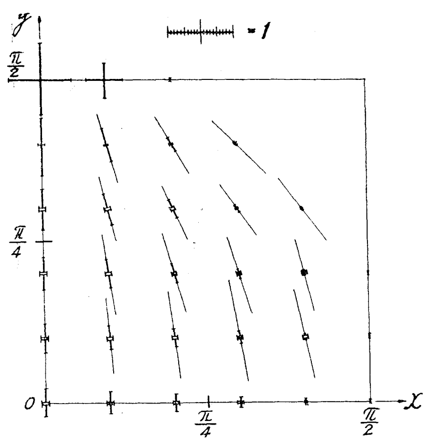


Abb. 6. Darstellung der Hauptspannungen sowie ihrer Richtung
Druckspannung
Zugspannung

Funktionen ψ_n und ϕ_n und damit die airysche Spannungsfunktion sich bestimmen lässt.

RESONANT VIBRATION IN STEEL BRIDGES

BY PROFESSOR B. P. HAIGH AND ALBERT BEALE (Whitworth Senior Scholar),
Royal Naval College, Greenwich, England.

In designing steel bridges, the stresses are calculated in terms of known dead-loads and moving-loads; and "impact factors" are commonly applied, in relation to the moving-loads, to cover recognized uncertainties as to their mode of action. Since the degree of uncertainty varies widely in different circumstances, it appears that the unrestricted use of such an impact factor may be a source of waste, or of danger, or of both.

Stresses and deflections induced in bridges by typical moving-loads have recently been measured by several responsible bodies*. The investigations of the Indian Railway Bridge Committee, in particular, have made it clear that the variations in stress depend on the speed and balance of the locomotive and on the mass and stiffness of the structure, as well as on the gross magnitude of the moving-load. Although the Committee found itself unable to recommend any improvement on the conventional practice of assuming that the impactive action is proportional to the moving-load, and varies with the loaded length in the manners allowed for by the Pencoyd and other formulae, the hope was expressed that further experiments might suffice to establish some other method of calculation, such as might more directly take account of the several contingencies.

The matter has recently been advanced by a mathematical analysis given by Professor Inglis†. Dealing with the case of a uniform beam, Professor Inglis demonstrated the dangers associated with "resonance", *i.e.*, that condition in which the frequency of one of the pulsating applied forces coincides with the frequency of one of the natural modes of vibration of the loaded structure. When resonance occurs, the cumulative effect of successive pulsations may produce quite abnormal stresses.

Hitherto, it has been sought to establish an impact factor applicable for all bridges alike. If such a factor is large enough to cover instances in which the state of resonance may occur. it is clearly unduly large for others, so that there is a real need for discrimination between the cases.

It is suggested therefore, that in analysing records of stresses and deflections obtained experimentally, attention should be directed to two distinct problems, viz:—

*American Railway Engineering Association, 1910. Indian Railway Bridge Committee, 1917-1921. Ministry of Transport, 1921.

†Institution of Civil Engineers, 18th March, 1924.

(1) What are the conditions for resonance, entailing the use of abnormally high impact factors?

(2) What lower value of the impact factor will suffice when these conditions are *not* realized?

It is further suggested, that in applying the results in bridge design, resonance should be avoided wherever reasonably possible—either by arranging the design so that the natural frequencies of the structure lie above those of any important disturbing forces, or by limiting train speeds. If, however, it is decided to accept the possibility of resonance in some cases where the methods for avoiding it are inconvenient, then the high impact factor appropriate to the case must be used. It should be remarked, in addition, that if the natural frequency of the spring-supported load is about the same as that of the bridge, a high impact factor is again necessary. On the other hand, if the possibility of resonance is avoided, the lower impact factor* may be used with safety.

Such a dual process in analysis and design would obviate the waste incurred by the adoption—for all bridges—of high impact factors based on exceptional results observed in a few; and at the same time would eliminate the possibility that, in a quite exceptional case, even such high impact allowances might not suffice to avert the danger of excessive vibration.

The dangers associated with resonance are mitigated, to some extent, by frictional forces which absorb some of the energy which otherwise would be accumulated by the repeated application of the pulsating forces. Any considerable measure of internal friction, however, would probably lead to abrasion or to fatigue; and it appears unlikely that frictional forces will play any very important part in damping the forced vibrations in bridges. For any particular bridge the importance of friction may be judged by observing the manner in which vibration dies away after the passage of a train.

The conditions for resonance may be ascertained, in the case of existing bridges, by experiment; and in the case of a structure in course of design, by means of calculations detailed hereafter. Since it is desirable to determine such conditions in the design stage, the present article has been developed to describe a convenient method of ascertaining the several frequencies of natural vibration of a framed structure, without further reference to the more controversial aspects relating to the practical application of such calculations.

The methods may readily be adapted for plate girders; but the present system has been adjusted in detail so as to employ, as far as possible, data that will usually be available already in the case of a *framed* structure.

Natural Vibration in Bridges

A steel bridge is a massive and elastic structure and, as such, is able to vibrate in a number of different ways, with different "natural" frequencies

*This impact factor must cover the effects due to the sudden application of the load to the elastic structure, and to forced, but non-resonant, vibrations due to incomplete balance of locomotives, to rail joints, and to flats on wheels; to centrifugal action due to bending the span, and to lateral surging on the track.

corresponding to the notes and overtones making up the sound emitted by a bell. As in the case of a bell, the several frequencies are only distantly related, except in the respect that all are proportional to the square root of the quotient obtained by dividing the Modulus of Elasticity by the Density. In a structure as complex as a bridge, the possible frequencies are so numerous and varied that the noise emitted may not indicate a dominant note; and, indeed, the important vibrations, which strain the main members and have natural frequencies approaching those of any probable periodic forces acting on the structure, lie below the audible range.

Of the possible modes of vibration, the primary transverse one alone appears to have been extensively noted by maintenance engineers and other observers. The primary frequency is held by many to be given with fair accuracy by the empirical formula:

$$n = 4\sqrt{\frac{L}{(D+L)\Delta}},$$

where L = Live Load,
D = Dead Load,
Δ = Static Deflection, in feet, produced by a steady load equal to the Live Load.

This formula would be almost exactly accurate in the case of a uniform beam, uniformly loaded.

While the above formula, and other simple expressions, may be of service in a limited scope of application, the problem merits more detailed consideration. The rational method of calculation described hereafter is valid for bridges having proportions such as would render the above formula valueless or misleading; and may be applied, also, for calculating the frequencies of the secondary and other harmonic vibrations, and the torsional frequencies which, in many cases, may be of practical importance.

The method does not fully cover the case in which the spring-supported load partakes in the natural vibration of the bridge; but when the bridge is stiff in comparison with the springs supporting the load, the degree of approximation is high. In other cases the effects of the springing, both on the natural vibrations and on the forced, merit special investigation.

Section I gives a concise account of the routine of calculation for the primary transverse mode of vibration.

Section II explains the basis of calculation, shows how certain factors (mentioned in I) are determined; and indicates how the method may readily be extended to the calculation of frequencies of the higher modes of vibration in bridges or other structures.

Section I

Routine Calculation for Frequency of Primary Transverse Vibration of Bridges of Ordinary Types

(1) Since the vibration of the bridge as a whole is governed directly by the span and by the properties of the metal of the main girders, calculate first the

so-called "longitudinal frequency", n_L, which may be described as that of a sound wave echoing to and fro in a bar of solid metal, *e.g.*, a rail, of length equal to the span. This frequency is

$$n_L = V/2s$$

where s is the span of the bridge, and V is the velocity of sound in the metal, given by:

$$V = \sqrt{gE/w},$$

where

E = Young's Modulus of Elasticity for the metal,

w = weight of the metal, per unit-volume.

Thus, for constructional steel

$$V = \sqrt{\frac{32.2 \text{ feet}}{\text{sec.}^2} \cdot \frac{12{,}500 \text{ Ton}}{\text{in.}^2} \cdot \frac{\text{in.}^3}{0.28 \text{ lb}} \cdot \frac{\text{ft.}}{12 \text{ in.}} \cdot \frac{2240 \text{ lb}}{\text{Ton}}} = 16{,}400 \text{ ft./sec.}$$

In the case of a 200 feet span, therefore,

$$n_L = (16{,}400 \text{ ft./sec.}) \div (2 \times 200 \text{ ft.}) = 41/\text{sec.}$$

The relation between this "longitudinal frequency" and the actual frequency depends on the design of the main girders, and on the load carried.

(2) Proceed to deduce the natural transverse frequency of the main girder, carrying no masses other than its own, by multiplication by a factor K_G which, since it depends solely on the design of the girder and the mode of vibration in question, may be termed the "girder-mode" factor.

The value of the factor may be calculated, for any given girder and mode, by a method described later. Alternatively, at least in the case of vibration in the primary mode, the value may be taken from tables summarizing calculations for girders of approximately similar design. The value varies with the ratio of the depth of the girder to its span, with the type of panel, with the relative stiffness of the individual members, and, in some degree, with the rigidities of the joints and the actions of the secondary stresses.

For example, in the case of a Warren girder made up of nine panels of equilateral profile, the value of the girder-mode factor may be taken—from Figure 1, discussed later—as

$$K_G = 0.11.$$

So that, in the case of a 200 feet span with girders of this type, the "girder frequency" would be

$$n_G = K_G \,.\, n_L = 0.11 \times 41/\text{sec.} = 4.5/\text{sec.}$$

The natural frequency of the bridge as a whole, however, will be less than this "girder frequency" because so far, the only masses taken into account are those of the girder itself.

(3) Proceed to deduce the frequency of the complete bridge, carrying the additional masses of the cross-bracing, cross-girders, flooring, etc., and the unsprung portion of the moving load—in fact all the loads that vibrate with

the girder, except its own mass, which has already been taken into account—by means of a second factor, viz.:

$$\sqrt{K_M . R + 1}.$$

The symbol R denotes the ratio between the additional load vibrating with the bridge and the mass of the main girders. If the whole moving load were

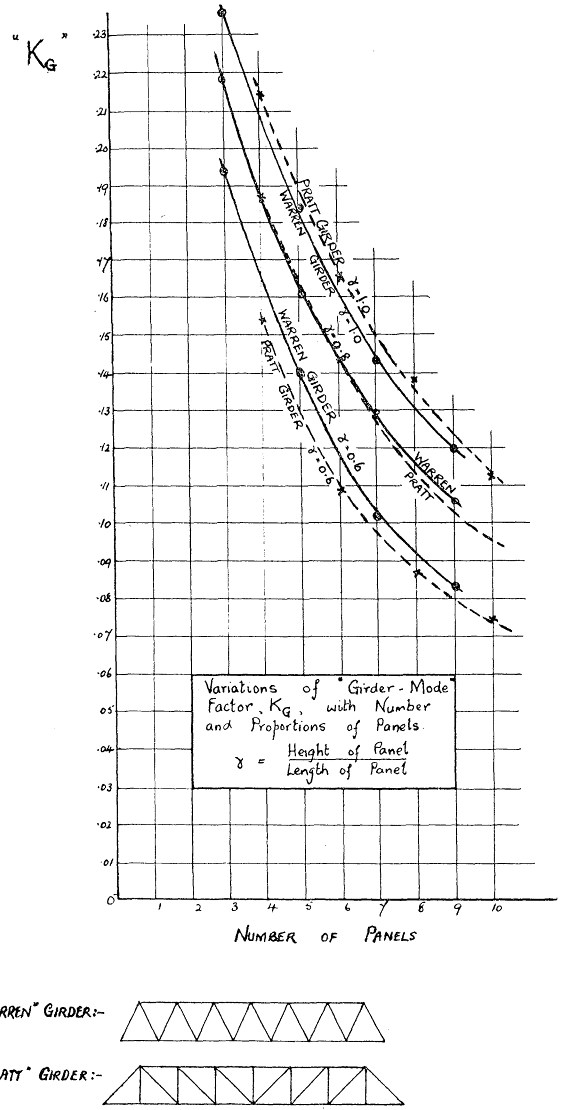


Fig. 1

unsprung R would be equal to the "Utility Ratio", U, defined as the ratio between the additional load carried and the mass of the main girders.

The factor K_M allows for the manner in which this load is distributed, relatively to the mass of the main girder. If the two loads are distributed in the same manner, whether uniformly or non-uniformly, the value of K_M is unity. For a single concentrated load at different positions ranging from the end to the middle of any bridge of ordinary design, K_M will range from 0 to about 2, and for a fully loaded bridge it is not likely to be very widely different from unity.

Thus, if the 200 feet Warren girders already mentioned carry an unsprung load of six times their own mass, and if the distributions of the masses are more or less similar, the frequency of primary transverse vibration of the bridge is

$$n = n_G \div \sqrt{K_M R + 1}$$
$$= (4.5/\text{sec.}) \div \sqrt{1 \times 6 + 1} = 1.7/\text{sec.}$$

(4) Summarizing the three preceding stages, the frequency of primary transverse vibration of the complete bridge is

$$n = (V/2s)\ (K_G) \div \sqrt{K_M R + 1}.$$

The formula has been arranged so as to embody the principles of dynamical similarity as far as may be possible in a convenient expression; and the use of the two factors K_G and K_M, in conjunction with the span s, velocity of sound V, and the ratio R, may be regarded as a method of allowing for complex effects in a simple manner that is amply accurate in the case of bridges of ordinary proportions and design.

The values of these factors may be determined by experiment or by calculation. A number of values of K_G, calculated to ascertain the effect of certain changes in design, are summarized in Figure 1. The values are for Warren and Pratt girders with different numbers of panels from three to ten, and with different ratios between the depth and length of panel. In all cases compared, the members have been proportioned so that, under a uniformly distributed load, the stresses in tension members are twice as great as those in compression members.

The foregoing routine, with other values of K_G and K_M, might be applied to calculate the frequencies of higher modes of vibration. But, in general, since the appropriate values of K_G and K_M are difficult to tabulate in a rational manner, it appears preferable to calculate higher modes by a fundamental method, which may be used, also, in determining the values of K_G and K_M applicable for the primary mode.

Section II

General Method for Calculation of Frequency of Natural Vibrations

A general method of calculation, taking direct account of the leading deta ls of the design, and applicable for any mode of vibration, may be summarized as follows:

(1) Sketch the profile of the structure flexed to an extreme position corresponding to the mode of vibration under consideration; and add, in the sketch, arrows representing a system of imaginary loads such as would produce an approximately similar profile of flexure—see Figures 2*a*, 2*b*, and 3. As will

be shown, it is not necessary that the approximation should be very close, provided that the general characteristics are reproduced. In choosing a suitable imaginary load, therefore, it will be permissible to consider the convenience of subsequent calculation. The number of forces employed will be no greater than may be necessary, and the forces will be applied at joints.

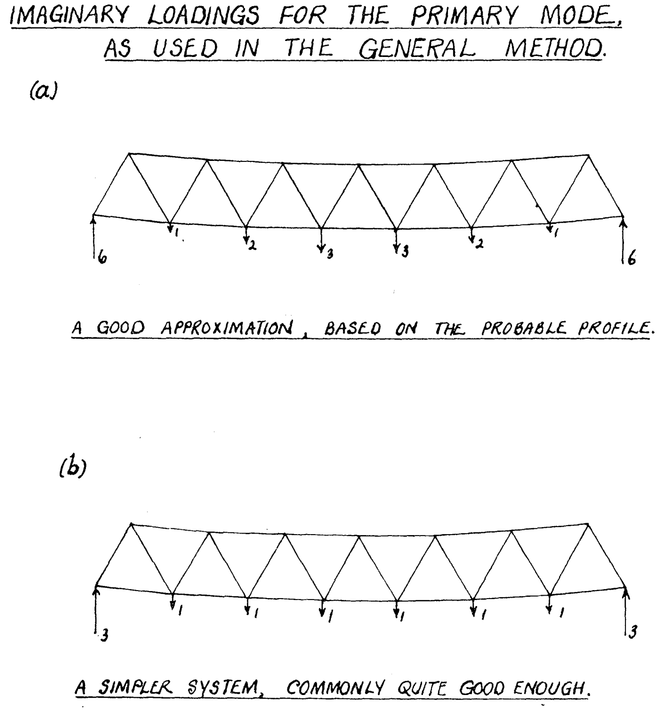


Fig 2

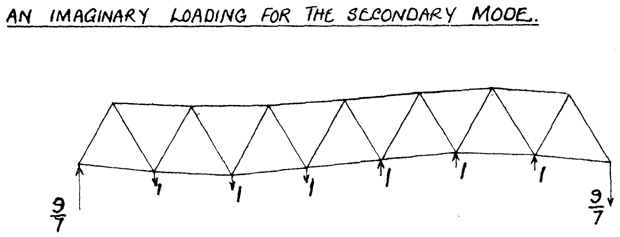


Fig 3

(2) Using methods of investigation now generally current in designing offices, proceed to determine the forces, F (in the members) and the deflections, D (of the joints), induced by the application of the above imaginary system of loading. For example, the stresses may be determined graphically, by means of the Clerk Maxwell Reciprocal Diagram (see Figure 4); or algebraically

using R. V. Southwell's notation*. If the structure be "simply-firm", only the principles of equilibrium need be applied; but if it be "redundant", the effects of the additional members must be taken into account if their importance warrants it. Likewise, the displacements of the joints may be determined graphically, by the Williot Diagram (kinematic); or algebraically, by the Re-

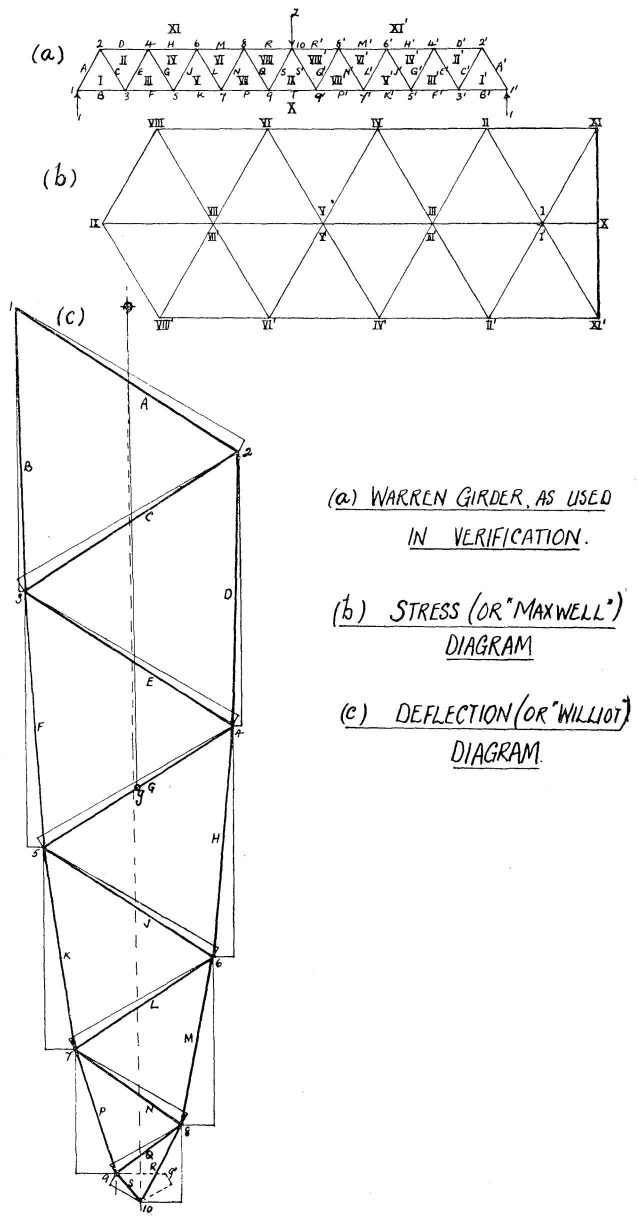


Fig 4

Primary Stress-Determination in Space Frames, Engineering, February 6, 1920.

silience method: but it may be noted that, since the displacements required are the vector sums of the vertical and horizontal components, and since the magnitudes and directions are required for all joints in the structure, the Williot diagram is usually much more convenient than the Resilience method. In either case it will be necessary to calculate the stiffness of the individual members; and these values will be used again in a subsequent stage.

$$S = \text{Stiffness} = \frac{\text{Axial force applied to member}}{\text{Consequent change in length}} = \frac{AE}{L}.$$

For example, if $E = 12{,}000$ Ton/in.², $A = 10$ in.² and $L = 240$ in.,

$$\text{then } S = (12{,}000 \text{ Ton/in.}^2) \times (10 \text{ in.}^2) \div (240 \text{ in.}) = 500 \text{ Ton/in.}$$

(3) Evaluate, for all members of the main girders and for any auxiliary members of sufficient importance, the products of the forces and extensions (or forces and compressions), equal to (F^2/S). Summate these quantities, all of which are positive, whether the forces act as tension or as compression. For example, if $F = 50$ Ton and $S = 500$ Ton/in. for one of the members, the summation is of quantities such as

$$F^2/S = (50.0 \text{ Ton})^2 \div (500 \text{ Ton/in.}) = 5 \text{ ins. Ton.}$$

(4) Apportion, among the joints of the structure, all the masses that are to be taken into account; viz., all that vibrate with the structure. If the object be to calculate the frequency of the complete bridge carrying a definite live-load, the unsprung portion of this live-load as well as the masses in the main-girders, bracing, roadway, etc., will be included. If the object be to calculate the "girder frequency", $n_G = K_G . n_L$, only the masses of the girder will be included.

Proceed to evaluate, for all the loaded joints of the structure, the products WD^2. That is, multiply the load at each joint by the square of its displacement. Summate these products for all joints, giving a total $\Sigma(WD^2)$; which may be expressed in tons-in.²

(5) Deduce the natural frequency, n, from the quotient of summations (3) and (4); introducing $g = 32.19 \times 12$ in./sec.² as below:

$$(2\pi n)^2 = g\Sigma(F^2/s) \div \Sigma(WD^2).$$

For example: If $\sum(F^2/S) = 153$ ins. Ton
and $\sum(WD^2)^\circ = 64$ Ton in.²

$$\text{then } (2\pi n)^2 = 32.19 \times 12 \frac{\text{ins.}}{\text{sec.}^2} \times \frac{153 \text{ ins. Ton}}{64 \text{ Ton in.}^2} = \frac{920}{\text{sec.}^2}.$$

Whence $n = 4.84$/sec.

The value of n deduced at this stage may be very accurate if the system of imaginary loads assumed in stage (1) is approximately suitable. In any case, even if the imaginary system is but moderately appropriate, the result will show whether the value of n lies near to or well above the zone of probable applied frequencies. In the latter case, no further attention need be given to the problem of resonance.

(6) If a close determination of the natural frequency is required, the validity of the first imaginary loading may be gauged by comparing it with a modified system deduced as follows: To each joint of the structure apply an imaginary load acting in the direction of the displacement calculated in stage (2) above, and of a magnitude proportional to the product WD.

If the two imaginary loadings are unduly different, a revised calculation may be carried out with the modified loading, leading to a slightly different value of n. If any important difference in frequency is found, the process may be carried a step further, taking a revised imaginary loading based on the second displacement diagram, and leading to a still closer approximation; but as a rule the result obtained on completing the second calculation will differ only slightly from the first.

At the outset, judgment will usually indicate whether the circumstances require a first quick trial with a simple imaginary load, followed by a more detailed second calculation; or admit of a detailed initial calculation with a system that is likely to give an almost accurate result. In special cases, *e.g.*, very large cantilever bridges, or structures having distinctive features, *e.g.*, opening bridges with balance weights, two stages of calculation may commonly be necessary until experience has indicated the probable character of an imaginary load suitable for use in one operation. A sketch of the flexed structure is always of considerable assistance in choosing an appropriate imaginary load; and when this sketch can be drawn with confidence, an appropriate load may be ascertained as follows: To each joint apply a load acting in the direction of its displacement, and proportional to the product of this displacement and the allocated mass.

Theory

The theory underlying the above method is, probably, sufficiently indicated by the form of the equation employed in stage (5) above. The fundamental principle is that of the Conservation of Energy in undamped natural harmonic vibration.

When an elastic structure stands at rest in a position of *stable* equilibrium, it can be flexed from that position only if an additional quantity of energy is imparted to it, by forces applied by external means—(An application of the Principle of Minimum Strain-energy). If the structure be allowed to return freely to its position of stable equilibrium, this additional elastic strain-energy is converted to kinetic energy; and its momentum carries it beyond the position of stable equilibrium, about which it continues to oscillate.

To justify the assumption that the additional strain-energy in the flexed position is equal to the kinetic energy in the position of stable equilibrium, certain conditions must be fulfilled as follows:

(*a*) The vibration should occur freely, without frictional resistance. Actually, however, considerable frictional resistance may be present without appreciably vitiating the accuracy of the calculated frequency.

(*b*) All parts of the structure should come to rest simultaneously, in the flexed position; and should cross the position of stable equilibrium simultaneously. Actually, this limitation signifies, merely, that only one mode of vibration can be calculated in one operation, a separate calculation being required for each higher mode.

In estimating the two quantities of energy to be equated, great difficulty would be experienced in starting *ab initio* without a preliminary knowledge of the profile of flexure. It is on this account that the approximate method is necessary, in which a probable profile is assumed, and a corresponding imaginary load system used. When the profile and load system are known, to a sufficient degree of approximation, the two quantities of energy are readily evaluated.

Thus, the additional strain-energy is simply $\Sigma(^1/_2 F^2/s)$ in the case of a framed structure, assuming, as usual, that the joints may be regarded as freely pivoted; or $\int\left(\frac{f^2}{2E}\right)dv_f + \int\left(\frac{q^2}{2c}\right)dv_w$ in the case of a plate girder, or of a framed structure when allowance is made for Secondary Stresses. In the latter expression, f and q signify, respectively, the tensile (or compressive) stresses in the flanges, and the shear stress in the webs; E and C are the values of the elastic moduli for tensile and shear stresses; and v_f and v_w are the volumes of metal in the flanges and webs.

The kinetic energy may be expressed

$$^1/_2\int dm\,.\,u^2 = {}^1/_2\int dm\,.\,(2\pi nD)^2$$

where $u = 2\pi nD$ is the velocity with which the element of mass dm crosses the position of stable equilibrium.

Thus the energy equation may be written

$$\Sigma(^1/_2\,.\,F^2/s) = \Sigma(^1/_2 Mu^2)$$

or

$$\Sigma(F^2/s) = (2\pi n)^2\Sigma(WD^2) \div g$$

where W is the weight of the mass M.

Although the application of the "energy balance", in estimating frequencies of natural vibration, entails the presumption of the profile of flexure, no high degree of approximation is essential to ensure a reasonably accurate result. The late Lord Rayleigh pointed out that in calculating the frequency of the primary mode for a uniform beam, no error of importance is introduced if the distribution of the inertia forces is assumed to be uniform instead of in a sine function as is actually the case. Calculations made by the writers show that the latitude of choice is also great in framed structures, and for the higher modes of vibration. For example, in the case of the secondary mode in the seven-panelled Warren girder shown with three alternative loadings in Figure 5, the values calculated for the girder-mode factor are as follow:

(*a*) Assuming that the vibration profile corresponds to the load shown, with only six equal forces, then $K_G = 0.338$.

(*b*) With the carefully revised imaginary load, deduced from the deflections as for load (*a*), then $K_G = 0.336$.

(*c*) Using a still simpler imaginary load than at (*a*), with only two forces (in addition to reactions at the supports) the result is $K_G = 0.328$, so that even in this case the error is only $2\frac{1}{2}\%$, whilst the quite simple consideration as at (*a*) leads to accuracy within 1%.

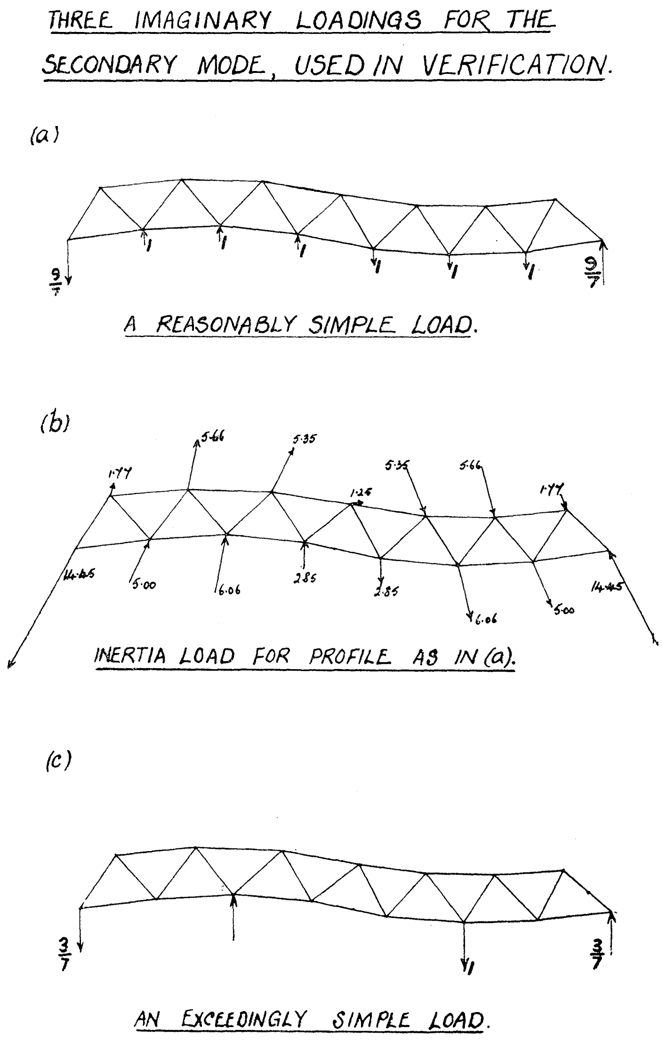


Fig. 5

The considerable latitude permissible in the choice of a suitable imaginary loading justifies a great reduction in the bulk of special calculation required to determine the frequency of *primary* transverse vibration. In the ordinary course of design, stresses and deflections will already have been calculated for one or more distributions of load that flex the structure in a manner approximating to the profile of the primary mode of vibration. Since the stiffnesses of the members and the allocations of the masses to the joints will also have been calculated, nearly all the data required for the calculation of the primary frequency are at hand, and require only multiplication and summation.

To investigate the validity of using such convenient loads as the "imaginary" load, calculations have been carried out for two different loadings on the girder represented in Figure IV, in which all the members have been taken to be of the same section.

The deflections considered were, in the first case, those due to a single load at joint 10; and in the second case, those corresponding to the weights of the members themselves, allocated to the full number of joints. The difference between the two values of the frequency calculated for the two cases was less than 1%.

Since these cases represent extremes, between which must lie the actual inertia forces in vibration, and also, in all probability, the hypothetical static loading which it will be convenient to consider in practice—namely, that for which Maxwell and Williot Diagrams are already drawn—it may be taken as proven that the use of such a hypothetical loading will involve no serious error.

It has been suggested that in calculating the kinetic energy of vibration, the masses of the structure may be regarded as concentrated at the joints. Since the masses are actually distributed, and are subject to rotat onal movements as well as translational, this step requires justification. To provide this, three sets of calculations were carried out for the first type of imaginary load mentioned above; and it was found that the kinetic energy of rotation was no more than1% of that of translation, corresponding to a reduction in the value of n of only $\frac{1}{2}$%. It was found further, that the translational kinetic energy calculated on the assumption that the masses of the members were concentrated at the joints was practically equal to that calculated on the assumption that the masses were concentrated at the centroids. So that the convenient process recommended may be regarded as justified.

The Girder-Mode Factor K_G

While the general method described above is applicable for all modes of vibration, and for structures of all types, the work of calculation is such that the simpler routine described in Section I offers considerable advantages. Since the application of the method of Sect on I requires the use of tables giving reference values of K_G and K_M it is of interest to show how such tables may be constructed as directly as possible.

When the girder frequency, n_G, has been calculated by the general method, taking account of the masses of the girder only, the corresponding value of K_G may be deduced immediately from definition (Section I) thus:

$$K_G = n_G/n_L = 2sn_g\sqrt{w/gE}.$$

In determining a series of values of K_G, for girders of different designs—for example, the series of girders referred to in Figure I—it is convenient to eliminate the use of the quantities g, E and w, by using a modified formula obtained by substituting for n_G and cancelling these terms:

$$K_G = \frac{1}{\pi}\sqrt{\frac{\sum\{v_m e^2/(L/s)^2\}}{\sum(v_j D^2)}} \quad \begin{array}{l}\text{for all members,}\\ \text{for all joints,}\end{array}$$

$$= \frac{1}{\pi}\sqrt{\frac{\sum\{v_m e^2/(L/s)^2\}}{\sum(v_m D_c^2)}} \quad \begin{array}{l}\text{for all members,}\\ \text{for all members.}\end{array}$$

In these alternative expressions, v_m signifies the volume of metal in any member, and v_j the volumes of metal regarded as concentrated at any joint whose displacement is D. In the second expression the mass of each member is regarded as concentrated at its centroid, whose displacement D_c is deduced from the same Williot Diagram as would give the values of D. For example, in Figure IV, the deflection of the centroid of member G, joining joints 4 and 5, is represented in the Williot Diagram by the line Og, where g lies midway between 4 and 5. The symbols e, which represent the extensions of the members, and cause the displacements D and D_c, need not be calculated completely, but may be taken as FL/A, instead of FL/AE; since the value of E influences e and D in the same ratio, and K_G not at all.

The Factor K_M, for Distribution of Mass

The factor K_M signifies, fundamentally, the ratio between the kinetic energy of the load, per unit mass, and that of the metal in the main girders, per unit mass; the displacements of the elements of mass, of both groups, corresponding to those of the actual structure in the mode under consideration.

The value of K_M, for any given distribution of the masses of the main girders and loads, may be determined by means of the general method, reduced to a simpler routine that involves only the use of the Williot diagram. The diagram having been drawn for an imaginary load that sufficiently well represents the mode of vibration (*e.g.*, for the total weight of the structure and live-load, if the mode considered is the primary) the displacements are measured and used in the formula below:

$$K_M = \frac{\dfrac{\sum(MD^2)}{\sum(M)}}{\dfrac{\sum(mD^2)}{\sum(m)}} \quad \begin{array}{l}\text{for all masses, } M\text{, included in load,}\\[2em] \text{for all masses, } m\text{, in main girders.}\end{array}$$

It may be noted that the quotient $\sum(M)/\sum(m)$ is the load ratio R, mentioned in Section I.

In conclusion, attention may be drawn once more to the fact that when values of K_G and K_M become known for typical designs covering the range commonly used in practice, the routine of Section I will render the estimation of natural frequencies for all ordinary bridges strikingly simple, in spite of the initial complexity of the problem.

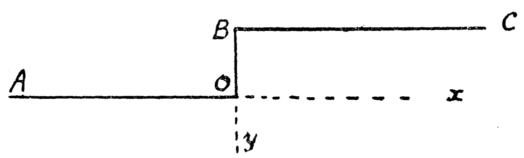
THE ELASTIC EQUIVALENCE OF STATICALLY EQUIPOLLENT LOADS

BY MR. S. D. CAROTHERS,
Officer-in-charge of Works, H.M. Naval Yard, Hong Kong.

According to this principle, the stresses that are produced in a body by the application, to a small part of its surface, of a system of forces statically equivalent to zero force and zero couple are of negligible magnitude at distances which are large compared with the linear dimensions of the part.*

Now it is submitted that in certain problems this principle can be extended to apply to forces statically equivalent to zero force and zero couple applied to an indefinitely large part of a body, provided there is perfect equilibrium in respect to forces and couples at every element except for a small area where equilibrium is imperfect.

Suppose for example that the question refers to an elastic solid of weight w per unit volume, having its upper surface horizontal and at level $y=0$ to the left of the origin and $y=h$ to the right of the origin, as shown in the figure, and it is required to determine the stresses due to gravity only with the surface $AOBC$ free from stress.



It is assumed that uniform normal pressure $=wh$ applied to the plane $y=0$ to the right of the origin will give a first approximation to the stresses for the region $y>0$, for the reason that this system reversed would be in perfect equilibrium with the weight of the material above $y=0$ for all points not too close to the origin.

In the case of a semi-infinite solid having a plane upper boundary solutions for the stresses for certain distributions of normal pressure were given by the present writer in the Proceedings of the Royal Society A, vol. 97, 1920; but it was not realized at the time that the stresses could be expressed in a much more simple manner and in a form more suitable for easy graphical or arithmetical calculation.

* *Vide*, Love, *Elasticity*, 2nd Ed., p. 129.

It was, however, known to the writer that these exact solutions could be each used as a first approximation to the solution of a corresponding problem of importance in practical engineering, and it was thought that the present opportunity should not be missed of bringing these solutions in their simplest form to the notice of engineers on the one hand and mathematicians on the other.

They are brought to the notice of engineers as they afford a simple, if approximate, treatment of some earth problems which are constantly recurring in practical engineering and to the notice of mathematicians with a view to their co-operation in obtaining complete solutions of these problems.

The Elastic Equivalence of Statically Equipollent Loads

There are many problems for which Engineers require solutions and in the absence of these they have sometimes no option but to resort to empirical formulae of some sort and this applies with great force to earth problems. If it is assumed that earth is an elastic solid some of these problems can be partially resolved and in the absence of some such assumption there is no adequate theory either to calculate the stresses or otherwise estimate their magnitude and direction.

It is the object of the present paper to call attention to a few of those problems, chiefly relating to earth, and indicate the nature of the solution which the elastic theory affords.

Case 1. Retaining Walls.

A very common case is a stretch of ground having its surface at a level h bounded on one side by a straight wall resting on ground at a lower level h_1. Such a case may be found at any harbour basin or wet dock. The usual method of treating this problem is to consider the retaining wall as a monolith with pressure on the back face. This pressure is then combined with the weight of the wall and the pressures on the base of the wall are then calculated by various methods and the problem is thus "solved."

Now it is claimed that this misses to a large extent the spirit of the problem. The question at issue is a very important one as failure of such walls as those at the Kidderpur Docks, bears testimony.

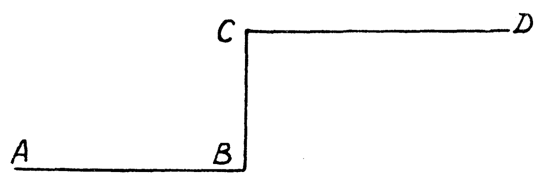


Fig. 1

Suppose the upper surface of the ground is denoted by $ABCD$ as shown in Fig. 1 and for the moment assume that the earth is of sufficient cohesion to retain itself for a short period without any retaining wall. The boundary $ABCD$ is then the free surface and the problem is to obtain a solution of the stress in the solid due to volume forces with the whole of the surface $ABCD$ free from pressure or shear. The points A and D may be considered as $x = \mp \infty$

respectively. The mathematician will readily see along what line a solution may be expected, but he will, I think, as readily admit that the transformations required will at the end probably not be very easily reduced to arithmetic.

If, however, the surface AB produced is taken and it is assumed that to the left of B the surface is free while to the right it is subjected to a uniform pressure $P=(h-h_1)w$, a first approximation to the stresses under AB and AB

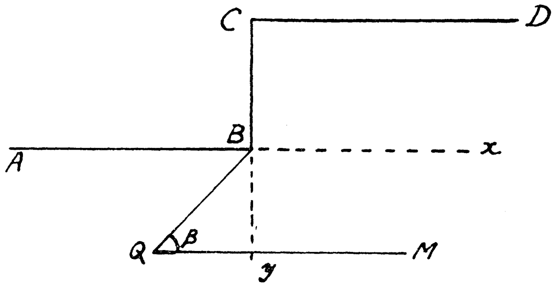


Fig. 2

produced can be obtained. If axes are selected as shown in Fig. 2 and any point Q is taken, QM being drawn horizontal and angle $BQM=\beta$, then it is easily shown that the stresses due to the uniform load on x positive are

$$\widehat{xx}=-P[\beta-xy/r^2]/\pi,$$
$$(1)\qquad \widehat{yy}=-P[\beta+xy/r^2]/\pi,$$
$$\widehat{xy}=-P[-y^2/r^2]/\pi,$$

where

$$r^2=x^2+y^2.$$

Also the directions of the principal stresses are along and at right angles to the bisector of the angle β at every point and their magnitudes are respectively

$$(2)\qquad \begin{aligned} p_1&=-P(\beta+\sin\beta)/\pi,\\ p_2&=-P(\beta-\sin\beta)/\pi. \end{aligned}$$

If now it is assumed that the pressure in the ground is $\widehat{xx}=\widehat{yy}=-wy$ apart from the load on x positive, we can write out the stresses as follows:

$$\widehat{xx}=-P(\beta-xy/r^2)/\pi-wy,$$
$$(3)\qquad \widehat{yy}=-P(\beta+xy/r^2)/\pi-wy,$$
$$\widehat{xy}=-P(-y^2/r^2)/\pi,$$

and the principal stresses are simply

$$(4)\qquad \begin{aligned} p_1&=-P(\beta+\sin\beta)/\pi-wy,\\ p_2&=-P(\beta-\sin\beta)/\pi-wy. \end{aligned}$$

This indicates that at the base of a perpendicular cliff there is a very considerable pressure on vertical planes. The principal stresses cut the vertical plane through the precipice at 45°, bend round under the load and become vertical at $y=0$ for x positive and horizontal for $x=-\infty$. The equation of any line of stress (p_1) is

(5) $$r(1-\cos\beta)=\text{const.}$$

It will thus be seen that all these lines of stress are similar and any radius vector from the origin cuts them all at the same angle. They are shown in the diagram below (Fig. 3).

It will also be easily seen that the pressure on a vertical plane containing the precipice due to the load to the right of the origin produces uniform loading on this plane from $y=0$ to $y=\infty$ and a retaining wall reaching down a few feet does not in any way deal with this general, and what might be described as the permanent stress system in the foundation, in contradistinction to the local and comparatively evanescent stresses due to the retaining wall itself.

For earth in which appreciable tension is absent the assumption made at the beginning cannot be very far from correct and further, the ground at the back of retaining walls is often filled in as loose earth. For the general elastic problem the pressure at the angle B must be considerably higher than the average

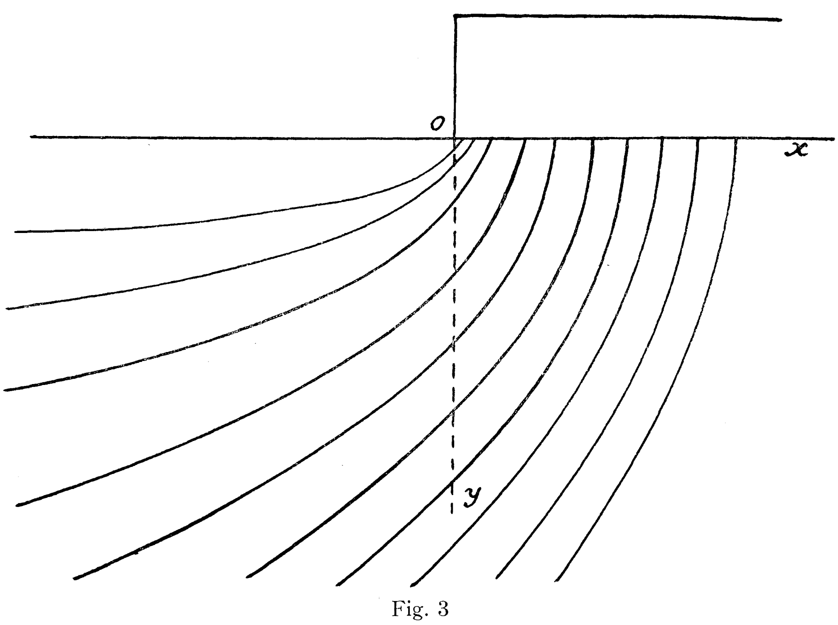


Fig. 3

taken, but this would produce evanescent stresses in the foundation which in no way affects the present argument.

It will thus be seen that where retaining walls are founded in earth there is a definite and, on the elastic theory, easily calculated tendency for the base to slide forward, which should be taken into consideration in their design.

Case 2. The pressure near the foot of a long uniform slope may prove of some interest. Consider the level plane AO produced as loaded to the right of O (Fig. 4) with a load increasing uniformly and select axes as shown. Take

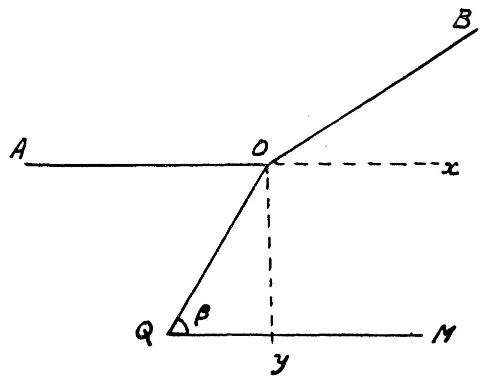


Fig. 4

any point Q and draw QM horizontal. Let angle $OQM=\beta$, then the stresses at Q can be expressed as follows:

$$(6)\qquad \begin{cases} \widehat{xx}=\dfrac{P}{\pi a}\,(-x\beta+2y\log r+y), \\ \widehat{yy}=\dfrac{P}{\pi a}\,(-x\beta-y), \\ \widehat{xy}=\dfrac{P}{\pi a}\,(y\beta). \end{cases}$$

In this case if plane polars are used with O as origin there is no difficulty in obtaining a solution which fits the whole of the boundary AOB giving throughout

$$(7)\qquad \widehat{\theta\theta}=\widehat{r\theta}=0.$$

Case 3. Embankment.

To find the stresses produced in the ground by an embankment it will be convenient to consider a triangular embankment in the first instance.

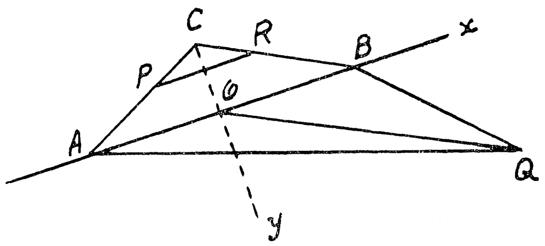


Fig. 5

Let $AQ=r_1$, $BQ=r_2$, $OQ=r_0$, $AO=OB=a$, and angles $AQO=\alpha_1$, $BQO=\alpha_2$. Then the stresses in the ground at any point Q can be written:

$$\widehat{xx} = -\frac{P}{\pi a}[a(\alpha_1+\alpha_2)+x(\alpha_1-\alpha_2)-2y\log(r_1r_2/r_0^2)],$$

(8)
$$\widehat{yy} = -\frac{P}{\pi a}[a(\alpha_1+\alpha_2)+x(\alpha_1-\alpha_2)],$$

$$\widehat{xy} = \frac{P}{\pi a}[y(\alpha_1-\alpha_2)].$$

If it is desired to calculate the stresses for an embankment of the usual form *APRB* as shown in the figure it will only be necessary to calculate the stresses for the two embankments *CPR* and *CAB* and deduct the stresses due to the former from those due to the latter and the work need not be here elaborated.

Case 4. The stresses produced in a semi-infinite solid by loading as shown in Fig. (6) can be written out direct. Axes are selected as shown in the figure and *QM* is drawn horizontal.

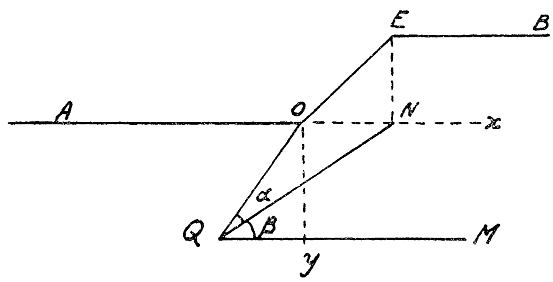


Fig. 6

Let $OQ=r_1$, $NQ=r_2$, $ON=a$, angle $OQN=\alpha$, $NQM=\beta$; then the stresses at Q are

$$\widehat{xx} = -\frac{P}{\pi a}[a\beta+x\alpha+2y\log(r_2/r_1)]$$

(9)
$$\widehat{yy} = -\frac{P}{\pi a}[a\beta+x\alpha]$$

$$\widehat{xy} = \frac{P}{\pi a}[y\alpha].$$

Now if the loading as above were produced by solid earth forming part of an infinite solid with an upper boundary *AOEB* the above stresses would be correct to a first approximation and it will not be difficult to extend this case to meet the next example.

Case 5. Suppose that it is required to find the stresses due to a large cutting as shown in Fig. 7. Let $CN=BM=a$. If $CQ=r_1$, $NQ=r_2$, $BQ=r_1'$, $MQ=r_2'$, $OB=OC=b$ and angles, $CQN=\alpha$, $CNQ=\beta$, $BQM=\alpha'$, $BMQ=\beta'$ and AM and DN are perpendicular to the horizontal, then the stresses at any point Q can be written as follows:

$$\widehat{xx} = -\frac{P}{\pi a}\left[a(\beta+\beta')-b(\alpha+\alpha')+x(\alpha-\alpha')+2y\log(r_2r_2'/r_1r_1')\right],$$

(10)
$$\widehat{yy} = -\frac{P}{\pi a}\left[a(\beta+\beta')-b(\alpha+\alpha')+x(\alpha-\alpha')\right],$$

$$\widehat{xy} = \frac{P}{\pi a}\left[y(\alpha-\alpha')\right].$$

For a point immediately under C we have

$$\alpha'=0,\ \alpha=\pi/2,\ \beta=\beta'=0,\ x=b,\ y=0$$

and thus $\widehat{xx}=\widehat{yy}=\widehat{xy}=0$.

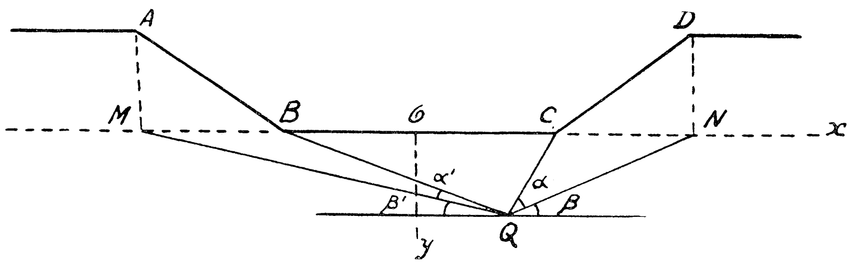


Fig. 7

For any point in the central plane $x=0$,

$$\widehat{xx} = -\frac{P}{\pi a}\left[2a\beta-2b\alpha+4y\log(r_2/r_1)\right],$$

(11)
$$\widehat{yy} = -\frac{P}{\pi a}\left[2a\beta-2b\alpha\right],$$

$$\widehat{xy}=0.$$

Case 6. Consider a masonry dam as shown, of Section $ABCD$, having a front straight face AB subjected to water pressure to a height AR and suppose it is required to obtain the stresses at any point Q for example. At Q draw a tangent to the curved back face meeting the front face in O; with OQ as radius draw a circular arc QP. Now the stresses at the section PQ will only be affected to a very slight extent by the precise shape of the dam at a distance from the arc PQ, and we may consider the dam as formed of a loaded triangle OPQ so far as the stresses at the arc PQ are concerned.

The stresses on the loaded triangle can be resolved into:

(*a*) Uniform normal pressure proportional to OR over OP.

(*b*) Pressure increasing uniformly as the distance from O.

(*c*) A force acting at O of amount X made up of the weight of the masonry in $BCQO$ and the pressure of the water over OR and finally

(*d*) a moment about O of amount M.

Now (*a*), (*b*), (*c*) and (*d*) can be very easily obtained by graphical methods and the stresses at *PQ* can then be written out by well known methods fitting all the conditions at the front and flank of the dam.

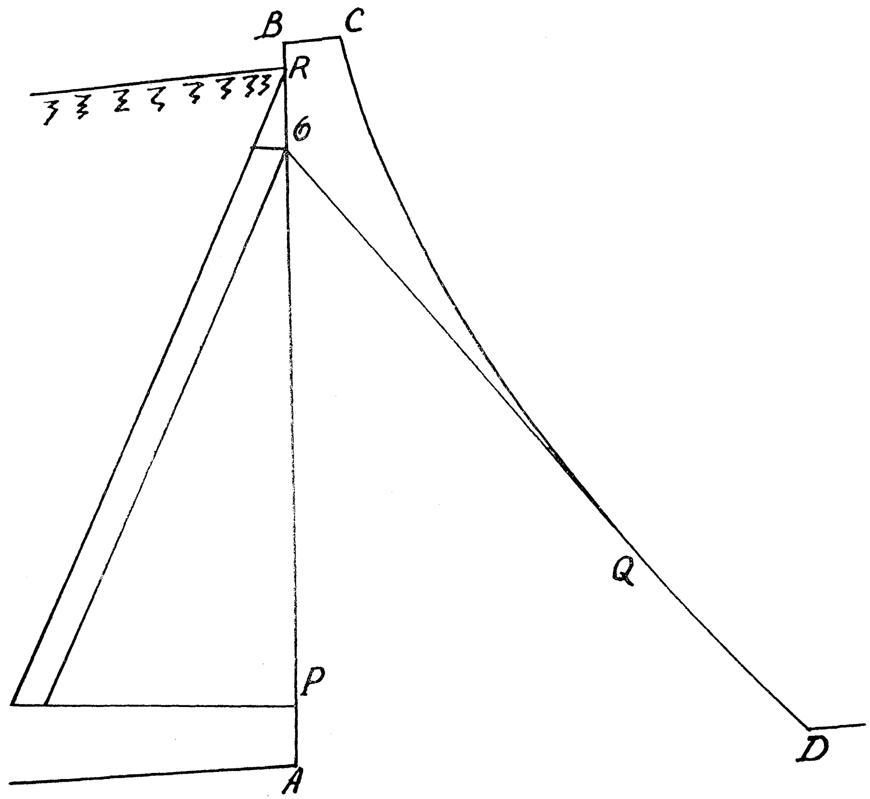


Fig. 8

The present writer has developed this by means of a process somewhat akin to the transference theorem in Basset's *Hydrodynamics*, Vol. I, p. 240, and the result is to show that this method based on equipollent loads leads in effect to Δ being very nearly harmonic, instead of being quite so.

TEST LOADS ON FOUNDATIONS AS AFFECTED BY SCALE OF TESTED AREA

By Mr. S. D. Carothers,
Officer-in-Charge of Works, H.M. Naval Yard, Hong Kong.

Introduction

The object of a former paper* (of which this is a sequel), and indeed of this paper, is to bring to the notice of engineers that in the case of a class of problems falling under the general heading of plane strain the stresses as found by the method of elasticity are independent of the elastic constants and that the study of this class both theoretically and practically is likely to afford the best line of approach to the study of earth problems in general.

The fact that the stresses in certain cases are independent of the elastic constants becomes more in accord with common sense when it is remembered that in such cases the displacements involve these constants and the difference in the behaviour of two substances with equal loads under similar plane strain conditions is made up by the difference in the movements. These movements if carefully analyzed might afford a means of determining the elastic constants in simple cases.

It is well known that the study of earth problems is beset with difficulties and the same remarks might be held to apply, in another sense, in respect to problems in elasticity. In elasticity there is sometimes very great difficulty in obtaining a solution which fits all the conditions of the problem in hand and even when all the conditions are met correctly the result may be so complicated that it is of little or no service in practical engineering and thus there has grown up a tendency to resort to all sorts of empirical formulae which by endless iteration become "current coin" and are accepted as correct whereas in reality they are of all degrees of accuracy or inaccuracy as the case may be.

The theory of elasticity when applied to elastic solids generally involves, at some point of the solution, the introduction of the well known and much discussed elastic constants λ and μ and one of the difficulties of applying the theory to earth is that the numerical values of these constants have not been found even for any of the well known varieties of earth. There is the further difficulty that in earth there is such an endless variety with all possible shades and gradations of difference that the obtaining and tabulation of the elastic constants, assuming that such constants can be found, would be a very formidable undertaking.

*Engineering, London, 4th July, 1924.

There is, however, the further consideration that there are many earth problems with which engineers are confronted for which apart from elasticity no adequate theory exists and as some solution is imperative the choice is between empirical methods and the more logical solution. It has also to be remembered that the treatment of earth problems apart from adequate theory, is likely to lead to as many solutions as there are varieties of earth and the experimental study of such problems apart from some theoretical basis is almost sure to be aimless and lead to considerable waste of time and money.

It will thus be seen that while the difficulties are serious the inducements to obtain a satisfactory solution seem to be still greater remembering that every engineering work of magnitude is affected directly or indirectly and that the question cannot be evaded even if we would.

If it could be definitely proved that the elastic solution for the stresses is correct when these do not involve the elastic constants, a very great impetus to the study of earth problems would have been made. It is almost self evident that if for a given problem the stresses as found by the elastic method are independent of the elastic constants the results should apply to a much greater variety of materials than if the stresses involved constants which strictly apply only to special material for which the particular values of the constants are valid.

The paper discusses the stresses in ground due to uniform pressure on a long narrow strip with parallel edges and also the stresses due to uniform pressure over a circular area and for which a new solution of great simplicity is given in the appendix.

A further object of the present paper is a discussion of the conditions which are necessary and which should be fulfilled as a preliminary to the application of any solution in plane strain to earth.

The fact that the paper discusses the question of test loads on foundations as affected by scale in some detail and as a main theme will it is hoped not be considered inappropriate when it is remembered that this would be a necessary preliminary to any practical tests which might be carried out with a view to the determination of the accuracy or otherwise of the general theory of elasticity as applied to earth.

In conclusion it should be stated that the problems here dealt with have been suggested to the writer by the Civil Engineer-in-Chief at the Admiralty, London, and it is hoped that this paper will lead to a discussion which will be of service to the cause of progress in engineering generally.

Test Loads

It seems to be more or less tacitly accepted that if a test load of a given intensity is just successfully carried on a small test area (A) a load of similar intensity ought to be successfully carried on any larger area (B). In other words and to be more exact if an area A is able to carry successfully a maximum uniform load of A.P. tons then it is more or less generally accepted that any larger area B will carry a uniform load of B.P. tons.

The facts appear to be that while the larger area may in some cases successfully carry the same maximum load per foot super as the smaller area there are other cases where it will certainly fail to do so and the results obtained from small test areas are in such cases misleading. If earth is treated as an elastic solid this question can be dealt with in a reasonably satisfactory manner.

Two special cases will be dealt with in some detail, if space permits, and these are taken for the reason that they are both important practical cases and at the same time they are the simplest cases known to the writer. They are:

(*a*) the stresses in a semi-infinite solid, having a plane upper surface, due solely to uniform normal pressure applied to a long narrow strip with parallel edges.

(*b*) the stresses in a similar solid due solely to uniform normal pressure applied over a circular area.

In each case the stresses will be given under the following conditions:

(1) when the solid extends to infinity without change in character;

(2) when at a depth h below the surface there is a horizontal layer of frictionless material;

(3) when at a depth h below the surface there is a horizontal surface on which the value of the friction is not restricted.

Both cases (*a*) and (*b*) have been treated formerly, (*a*) by Professor J. H. Michell and (*b*) by Boussinesq, but the method used in the present paper will enable the problem to be solved under each subdivision and the stresses written out direct by the method of images in a particularly simple manner as was done in *Engineering* (London) of July 4th, 1924 by the present writer for certain other cases. Indeed this paper is intended as a sequel to that just referred to. Every effort, however, will be made to make the present paper self contained and at the same time repetition will be avoided as far as possible.

Before treating the above cases in detail it will be necessary to fix on criteria for stability as applied to earth. In this connection it may be permitted to state that this is as necessary in the case of a theoretical investigation as mechanical tests are in the case of any practical investigation. This depends on the very simple statement that you cannot get out of any investigation more than is put into it and this applies whether in the domain of theory or experiment. The criteria decided on may be erroneous, inadequate or otherwise faulty, as indeed may be the mechanical tests, but their necessity in any case cannot be doubted.

Theories of Rupture

There are various theories of rupture as applied to elastic solids and a synopsis of these will be found in Love's *Treatise on the Mathematical Theory of Elasticity*, 2nd Edition, p. 117. These might be very briefly enumerated as follows:

(*a*) Greatest extension to be below a certain limit,
(*b*) Greatest tension to be below a certain limit,
(*c*) Greatest stress hypothesis,

(*d*) Greatest stress-difference to be below a certain limit,

(*e*) Greatest shear to be below a certain limit;

to these might be added

(*f*) Rankine's criterion of stability as applied to loose granular material.

Now in the case of earth it is quite clear that both tension and extension must be insensible as earth could not sustain a tension of any appreciable amount and this disposes of (*a*) and (*b*). As regards (*c*) Professor Love's remarks appear to be very much to the point. He indicates that if it were proposed that rupture would take place if the greatest stress exceeds a certain limit then the experiments of A. Föppl on bodies subjected to very great pressures uniform over their surface would be very important as it appeared that rupture could not be produced by such pressures as could be applied. This statement is in strict accordance with common sense and the experience of practical engineers as it is perfectly well known that the most unstable materials, such as quicksands, if confined in all horizontal directions can be made to sustain very great vertical loads.

There remain (*d*), (*e*) and (*f*) and it will be convenient to consider the latter first. According to Rankine's theory as applied to loose granular material it is the ratio of the principal stresses (p_1/p_2) which constitutes the vital factor. Now the ratio of two dissimilar pressures p_1 and p_2 would remain the same if we multiplied each by a factor m, and however large m might be Rankine's criterion would remain unaffected whereas the stress difference and the maximum shear would be each increased in the ratio of 1 to m. It can hardly be denied that a substance might easily stand a shear of $(p_1-p_2)/2$ and fail under a shear of $m(p_1-p_2)/2$ if m is large. It will thus be seen that for a substance such as earth obviously weak in shear the ratio of the principal stresses affords no safe criterion of stability.

There remain then (*d*) and (*e*) and in a very large measure it is a matter of indifference whether we take as a criterion of stability one or other of these as they are to a large extent interdependent.

The writer submits that applied to earth the necessary conditions can be fairly reduced to two:

(*a*) absence of tension everywhere,

(*b*) the maximum stress difference at any point to be below a certain limit.

If, then, the solution of any earth problem by the elastic theory indicates that tension exists, or that the stress difference is greater than a maximum value Q, say, instability will be taken as proved. In testing for instability it will always be necessary to take into account the original stresses in the ground due to the weight of the foundation.

It has been shown that in the case of earth formed by gradual accretion in horizontal layers the stresses due to the weight of the earth alone can be expressed as follows:

$$\widehat{xx} = -\lambda wy/(\lambda+2\mu),\qquad (1)$$
$$\widehat{yy} = -wy,$$
$$\widehat{xy} = 0,$$

where $\widehat{xx}$ is the pressure on any vertical plane at a distance y below the surface, and $\widehat{yy}$ is the pressure on any horizontal plane at the same depth. In this case there is absence of tension everywhere and the stress difference due to the weight of the foundation alone is

$$\widehat{yy}-\widehat{xx}=-2\mu wy/(\lambda+2\mu)=-mwy \text{ (say)}. \tag{2}$$

There must be many cases where the stress-difference is negligible as for example where the earth has been formed from a semifluid impalpable mass and has become solidified by gradual degrees from the top downwards. In such cases the material in the semifluid mass must be under hydraulic conditions during the process of solidification.

There would be other cases where a layer of earth had once been overlaid with strata now denuded and in such a case it is conceivable that the pressure on vertical planes is greater than that on horizontal planes. The case of slate formation affords a conspicuous example of strata where the pressure on vertical planes is much greater than that on horizontal planes.

It will therefore be convenient to express the stress-difference in the original ground due to gravity as

$$\text{stress-difference}=-mwy$$

where m may vary from zero to a considerable positive or negative fraction, depending on the relative values of λ and μ, the history of the foundation and other causes. The following results of which proofs will in some cases be found in the appendix are stated here in order.

Surface Loading Only

Case (*a*). A Long Narrow Strip with Parallel Edges Uniformly Loaded (Breadth $=2a$)

(1) The principal stresses at any point Q due solely to uniform normal pressure applied to a long narrow strip with parallel edges are:

$$p_1=-P(\alpha+\sin\alpha)/\pi,$$
$$p_2=-P(\alpha-\sin\alpha)/\pi, \tag{3}$$

where α is the angle subtended by the width AB $(=2a)$ of the strip at the point Q.

(2) The directions of the principal stresses at the point Q bisect the angles made by the lines QA and QB.

(3) As α is always greater than $\sin\alpha$ there is absence of tension everywhere.

(4) The stress difference

$$p_1-p_2=-2P\sin\alpha/\pi \tag{4}$$

everywhere.

(5) The stress difference is thus constant when α is constant and as α is constant along any circular arc passing through the points A and B the lines of constant stress difference are arcs of circles passing through these points.

(6) The stress difference is a maximum when sin α is a maximum and as sin α is a maximum for $\alpha=\pi/2$ the arc of maximum stress difference is the semicircle $r=a$ having AB for diameter, or

$$\text{maximum stress-difference} = -2P/\pi, \tag{5}$$

which for rough approximate purposes might be taken as

$$-2P/3. \tag{6}$$

Combined Surface Load and Weight of Foundation

(7) The principal stresses in the original ground solely due to the weight of the earth are

$$\begin{aligned} \widehat{xx} &= -\lambda wy/(\lambda+2\mu), \\ \widehat{yy} &= -wy, \\ \widehat{xy} &= 0. \end{aligned} \tag{7}$$

The stress difference is $2\mu wy/(\lambda+2\mu)$ and it can be shown that the maximum stress difference for the central plane occurs at the point

$$y=a\,.\sqrt{\frac{m}{n}-1-\sqrt{\frac{m^2}{n^2}-4\frac{m}{n}}} \tag{8}$$

where

$$m=P/\pi,\ n=\mu wa/(\lambda+2\mu).$$

For large values of m/n this gives a maximum stress difference at a point slightly outside the circle $r=a$. For the value $m/n=4$ the maximum stress difference occurs at the point $y=\sqrt{3}\,.a$, and for values of $m/n<4$ there is no maximum, the difference continually increasing with increase of y.

It will perhaps best illustrate the point now under consideration, if we assume an earth foundation $60'-0''$ deep and having various widths of strip loaded with one ton per foot super. Working out the stress difference for various cases it will be seen that for quite narrow strips the maximum stress difference in the central plane occurs near the surface whereas in the case of much wider strips the maximum stress difference occurs at deeper levels; when the width of the strip becomes 120 feet or twice the thickness of the earth the maximum stress difference in the central plane occurs at the earth foundation level and the whole thickness of the foundation is subjected at every level to the full maximum at some point, whereas with the narrower strips only a portion of the thickness is subjected to this maximum.

(8) If, however, it is assured that there is a small stress difference in the original foundation increasing with depth it will be found that the stress difference throughout is considerably modified.

For the narrow strips the stress difference near the surface and at the bottom of the earth foundation may be subjected to nearly equal stress difference with an intervening region under considerably less stress difference.

For wider strips of say 20′ or thereabouts, the stress difference is nearly constant for any width from $y=a$ downwards, but in general the stress difference increases considerably with increase in width largely due to the fact that the falling off of that difference due to the load is not serious till much greater depths than 60 feet are attained.

The point above referred to will be illustrated by the following tables:

TABLE I

Depth in feet	Stress difference due to load only, $P=20$ cwt., Width $=2a$.				
	$2a=5$ ft.	$2a=10$ ft.	$2a=20$ ft.	$2a=40$ ft.	$2a=60$ ft.
$2\frac{1}{2}$	12.73	$\overline{10.19}$	$\overline{5.99}$	$\overline{3.13}$	$\overline{2.10}$
5	10.19	12.73	$\overline{10.19}$	$\overline{5.99}$	$\overline{4.13}$
10	5.99	10.19	12.73	$\overline{10.19}$	$\overline{7.64}$
20	3.13	5.99	10.19	12.73	$\overline{11.75}$
30	2.10	4.13	7.64	11.75	12.73
40	1.58	3.13	5.99	10.19	12.22
50	1.27	2.52	4.90	8.78	11.23
60	1.06	2.10	4.13	7.64	10.19

These stress differences all refer to the central plane under the loaded strip and those differences marked with a bar are not maximum values for the corresponding depth. The corresponding maximum value in each case is, of course, $2P/\pi=12.73$ at the points $x=\pm\sqrt{a^2-y^2}$.

TABLE II

Depth in feet	Stress difference in ground (assumed)	Stress difference due to a surface load of $P=20$ cwts. per foot super. and the original stress difference in ground				
		$2a=5$ ft.	$2a=10$ ft.	$2a=20$ ft.	$2a=40$ ft.	$2a=60$ ft.
$2\frac{1}{2}$	0.45	13.18	$\overline{10.64}$	$\overline{6.44}$	$\overline{3.58}$	$\overline{2.55}$
5	0.91	11.10	13.64	$\overline{11.10}$	$\overline{6.90}$	$\overline{5.04}$
10	1.82	7.81	12.01	14.55	$\overline{12.01}$	$\overline{9.46}$
20	3.64	6.77	9.63	13.83	16.37	$\overline{15.39}$
30	5.46	7.56	9.59	13.10	17.21	18.19
40	7.28	8.86	10.41	13.27	17.47	19.50
50	9.10	10.37	11.62	14.00	17.88	20.33
60	10.90	11.96	13.00	15.03	18.54	21.09

These stresses again refer to the central plane only and those stresses marked with a bar are again not maximum stress differences for the corresponding depth below the surface. The maximum values, if such exist, would vary from 12.73 at $y=0$ to the maximum values given in each case for $y>a$.

If the earth is just capable of sustaining a stress difference of 15 cwts. per foot super it will be found that the loads which can be carried by the various widths of strip are:

5′ strip = 22.7 cwt. per foot super.
10′ strip = 22.1 cwt. per foot super.
20′ strip = 20.0 cwt. per foot super.
40′ strip = 10.7 cwt. per foot super.
60′ strip = 8.0 cwt. per foot super.

(9) If in equation (7) $\widehat{yy} < \widehat{xx}$ it is easily seen that the maximum stress difference does not occur in the central plane of the loaded strip; it is transferred to the edges of the loaded area at A and B.

Semi-infinite Solid with a Horizontal Frictionless Plane

(10) If a frictionless layer occurs at a depth h below the surface the stresses at this plane, solely due to surface loading, are greatly modified and it will be shown that $\widehat{xx}$ and $\widehat{yy}$ are tension and compression respectively and equal in magnitude at every point of this surface ($y=h$.) The actual stresses at any point on the plane $y=h$ are

$$(9) \qquad -\widehat{xx} = \widehat{yy} = -\tfrac{1}{2}P\left[\tanh\frac{x+a}{h}\frac{\pi}{2} - \tanh\frac{x-a}{h}\frac{\pi}{2}\right]$$

which at the point $x=0$ becomes

$$(10) \qquad -\widehat{xx} = \widehat{yy} = -P\left[\tanh\frac{a}{h}\frac{\pi}{2}\right],$$

giving

$$(11) \qquad \text{Maximum stress difference} = -2P\tanh\frac{a}{h}\frac{\pi}{2},$$

$$= -2P \text{ (approx.) if } h<a,$$
$$= -1.84P \quad \text{" if } h=a,$$
$$= -0.65\,P \quad \text{" if } h=2a.$$

For the present purpose it will be sufficient to examine the case where $h=a$. In this case the stresses are

$$(12) \qquad \begin{aligned} \widehat{xx} &= 0.92P - \lambda wy/(\lambda+2\mu) \text{ at } y=h, \\ \widehat{yy} &= -0.92\,P - wy, \\ \widehat{xy} &= 0. \end{aligned}$$

To avoid tension it is easily seen that

$$(13) \qquad P \lessdot 1.08\lambda wy/(\lambda+2\mu)$$

and maximum stress difference $= 1.84P - 2\mu wy/(\lambda+2\mu)$.

A little consideration will show that for moderate values of P it will be difficult, in this case, to avoid tension and at the same time retain the maximum stress difference within reasonable limits.

(11) It is worth noting that immediately under the edges of the loaded strip the stress $\widehat{yy}$ is precisely $-P/2$ and this value of $\widehat{yy}$, due to surface loading only, persists with quite negligible diminution right down to the plane $y=h=a$. For larger values of h the pressure falls off somewhat faster as may be seen from the following:

$$(14) \qquad \begin{aligned} \widehat{yy} &= 0.500P, \text{ for } y=0, \text{ at } x=\pm a, \\ \widehat{yy} &= 0.498P, \text{ for } y=a, \text{ at } x=\pm a, \\ \widehat{yy} &= 0.458P, \text{ for } y=2a, \text{ at } x=\pm a. \end{aligned}$$

Under the edges of a loaded strip $y=h=a$ the stress difference may thus be taken as $-P$, due to the surface load but this is not of great importance.

(12) This particular case affords a very good example of what Professor Michell has insisted on that stresses in such circumstances as are here assumed are transmitted laterally to a perfectly negligible extent.

For example if $y=h=a$

$$\begin{aligned} -\widehat{xx}=\widehat{yy}&=-0.92P \text{ at } x=0,\ y=h,\\ -\widehat{xx}=\widehat{yy}&=-0.82P \text{ at } x=\pm a/2,\ y=h,\\ -\widehat{xx}=\widehat{yy}&=-0.50P \text{ at } x=\pm a,\ y=h,\\ -\widehat{xx}=\widehat{yy}&=-0.04P \text{ at } x=\pm 2a,\ y=h. \end{aligned} \tag{15}$$

(13) In the case of an earth foundation resting on rock at a depth $y=h$ below the surface with no restriction on friction at the rock surface the stresses $\widehat{xx}$ and $\widehat{yy}$ are shown to be equal in magnitude and sign. In this case if $\bar{a}$ denotes the algebraic sum of a series of angles, which will be defined in the appendix, it is shown that at the surface $y=h$:

$$\begin{aligned} \widehat{xx}=\widehat{yy}&=-2P\,.\,h\,.\,\bar{a}/\pi,\\ \widehat{xy}&=(2/\pi)\,.\,P\,.\,h\,.\,\partial\bar{a}/\partial x. \end{aligned} \tag{16}$$

In this case the series for $\partial\bar{a}/\partial x$ when set out is far from being satisfactory from the point of view of convergence and it is necessary to sum it in a form for easy calculation. The result I have obtained is

$$\widehat{xy}=(P/2)\,.\left[1/\cosh\frac{\pi}{2}\frac{x+a}{h}-1/\cosh\frac{\pi}{2}\frac{x-a}{h}\right] \tag{17}$$

but this will be discussed in the appendix. If $h=a$ this has a maximum in the neighbourhood of $(x+a)/h=2.06$, $(x-a)/h=0.06$ or at a value of x slightly greater than a. The value of $\widehat{xy}$ at this point may for practical purposes be taken as $\widehat{xy}=-P/2$. The stresses have been roughly calculated for $x=\pm a$, $y=h=a$, and if combined with the weight are:

$$\begin{aligned} \widehat{xx}&=-0.45P-\lambda wy/(\lambda+2\mu),\\ \widehat{yy}&=-0.45P-wy,\\ \widehat{xy}&=-0.50P, \end{aligned} \tag{18}$$

giving

$$\text{Stress difference}=\sqrt{P^2+[2\mu wy/(\lambda+2\mu)]^2}\geqq P. \tag{19}$$

If it is assumed that the stress on horizontal planes is greater than that on vertical planes the following deductions can be made:

1. The narrower the strip the greater the load which can be carried per foot super on any given foundation.

2. If at a depth $y=h$ there is a layer of frictionless material it will be necessary if $h=a$ to limit the load on the surface to a very small amount in order to avoid tension and retain the stress difference within reasonable limits. If $h>$ or

$<a$, h and a being of the same order of magnitude, the same remarks would apply.

3. If test holes or bore holes indicate that the quality of the ground becomes softer with depth then the wider the strip the less average load per foot super will it carry.

4. If the pressures on horizontal and vertical planes are equal at all points then all widths of strip should carry the same intensity of loading, provided the ground does not vary with depth.

5. If the pressures on vertical planes is greater than that on horizontal planes then the maximum stress difference occurs at the edges of the loaded strip and again all widths of strip should carry the same intensity of loading provided the ground does not vary with depth.

6. If the conditions are as stated in 1 above, then these could be materially improved by driving piles of length somewhat greater than a or alternatively driving skeleton piles and filling up the spaces with earth.

7. If at a depth $y=h$ there is a rock surface where the friction is not restricted in value then it appears that the stress difference at the rock surface is a maximum at points $\pm x$ slightly greater than a, and if $h=a$ this stress difference appears to be greater than P and under the present theory earth standing on rock would present a less favourable foundation than the same earth extending indefinitely. This at first sight seems to traverse the common sense of the situation but it is not altogether inconsistent with the fact that a soft material is more easily crushed if placed between two hard substances. In any case it is hoped that this will lead to a good discussion.

Conditions Necessary for Similarity

These can be investigated in a very simple and general way and the results will apply equally whether there is a stress difference in the ground originally or otherwise.

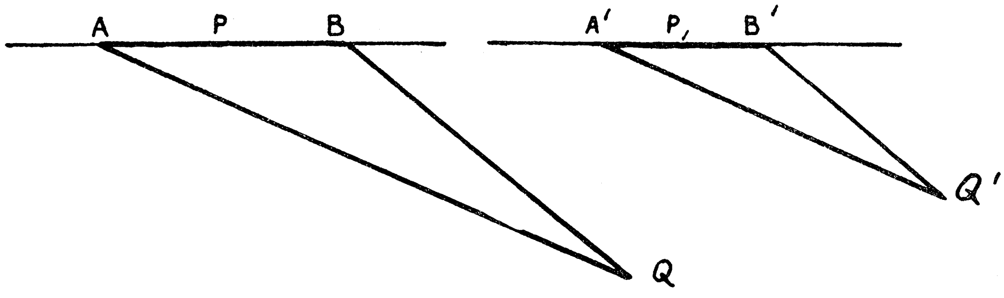


Fig. 1

Consider two entirely independent strips AB and $A'B'$ having uniform loads P and P_1 per foot super and of widths $2a$ and $2a_1$, respectively.

Take the points Q and Q' in the same relative position in respect to AB and $A'B'$ respectively, *i.e.*, angle $QAB=Q'A'B'$ and angle $QBA=Q'B'A'$.

The principal stresses due to the loads on the strips only are as follows:

$$-\frac{P}{\pi}(\alpha+\sin\alpha),\quad -\frac{P}{\pi}(\alpha-\sin\alpha),\quad (\text{at } Q),$$

$$-\frac{P_1}{\pi}(\alpha+\sin\alpha),\quad -\frac{P_1}{\pi}(\alpha-\sin\alpha),\quad (\text{at } Q'),$$

which when $P=P_1$ are precisely equal. Q is however considerably further below the surface than Q' and in respect to the pressure of the earth foundation is subjected to greater loading and owing to similarity of position the depths can be represented by la and la_1 (l being a common factor) as these are proportional to AB and $A'B'$.

We have thus to combine the principal stress at Q and Q', with the stresses due to the depths la and la_1. This need not actually be done as it is clear that the resultant principal stresses would be different at the point Q from those at Q' both in magnitude and direction.

Now replace the loads P and P_1 by loads aP and a_1P respectively and there is obtained for the principal stresses due to the loaded strips

(20)

$$-\frac{aP}{\pi}(\alpha+\sin\alpha),\quad -\frac{aP}{\pi}(\alpha-\sin\alpha),\quad (\text{at } Q),$$

$$-\frac{a_1P}{\pi}(\alpha+\sin\alpha),\quad -\frac{a_1P}{\pi}(\alpha-\sin\alpha),\quad (\text{at } Q').$$

Combining these with the stresses in the foundation due to the depths la and la_1 it is clear that the combined systems have components which are proportional to a and a' and the resultant principal stresses at Q and Q' will be proportional to a and a_1 and have the same directions thus producing similarity but not equality.

Again it might be useful to inquire what intensity of loading (P_1 per foot super) on a narrow strip of width $2a_1$, would produce the same stress difference at a depth D as would be produced by a load of given intensity (P) on a given width ($2a$) at the same depth.

As the depth is the same in each case the stress difference in the ground need not be considered. The quantities P, a, and D are known and one of the quantities P_1 or a_1 may be assumed.

The condition for equal stress difference is

$$\frac{2}{\pi}P\sin\alpha=\frac{2}{\pi}P_1\sin\alpha_1.$$

Now $\sin\alpha=2aD/(a^2+D^2)$ and $\sin\alpha_1=2a_1D/(a_1^2+D^2)$, and thus the condition to be satisfied is that

$$PaD/(a^2+D^2)=P_1a_1D/(a_1^2+D^2),$$

or

$$P:P_1::a_1/(a_1^2+D^2):a/(a^2+D^2).$$

If a is large in comparison to a_1 it is easily seen that P_1 must be large in comparison to P and thus it will be difficult to arrange for practical tests to fulfil the necessary conditions.

Appendix I to Case (*a*)

It has been shown that the stresses produced in two dimensional stress systems can generally be resolved into four distinct types. Only one of these types will be required for the present purpose, it is

$$\widehat{xx}=\phi+y\frac{d\phi}{\partial y},$$

$$\widehat{yy}=\phi-y\frac{d\phi}{\partial y},$$

$$\widehat{xy}=\quad -y\frac{\partial\phi}{\partial x},$$

where ϕ is any plane harmonic function of x and y only, satisfying

$$\nabla_1^2\phi=\frac{\partial^2\phi}{\partial x^2}+\frac{\partial^2\phi}{\partial y^2}=0.$$

If ϕ is suitably chosen it has been shown that the above stress system can be made to represent uniform normal pressure over a long narrow strip with parallel edges.

If the width of the strip is denoted by AB where $AO=OB=a$ and angle $AQB=\alpha$ and axes are selected as shown in the figure, then it has been shown that the stresses at any point Q can be expressed in the form

$$\widehat{xx}=-P\left(\alpha+y\frac{\partial\alpha}{\partial y}\right)/\pi,$$

$$\widehat{yy}=-P\left(\alpha-y\frac{\partial\alpha}{\partial y}\right)/\pi,$$

$$\widehat{xy}=-P\left(\quad -y\frac{\partial\alpha}{\partial x}\right)/\pi,$$

and it is seen at once that as Q approaches any point between A and B, α approaches π, while if Q approaches any point on the surface outside AB, α approaches zero, giving stresses on the surface

$\widehat{yy}=-P$ between A and B,
$\widehat{yy}=0$ outside AB,
$\widehat{xy}=0$ everywhere on $y=0$.

Appendix II to Case (*a*)

Suppose that at a depth $y=h$ there is a thin layer CD of soft slippery material and the surface of the ground over AB is loaded.

Retaining the same scheme of stresses as before we may assume an image A_1B_1 of the load AB in the slippery plane CD.

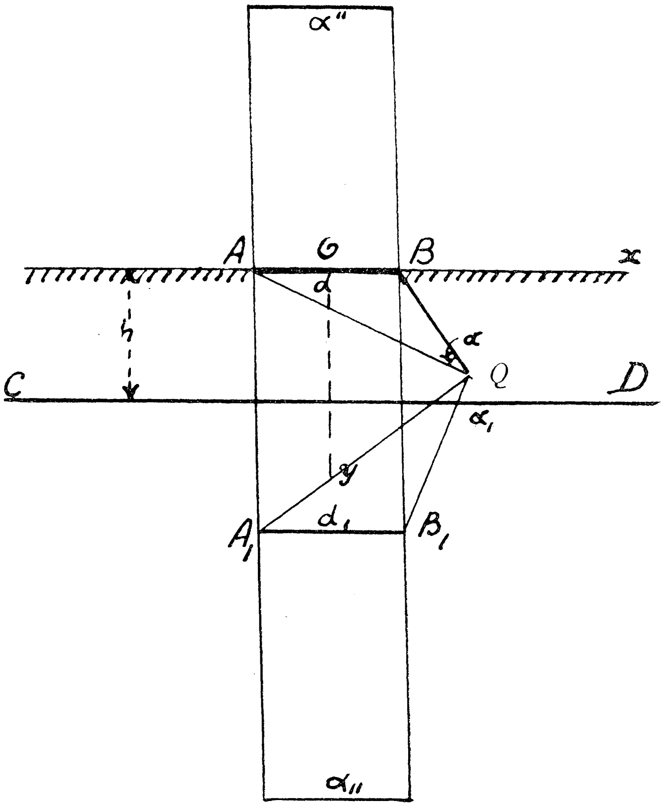


Fig. 2

We can write the stresses

$$\widehat{xx} = -\pi^{-1}P\left[\alpha-\alpha_1+y\frac{\partial}{\partial y}(\alpha-\alpha_1)\right]$$

(22) $$\widehat{yy} = -\pi^{-1}P\left[\alpha-\alpha_1-y\frac{\partial}{\partial y}(\alpha-\alpha_1)\right]$$

$$\widehat{xy} = -\pi^{-1}P\left[\qquad -y\frac{\partial}{\partial x}(\alpha-\alpha_1)\right]$$

Now on any point of the plane CD we have $\alpha-\alpha_1=0$ and therefore $\frac{\partial}{\partial x}(\alpha-\alpha_1)=0$ and thus the shear vanishes on both the planes $y=0$ and $y=h$, but as above written $\widehat{yy}$ gives loading over the whole surface of the plane $y=0$ and another image must be introduced at a height of $2h$ above AB to correct this and so on.

The final result is

$$
(23)\qquad
\begin{aligned}
\widehat{xx} &= -\pi^{-1}P\left[\alpha-\alpha_1+\alpha''-\alpha_{11}+\quad +y\frac{\partial}{\partial y}(\alpha-\alpha_1+\alpha''-\alpha_{11}+\ldots)\right],\\
\widehat{yy} &= -\pi^{-1}P\left[\alpha-\alpha_1+\alpha''-\alpha_{11}+\quad -y\frac{\partial}{\partial y}(\alpha-\alpha_1+\alpha''-\alpha_{11}+\ldots)\right],\\
\widehat{xy} &= -\pi^{-1}P\left[\qquad -y\frac{\partial}{\partial x}(\alpha-\alpha_1+\alpha''-\alpha_{11}+\ldots)\right].
\end{aligned}
$$

At the plane $y=0$ there obtains: $\alpha=\pi$ between the points A and B and $\alpha=0$ outside AB, also $-\alpha_1+\alpha''=0$, $-\alpha_{11}+\alpha'''=0$ at every point, and

$$\frac{\partial\alpha}{dx}=0,\ \frac{\partial}{\partial x}(-\alpha_1+\alpha'')=0,\ \frac{\partial}{\partial x}(-\alpha_{11}+\alpha''')=0, \text{ etc.}$$

giving finally on this plane

$$\widehat{yy}=-P,$$

the remainder of the surface being free from stress. At the surface $y=h$ the stresses are easily seen to be

$$
(24)\qquad
\begin{aligned}
\widehat{xx} &= -2\pi^{-1}Ph\frac{\partial}{\partial y}(\alpha+\alpha''+\alpha'''+\ldots),\\
\widehat{yy} &= +2\pi^{-1}Ph\frac{\partial}{\partial y}(\alpha+\alpha''+\alpha'''+\ldots),\\
\widehat{xy} &= 0.
\end{aligned}
$$

It will be seen that as $\frac{\partial\alpha}{\partial y}$, $\frac{\partial\alpha''}{\partial y}$, etc., are in general each negative, $\widehat{xx}$ the stress on vertical planes at $y=h$ is tension and for earth this would not be permissible, and this example affords an instance of definite instability if tension in earth is considered inadmissible as it must be. Common sense indicates that in this case the pressure cannot be transmitted in the horizontal direction to any extent worth consideration and that the pressure on the plane $y=h$ must be highly concentrated directly under AB. It will be interesting to see how the stresses on this plane figure out numerically as compared to the case where no such slippery plane exists (*vide* Equation 9).

In the above equations the weight of the earth itself has not been taken into account but this can be done by adding to equations (24) the value due to this weight, giving finally

$$
(25)\qquad
\begin{aligned}
\widehat{xx} &= -2\pi^{-1}Ph\frac{\partial}{\partial y}(\alpha+\alpha''+\alpha'''+\quad)-\lambda wy/(\lambda+2\mu),\\
\widehat{yy} &= -2\pi^{-1}Ph\frac{\partial}{\partial y}(\alpha+\alpha''+\alpha'''+\quad)-wy,\\
\widehat{xy} &= 0.
\end{aligned}
$$

9. Suppose that instead of assuming that the plane $y=h$ is a slippery plane it is assumed that the vertical movement at this plane is zero. This would correspond to the case of a comparatively soft layer of earth of depth h resting on rock and carrying a load on its upper surface. Referring to Figure 2 the stresses can be written out at once as follows:

Let

$$\bar{\alpha}=\alpha+\alpha_1-\alpha''-\alpha_{11}+\alpha'''+\alpha_{111}-\ldots,$$

then

$$(26)\qquad \begin{aligned} \widehat{xx}&=-\pi^{-1}P\left[\bar{\alpha}+y\frac{\partial}{\partial y}\bar{\alpha}\right],\\ \widehat{yy}&=-\pi^{-1}P\left[\bar{\alpha}-y\frac{\partial}{\partial y}\bar{\alpha}\right],\\ \widehat{xy}&=-\pi^{-1}P\left[\quad-y\frac{\partial}{\partial x}\bar{\alpha}\right]. \end{aligned}$$

At the surface $y=0$ we have

$$\alpha_1=\alpha'',\ \alpha_{11}=\alpha''',\ \ldots,$$

and thus

$$(27)\qquad \begin{aligned} &\widehat{yy}=-P,\ \widehat{xy}=0 \text{ for the interval } a>x>-a,\\ &\widehat{yy}=\widehat{xy}=0 \text{ for } x \text{ outside the limits } \pm a. \end{aligned}$$

At the plane $y=h$ we have

$$\alpha=\alpha_1,\ \alpha''=\alpha_{11},\ \alpha'''=\alpha_{111},\ \ldots$$

Also

$$(28)\qquad \begin{aligned} &\frac{\partial}{\partial y}(\alpha+\alpha_1)=0,\quad \frac{\partial}{\partial y}(\alpha''+\alpha_{11})=0,\\ &\frac{\partial}{\partial x}(\alpha+\alpha_1)=2\frac{\partial\alpha}{\partial x},\quad \frac{\partial}{\partial x}(\alpha''+\alpha_{11})=2\frac{\partial\alpha''}{\partial x}, \end{aligned}$$

and thus at the surface of the rock $y=h$,

$$(29)\qquad \begin{aligned} \widehat{xx}&=\widehat{yy}=-2\pi^{-1}Ph(\alpha-\alpha''+\alpha'''-\alpha''''+\ldots),\\ \widehat{xy}&=2\pi^{-1}Ph\frac{\partial}{\partial x}(\alpha-\alpha''+\alpha'''-\alpha''''+\ldots). \end{aligned}$$

A few images would give results sufficiently accurate for most practical purposes and it will be seen that $\widehat{xy}$ and $\widehat{yy}$ can be obtained very easily by graphical methods.

To obtain the result shown in Equation 9 we may, starting with Equations 24, proceed as follows:

$$\begin{aligned} \alpha\ &=\tan^{-1}(x+a)/y-\tan^{-1}(x-a)/y,\\ \alpha''&=\tan^{-1}(x+a)/(2h+y)-\tan^{-1}(x-a)/(2h+y),\\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \end{aligned}$$

where

$$\tan^{-1}(x+a)/y \equiv \tan^{-1}\frac{x+a}{y},$$

. . . .

Then

$$y\frac{\partial\bar{\alpha}}{\partial y} = y\frac{\partial}{\partial y}(\alpha+\alpha''+\alpha'''+\ldots\ldots)$$

$$= y(x-a)\left[\frac{1}{(x-a)^2+y^2}+\frac{1}{(x-a)^2+(2h+y)^2}+\frac{1}{(x-a)^2+(4h+y)^2}+\ldots\right]$$

$$-y(x+a)\left[\frac{1}{(x+a)^2+y^2}+\frac{1}{(x+a)^2+(2h+y)^2}+\frac{1}{(x+a)^2+(4h+y)^2}+\ldots\right].$$

At the plane $y=h$ this becomes:

$$(30)\quad h\frac{\partial\bar{\alpha}}{\partial y} = h(x-a)\left[\frac{1}{(x-a)^2+h^2}+\frac{1}{(x-a)^2+(3h)^2}+\frac{1}{(x-a)^2+(5h)^2}+\ldots\right]$$

$$-h(x-a)\left[\frac{1}{(x+a)^2+h^2}+\frac{1}{(x+a)^2+(3h)^2}+\frac{1}{(x+a)^2+(5h)^2}+\ldots\right]$$

$$=\frac{\pi}{4}\left[\tanh\frac{\pi}{2}\frac{x-a}{h}-\tanh\frac{\pi}{2}\frac{x+a}{h}\right],$$

from which the result in the text (Equation 9) is easily deduced.

Again starting from the last of equations (29) we may proceed as follows: letting $\beta=\alpha-\alpha''+\alpha'''-\alpha''''+\ldots$ we have

$$h\frac{\partial\beta}{\partial x} = h\frac{\partial}{\partial x}(\alpha-\alpha''+\alpha'''-\alpha''''+\ldots)$$

$$= hy\left[\frac{1}{(x+a)^2+y^2}-\frac{3}{(x+a)^2+(2h+y)^2}+\frac{5}{(x+a)^2+(4h+y)^2}-\ldots\right]$$

$$-hy\left[\frac{1}{(x-a)^2+y^2}-\frac{3}{(x-a)^2+(2h+y)^2}+\frac{5}{(x-a)^2+(4h+y)^2}-\ldots\right],$$

which at $y=h$ becomes

$$h\frac{\partial\beta}{\partial x} = \left[\frac{1}{\left(\frac{x+a}{h}\right)^2+1^2}-\frac{3}{\left(\frac{x+a}{h}\right)^2+3^2}+\frac{5}{\left(\frac{x+a}{h}\right)^2+5^2}-\ldots\right]$$

$$-\left[\frac{1}{\left(\frac{x-a}{h}\right)^2+1^2}-\frac{3}{\left(\frac{x-a}{h}\right)^2+3^2}+\frac{5}{\left(\frac{x-a}{h}\right)^2+5^2}-\ldots\right].$$

Now it is known that

$$\int_0^\infty e^{-n\lambda}\cos m\lambda\, d\lambda = \frac{n}{m^2+n^2},$$

and thus

$$h\frac{\partial\beta}{\partial x}=\int_0^\infty(e^{-\lambda}-e^{-3\lambda}+e^{-5\lambda}-\ldots)\cos m\lambda d\lambda$$
$$-\int_0^\infty(e^{-\lambda}-e^{-3\lambda}+e^{-5\lambda}-\ldots)\cos m_1\lambda d\lambda$$

where $m=(x+a)/h$ and $m_1=(x-a)/h$.
Again

$$e^{-\lambda}-e^{-3\lambda}+e^{-5\lambda}-\ldots.=e^{-\lambda}/(1+e^{-2\lambda})=1/2\cosh\lambda=\sinh\lambda/\sinh 2\lambda.$$

And thus at $y=h$:

$$h\frac{\partial\beta}{\partial x}=\int_0^\infty\frac{\sinh\lambda}{\sinh 2\lambda}(\cos m\lambda d\lambda-\cos m_1\lambda d\lambda) \tag{31}$$
$$=\frac{\pi}{4}\left[\frac{\sin\frac{\pi}{2}}{\cosh\frac{\pi}{2}m+\cos\frac{\pi}{2}}-\frac{\sin\frac{\pi}{2}}{\cosh\frac{\pi}{2}m_1+\cos\frac{\pi}{2}}\right],$$

from which the result in the text (Equation 17) is deduced.

CASE (*b*). A UNIFORMLY LOADED CIRCULAR AREA OF RADIUS $r=a$

In this case it will be convenient to take cylindrical co-ordinates with the axis of z in the positive direction drawn downwards through the centre of the loaded circle and denote by r_1 the radius drawn horizontally through the z axis, so that

$$r^2=r_1^2+z^2=x^2+y^2+z^2. \tag{32}$$

The stresses can be found either by the general method given in Love's *Treatise*, 2nd edition, p. 190 or by the alternative method applicable to this special case and for which a solution is given hereafter.

The stresses have been calculated in detail from the general method, the various functions being expanded in zonal harmonics. It is found that the stress difference due to the load on the circular area only, omitting any stress difference in the ground, is a maximum at the edge of the loaded circle and this Maximum stress difference $=-0.63\,P=-2P/\pi$ apparently. On the axis of z all the stresses can be expressed in a very simple manner as follows:

$$\widehat{r_1r_1}=\widehat{\theta_1\theta_1}=-\frac{P}{4\pi}\cdot\left[(1+2\sigma)\,\omega+z\frac{\partial\omega}{\partial z}\right],$$
$$\widehat{zz}=-\frac{P}{2\pi}\left[\omega-z\frac{\partial\omega}{\partial z}\right], \tag{33}$$
$$\widehat{r_1z}=0,$$

where ω is the solid angle subtended by the loaded circle at the point, also for any point on the z axis.

$$\omega = 2\pi\left(1-\frac{z}{R}\right), \qquad (R^2=z^2+a^2),$$

(34)
$$\frac{\partial\omega}{\partial z} = -\frac{2\pi}{R}\left(1-\frac{z^2}{R^2}\right),$$

$$\frac{\partial^2\omega}{\partial z^2} = \frac{6\pi z}{R^3}\left(1-\frac{z^2}{R^2}\right).$$

From the above equations it is easily deduced that:

(a) The stress difference for any point on the axis of z is

(35)
$$\widehat{zz}-\widehat{r_1r_1} = -P[(1-2\sigma)\omega - 3z\partial\omega/\partial z]/4\pi.$$

(b) The point where this stress difference becomes a maximum is given by the equation

(36)
$$z^2 = \frac{2+2\sigma}{7-2\sigma}a^2 = \frac{2+2\sigma}{9}R^2.$$

or

(c) The axial stress difference can be expressed in the form

(37) Stress difference $= -\dfrac{P}{2}\left[1-2\sigma+(2+2\sigma)\dfrac{z}{R}-3\dfrac{z^3}{R^3}\right] = -0.5986P$ (maximum) at $z=0.6895a$, if $2\sigma=0.90$.

The above is given in some detail as it is of considerable interest to ascertain with certainty that there is no surface cutting the axis of z having constant stress difference $= -2P/\pi$ as obtains in the case of plane strain. On the surface $r=a$ and $r=2a/3$ the stress differences starting from the z axis are as follows:

	$r=a$	$r=2a/3$
0° stress difference	0.50 P	.56 P
30° " "	0.55 P	.58 P
45° " "	0.57 P	.60 P
60° " "	0.58 P	.55 P
75° " "	0.62 P	.40 P
80° " "	0.63 P	.30 P
85° " "	0.63 P	.18 P
90° " "	0.63 P	.05 P

It will be seen from the foregoing that the surface of maximum stress difference has a bounding edge in the circle enclosing the loaded area and that the maximum stress difference varies from 0.63 P at this edge to 0.60 P at the point $z=0.69a$ (approx.) on the axis of z. The stress differences on the z axis fall off very rapidly with increase of depth as may be seen from the following:

$z=2a/3$	stress difference	0.56 P,
$z=0.69a$	" "	0.60 P,
$z=a$	" "	0.53 P,
$z=\sqrt{2}a$	" "	0.46 P,
$z=2a$	" "	0.27 P,
$z=3a$	" "	0.14 P,
$z=4a$	" "	0.08 P.

(*d*) It may be convenient to note that for any point whatever the general solution leads to the following results:

$$\widehat{r_1 r_1}+\widehat{\theta_1\theta_1}+\widehat{zz}=-\frac{P}{2\pi}(2-2\sigma)\omega,$$

(38)
$$\widehat{zz}=-\frac{P}{2\pi}\left(\omega-z\frac{\partial\omega}{\partial z}\right),$$

$$\widehat{r_1 z}=\frac{P}{2\pi}z\frac{\partial\omega}{\partial r_1},$$

where ω is the solid angle subtended at the point.

(*e*) In respect to the loaded circular area the same general deductions can be made as in the case of a loaded strip with parallel edges and it is clear that the stresses $\widehat{r_1 z}$ and $\widehat{zz}$ can be written out by the method of images in exactly the same way as was done for the loaded strip and these need not be here elaborated.

It might be noted that the calculation of the stresses for the surface $r=a$ is of some interest as the functions are generally not convergent and are all quite unsuitable for calculation in the neighbourhood of this region. It is believed, however, that the stresses for this surface are more accurately stated than for any region except perhaps those relating to $r>2a$ where the functions are all highly convergent.

Appendix to Case (*b*)

STRESSES

	$\widehat{r_1r_1}$	$\widehat{\theta_1\theta_1}$	$\widehat{zz}$	$\widehat{r_1z}$
$r=0$	−.9500	−.9500	−1.0000	.0000
$r=2a/3$				
0°	−.2310	−.2310	−.7904	−.0000
30°	−.2548	−.2625	−.8376	−.0637
45°	−.3344	−.3243	−.8585	−.0529
60°	−.4129	−.4252	−.9062	−.0500
80°	−.6910	−.7213	−.9910	−.0270
85°	−.8134	−.8324	−.9991	−.0071
90°	−.9500	−.9500	−1.0000	.0000
$r=a$				
0°	−.1016	−.1016	−.6466	−.0000
30°	−.1406	−.1121	−.6283	−.1364
45°	−.1903	−.1409	−.6064	−.1980
60°	−.2607	−.1879	−.5769	−.2533
75°	−.3516	−.2676	−.5406	−.2996
80°	−.3852	−.3201	−.5272	−.3075
85°	−.4194	−.3935	−.5127	−.3136
90°	−.4500	−.5000	−.5000	−.3162
$r=\sqrt{2}a$				
0°	−.0265	−.0265	−.4810	.0000
30°	−.0944	−.0430	−.3979	−.1306
45°	−.1337	−.0470	−.3442	−.1599
60°	−.2144	−.0533	−.2117	−.1906
75°	−.2138	−.0498	−.0694	−.1112
80°	−.1564	−.0435	−.0384	−.0579
85°	−.0732	−.0348	−.0164	−.0160
90°	+.0250	−.0250	.0000	.0000

	$\widehat{r_1r_1}$	$\widehat{\theta_1\theta_1}$	$\widehat{zz}$	$\widehat{r_1z}$
$r=2a$				
0°	−.0119	−.0113	−.2825	.0000
30°	−.0538	−.0118	−.2234	−.1015
45°	−.0904	−.0128	−.1583	−.1200
75°	−.0932	−.0133	−.0200	−.0373
80°	−.0660	−.0129	−.0037	−.0181
85°	−.0257	−.0126	−.0013	−.0047
90°	+.0125	−.0125	−.0000	−.0000
$r=3a$				
0°	−.0013	−.0013	−.1463	−.0000
30°	−.0273	−.0013	−.1052	−.0534
45°	−.0482	−.0024	−.0643	−.0560
60°	−.0575	−.0031	−.0256	−.0388
75°	−.0361	−.0044	−.0033	−.0128
80°	−.0241	−.0046	−.0018	−.0065
85°	−.0090	−.0051	−.0000	−.0015
90°	+.0055	−.0055	.0000	.0000
$r=4a$				
0°	−.0008	−.0000	−.0863	−.0000
30°	−.0160	−.0002	−.0603	−.0325
45°	−.0317	−.0005	−.0313	−.0324
60°	−.0327	−.0011	−.0192	−.0202
75°	−.0203	−.0020	−.0017	−.0067
80°	−.0136	−.0023	−.0002	−.0031
85°	−.0052	−.0027	−.0001	−.0008
90°	+.0031	−.0031	−.0000	.0000

Types of Symmetrical Stress

It is shown in Elasticity that in the case of a solid of revolution being strained symmetrically in such a way that the displacement is the same in all planes through the axis of revolution all the quantities which occur can be expressed in terms of either one or two functions as desired. This case will be applied to the problem of a circular area uniformly loaded and resting on an infinite horizontal plane. It will be more convenient in this case to use two functions χ and Ω related in such a way that

$$\frac{\partial^2\Omega}{\partial z^2}=(1-\sigma)\nabla^2\chi,\ \nabla^4\chi=0,\ \nabla^2\Omega=0 \tag{39}$$

where

$$\nabla^2=\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{dy^2}+\frac{\partial^2}{\partial z^2}=\frac{\partial^2}{\partial r^2}+\frac{1}{r}\frac{\partial}{\partial r}+\frac{\partial^2}{\partial z^2}.$$

It is shown that

$$\Delta = \frac{1-2\sigma}{E}(1+\sigma)\nabla^2\chi,$$

(40)
$$U = -\frac{1+\sigma}{E}\left[\frac{\partial\Omega}{\partial r} + \frac{\partial\chi}{\partial r}\right],$$

$$w = \frac{1+\sigma}{E}\left[\frac{\partial\Omega}{\partial z} - \frac{\partial\chi}{\partial z}\right].$$

The following types of stress can now be written out:

Case 1. Let χ be harmonic and $\nabla^2\chi = 0$; we may without loss of generality take $\Omega = 0$ and write stresses of the following types:

(41)
$$\widehat{rr} = \frac{\partial^2\chi}{\partial r^2},$$

$$\widehat{\theta\theta} = \frac{1}{r}\frac{\partial\chi}{\partial r},$$

$$\widehat{zz} = \frac{\partial^2\chi}{\partial z^2},$$

$$\widehat{rz} = \frac{\partial^2\chi}{\partial r\partial z},$$

with the corresponding movements

(42)
$$U = \frac{1+\sigma}{E}\frac{\partial\chi}{\partial r}, \quad w = \frac{1+\sigma}{E}\frac{\partial\chi}{\partial z}.$$

Case 2. In this replace χ by $r\dfrac{\partial\chi}{\partial r}$; then it is easily shown that

(43)
$$\nabla^2\left(r\frac{\partial\chi}{\partial r}\right) = -2\frac{\partial^2\chi}{\partial z^2}, \quad \nabla^4\left(r\frac{\partial\chi}{\partial r}\right) = 0,$$

and the appropriate form for Ω is

(44)
$$\Omega = -2(1-\sigma)\chi,$$

giving stresses

$$(45)\quad \begin{aligned} \widehat{rr} &= (1-2\sigma)\frac{\partial^2\chi}{\partial z^2} + r\frac{\partial}{\partial r}\frac{\partial^2\chi}{\partial z^2} + 2(1-\sigma)\frac{\partial^2\chi}{\partial r^2},\\ \widehat{\theta\theta} &= (1-2\sigma)\frac{\partial^2\chi}{\partial z^2} + 2(1-\sigma)\frac{1}{r}\frac{\partial\chi}{\partial r},\\ \widehat{zz} &= -2(2-\sigma)\frac{\partial^2\chi}{\partial z^2} - r\frac{\partial}{\partial r}\frac{\partial^2\chi}{\partial z^2} + 2(1-\sigma)\frac{\partial^2\chi}{\partial z^2},\\ \widehat{rz} &= r\frac{\partial}{\partial z}\frac{\partial^2\chi}{\partial z^2}, \end{aligned}$$

and movements

$$(46)\quad U = \frac{1+\sigma}{E}\left[2(1-\sigma)\frac{\partial\chi}{\partial r} + r\frac{\partial^2\chi}{\partial z^2}\right],\quad w = -\frac{1+\sigma}{E}\left[2(1-\sigma)\frac{\partial\chi}{\partial z} + r\frac{\partial^2\chi}{\partial r\partial z}\right].$$

Case 3. In this case replace χ by $z\dfrac{\partial\chi}{\partial z}$ then it may be shown that

$$(47)\quad \nabla^2\left(z\frac{\partial\chi}{\partial z}\right) = 2\frac{\partial^2\chi}{\partial z^2},\quad \nabla^4\left(z\frac{\partial\chi}{\partial z}\right) = 0,$$

and the appropriate form for Ω is

$$(48)\quad \Omega = 2(1-\sigma)\chi,$$

giving stresses

$$(49)\quad \begin{aligned} \widehat{rr} &= 2\sigma\frac{\partial^2\chi}{\partial z^2} + z\frac{\partial}{\partial z}\frac{\partial^2\chi}{\partial z^2} + z\frac{\partial}{\partial z}\left(\frac{1}{r}\frac{\partial\chi}{\partial r}\right) - 2(1-\sigma)\frac{\partial^2\chi}{\partial r^2},\\ \widehat{\theta\theta} &= 2\sigma\frac{\partial^2\chi}{\partial z^2} + z\frac{\partial}{\partial z}\frac{\partial^2\chi}{\partial z^2} + z\frac{\partial}{\partial z}\left(\frac{\partial^2\chi}{\partial r^2}\right) - 2(1-\sigma)\frac{1}{r}\frac{\partial\chi}{\partial r},\\ \widehat{zz} &= (2-2\sigma)\frac{\partial^2\chi}{\partial z^2} - z\frac{\partial}{\partial z}\frac{\partial^2\chi}{\partial z^2} - 2(1-\sigma)\frac{\partial^2\chi}{\partial z^2},\\ \widehat{rz} &= -\frac{\partial^2\chi}{\partial r\,\partial z} - z\frac{\partial}{\partial r}\frac{\partial^2\chi}{\partial z^2}, \end{aligned}$$

with corresponding movements

$$(50)\quad U = -\frac{1+\sigma}{E}\left[z\frac{\partial^2\chi}{\partial r.\partial z} + 2(1-\sigma)\frac{\partial\chi}{\partial r}\right],\quad w = \frac{1+\sigma}{E}\left[2(1-\sigma)\frac{\partial\chi}{\partial z} - \frac{\partial\chi}{\partial z} - z\frac{\partial^2\chi}{\partial z^2}\right].$$

In Case 1 the stresses are all independent of the elastic constants of the material and it will be seen that in Case 2 and Case 3, $\widehat{zz}$ is independent of the elastic constants and this is important.

It is now necessary to select an appropriate value for χ for the particular purpose in hand.

If χ is taken as the logarithmic potential of a uniform distribution of matter over the circular area $(r=a)$ then $\partial^2\chi/\partial z^2 = -\omega$ where ω is the solid angle subtended at the point.

Now ω is constant for every point within the area of the loaded circle and therefore $\frac{\partial \omega}{\partial r} = 0$, but $\frac{\partial^2 \chi}{\partial r \partial z}$ is not zero over the loaded area. This can be rectified, however, as $z\frac{\partial \omega}{\partial r} = 0$ and thus the addition of Case 1 and Case 3 will furnish the requisite condition of normal pressure only over the loaded area.

This leads to the same solution as would be arrived at by taking the movements as given in Love's *Treatise*, p. 190, for the general case.

It is clear from the form of the three cases given above that solutions for the stresses involving a combination of pressure and shear over the circular area can be expressed in terms of ω only but $\widehat{zz}$ will not in general be independent of the elastic constants.

AZIONI SISMICHE SUSSULTORIE SU MONTANTI VERTICALI INCASTRATI ALLA BASE E CON CARICHI E VINCOLI ELASTICI ALL'ESTREMO SUPERIORE

Del Professore Umberto Puppini,
R. Università di Bologna, Bologna, Italia.

1. Studi iniziati qualche anno fa mi hanno condotto a considerare le tensioni interne provocate nel montante di un edificio dal moto sismico sussultorio del suolo, dapprima nella supposizione che il montante sia privo di carichi e di vincoli elastici in sommità, poi molto recentemente nella supposizione che il montante privo di vincolo elastico in sommità sia ivi gravato da un carico concentrato, e infine ora nella supposizione che il montante presenti in sommità vincolo elastico con altre strutture.

In questa communicazione riassumo i risultati di tali studi in parte già enunciati, in parte ancore inediti.

2. Pensiamo che il suolo compia un moto in direzione verticale.

Ammettiamo, come il buon senso costruttivo consiglia, che l'edificio sia ben solidale col suolo; cioè non prendiamo in considerazione appoggi formati di rulli, sfere, molle, e simili congegni che, rimasti immobili per mesi e per anni, presentanto certamente, al momento in cui occorrerebbe il loro funzionamento, una resistenza enorme, e non già quella minima resistenza che è condizione necessaria (non però sufficiente) perchè il soprastante edificio non risenta gravi sollecitazioni pel fatto del moto sismico.

Tale vincolo della base dell'edificio porta alla condizione che il moto della base del montante sia identico al moto del suolo.

3. Consideriamo un montante di un edificio, senza ammettere od escludere per ora carichi concentrati o vincoli elastici all'estremo superiore e senza distinguere se il montante sia o meno l'unica membratura dell'edificio.

Indichiamo con

l l'altezza del montante,

z l'altezza di una sezione orizzontale generica rispetto alla sezione di incastro alla base,

p la pressione per unità di area nella sezione generica di ordinata z all'istante generico t, pressione dovuta solo al fatto del moto del montante sotto l'azione sismica (p è la differenza fra la pressione vera e la pressione statica),

v la velocità con cui si sposta verticalmente verso il basso una sezione generica in un istante generico.

A l'area della sezione orizzontale del montante, area supposta costante, cioè indipendente da z,

E ρ, ω modulo di elasticità, densità, peso specifico del materiale di cui il montante è constituito.

Le equazioni differenziali fra le variabili dipendenti p, v e le variabili indipendenti z, t sono l'una l'equazione del moto, l'altra l'equazione di elasticità.

$$(1) \qquad \begin{cases} \dfrac{\partial p}{\partial z} = \rho \dfrac{\partial v}{\partial t}, \\ \dfrac{\partial p}{\partial t} = E \dfrac{\partial v}{\partial z}. \end{cases}$$

Il sistema (1) è sufficiente allo studio del moto delle singole sezioni e delle pressioni che in esse si generano, quando siano date le condizioni iniziali, le condizioni relative ai vincoli meccanici del sistema e il moto di una sezione del montante.

Porremo l'origine del tempo all'istante in cui il moto sismico si inizia. Allora, per $t=0$, sono nulli lo spostamento e la velocità in tutte le sezioni.

Quanto ai vincoli meccanici, trattandosi, come si è detto, di montante verticale solidale col suolo, si avrà per l'estremo inferiore ($z=0$) che il moto di esso estremo sia identico al moto del suolo. Per l'estremo superiore si avranno condizioni analitiche da specificarsi a seconda delle condizioni fisiche di vincolo elastico o di carico di esso estremo.

4. Il sistema integrale del sistema differenziale (1) è dato da

$$(2) \qquad \begin{cases} p = F\left(t - \dfrac{z}{a}\right) - f\left(t + \dfrac{z}{a}\right), \\ v = -\dfrac{1}{\sqrt{E\rho}} \left\{ F\left(t - \dfrac{z}{a}\right) + f\left(t + \dfrac{z}{a}\right) \right\}, \end{cases}$$

nel quale F, f sono simboli di due funzioni indeterminate, ed è $a = \sqrt{\dfrac{E}{\rho}}$.

Essendo, F, f funzioni rispettivamente degli argomenti $t - \dfrac{z}{a}$, $t + \dfrac{z}{a}$, evidentemente il sistema (2) definisce tanto la pressione quanto la velocità come risultanti dalla somma di due parti che si propagano lungo il montante, l'una con velocità a dalla base verso l'estremo superiore, l'altra colla stessa velocità dall'estremo superiore verso la base.

5. A partire dall'inizio della perturbazione sismica, cioè da $t=0$, la perturbazione stessa (funzione F) si propaga dalla base all'estremo superiore colla velocità a. Dunque occorre il tempo $\dfrac{l}{a}$ perchè la perturbazione F raggiunga l'estremo superiore, dopo di che la perturbazione f comincierà a propagarsi verso la base partendo dall'estremo superiore.

Occorre il tempo $\frac{2l-z}{a}$ perchè l'azione f abbia inizio in una sezione alta z sulla base. Dunque in una sezione generica, per il tempo $\frac{2l-z}{a}$ dall'inizio del moto sismico lo stato del montante, sia nei riguardi della pressione, sia nei riguardi della velocità, è rappresentato dalla sola funzione F. Con tale limitazione nel tempo, le (2) si riducono a

$$(3) \qquad \begin{cases} p = F\left(t - \frac{z}{a}\right) \\ v = -\frac{1}{\sqrt{E\rho}} F\left(t - \frac{z}{a}\right); \end{cases}$$

dalle quali si ricava

$$(4) \qquad p = -\sqrt{E\rho}.v.$$

Chiameremo periodo di colpo sismico diretto il periodo di tempo $\frac{2l-z}{a}$ in cui, per la sezione generica z, sono valide la (3) e la (4), mentre diremo periodo di contraccolpi sismici tutto il tempo successivo.

Si noti bene che le (3), (4) sono valide nel periodo di colpo sismico diretto, qualunque siano le condizioni di carico o di vincolo all'estremo superiore del montante. Per giungere, infatti, alle formule (3), (4) non si è mai supposta alcuna particolare condizione meccanica, dell'estremo superiore del montante.

Il montante può essere isolato o invece essere elemento di un complesso edificio, esso può essere o no gravato di carichi in sommità; e sempre le condizioni pressione e di velocità nelle varie sezioni, entro il periodo di colpo sismico diretto, sono rappresentate dalle equazioni (3), (4).

6. La formula (4) dà in qualche modo la valutazione della differenza di effetti fra un terremoto così detto con urto iniziale e un terremoto senza urto iniziale. Nel terremoto con urto iniziale lo spostamento del suolo, e quindi anche della base del montante che è solidale col suolo, avviene con velocità la quale, pur soddisfacendo alla condizione $v=0$ per $t=0$, assume valori assai grandi per piccolissimi valori di t. Nel terremoto senza urto inziale lo spostamento del suolo avviene con velocità la quale dal valore $v=0$ per $t=0$ cresce più gradatamente che non nel caso precedente.

Per il brevissimo intervallo di tempo uguale alla durata del colpo sismico diretto la (4) è valida ugualmente per terremoto con o senza urto iniziale. Ma nel terremoto con urto iniziale nella (4) si possono introdurre valori più elevati di velocità, nel terremoto senza urto iniziale la (4) è valida solo per più limitati valori di velocità.

L'attributo di velocità iniziale diverso da zero, anzi relativamente considerevole, che a rigore non è fisicamente concepibile, prende pertanto questo significato concreto di velocità considerevole acquistata nel periodo di colpo sismico diretto, periodo di durata sia pure brevissima, ma certamente diversa da zero.

7. Durante il periodo dei contraccolpi sismici, cioè, in una sezione alta z sulla base, per $t \gtreqless \frac{2l-z}{a}$, le equazioni (2) debbono ricevere le specificazioni inerenti alle condizioni particolari dell'estremo superiore.

All'estremo superiore, ove manchi affatto carico concentrato e vincolo elastico, si avrà, per $t \gtreqless \frac{l}{a}$:

$$p=0; \tag{5}$$

ove esista un carico concentrato di massa M:

$$p=-\frac{M}{A}\frac{\partial v}{\partial t}; \tag{6}$$

ove si abbia un vincolo elastico:

$$\frac{\partial p}{\partial t}=-\frac{K}{A}v, \tag{7}$$

con K coefficiente costante di elasticità.

Mentre sono bene evidenti le formule (5) e (6), abbisogna di un breve chiarimento la formula (7).

L'estremità superiore del montante (che è pure estremità delle membrature orizzontali colle quali nel caso della (7) il montante si collega) assume dall'istante $\frac{l}{a}$ in poi degli spostamenti che non sono immediatamente assunti da tutti i correnti orizzontali, perchè le azioni sismiche trasversali all'asse di una membratura si propagano assai più lente delle azioni longitudinali.

Il moto della sommità del montante è pertanto accompagnato da una deformazione degli assi orizzontali dei correnti, alla quale necessariamente si accompagna una reazione verticale dei correnti contro il montante che col suo moto produce la detta deformazione.

Indicando con y lo spostamento verticale della sommità del montante dalla posizione di equilibrio precedente l'azione sismica (y positivo se si tratti di abbassamento), avremo in via di approsimazione:

$$p=-\frac{K}{A}y,$$

dove la costante K dipenderà dalle luci orizzontali dei correnti solidali coi montanti, dai momenti di inerzia delle sezioni dei correnti e dal modulo di elasticità del materiale di cui i correnti sono constituiti.

Derivando la formula ora scritta rispetto a t si ottiene appunto la (7).

8. Portando le condizioni espresse dalle (5), (6), (7) nella prima della (2), si ottiene rispettivamente:
per carico e vincolo elastico nulli in sommità:

$$0 = F\left(t - \frac{l}{a}\right) - f\left(t + \frac{l}{a}\right); \tag{8}$$

per carico in sommità di massa M

$$\frac{M}{A\sqrt{E\rho}}\left\{\frac{\partial F\left(t-\frac{l}{a}\right)}{\partial t} + \frac{\partial f\left(t+\frac{l}{a}\right)}{\partial t}\right\} = F\left(t-\frac{l}{a}\right) - f\left(t+\frac{l}{a}\right); \tag{9}$$

per vincolo in sommità definito dal coefficiente K:

$$\frac{\partial F\left(t-\frac{l}{a}\right)}{\partial t} - \frac{\partial f\left(t+\frac{l}{a}\right)}{\partial t} = \frac{K}{A\sqrt{E\rho}}\left\{F\left(t-\frac{l}{a}\right) + f\left(t+\frac{l}{a}\right)\right\}. \tag{10}$$

Quanto sopra è sempre subordinato a che sia $t \gtreqless \frac{l}{a}$.

L'argomento $t + \frac{l}{a}$ della funzione f può anche essere sostituito, nelle equazioni (8), (9), (10), coll'argomento $t + \frac{z}{a}$ relativo a una sezione generica, purchè sia $t + \frac{z}{a} \gtreqless \frac{2l}{a}$.

Allora le (8), (9), (10) prendono la forma

$$0 = F\left(t + \frac{z}{a} - \frac{2l}{a}\right) - f\left(t + \frac{z}{a}\right), \tag{8$^{\text{I}}$}$$

$$\frac{M}{A\sqrt{E\rho}}\left\{\frac{\partial F\left(t+\frac{z}{a}-\frac{2l}{a}\right)}{\partial t} + \frac{\partial f\left(t+\frac{z}{a}\right)}{\partial t}\right\} \doteq F\left(t+\frac{z}{a}-\frac{2l}{a}\right) - f\left(t+\frac{z}{a}\right), \tag{9$^{\text{I}}$}$$

$$\frac{\partial F\left(t+\frac{z}{a}-\frac{2l}{a}\right)}{\partial t} - \frac{\partial f\left(t+\frac{z}{a}\right)}{\partial t} = \frac{K}{A\sqrt{E\rho}}\left\{F\left(t+\frac{z}{a}-\frac{2l}{a}\right) + f\left(t+\frac{z}{a}\right)\right\}. \tag{10$^{\text{I}}$}$$

e per $z = 0$ (sezione base):

$$0 = F\left(t - \frac{2l}{a}\right) - f(t), \tag{8$^{\text{II}}$}$$

$$\frac{M}{A\sqrt{E\rho}}\left\{\frac{\partial F\left(t-\frac{2l}{a}\right)}{\partial t} + \frac{\partial f}{\partial t}\right\} = F\left(t-\frac{2l}{a}\right) - f(t), \tag{9$^{\text{II}}$}$$

$$\frac{\partial F\left(t-\frac{2l}{a}\right)}{\partial t} - \frac{\partial f(t)}{\partial t} = \frac{K}{A\sqrt{E\rho}}\left\{F\left(t-\frac{2l}{a}\right) + f(t)\right\}. \tag{10$^{\text{II}}$}$$

9. Indicando con t_1 un istante del periodo di colpo sismico diretto (periodo che, per $z=0$, dura il tempo $\frac{2l}{a}$ dall'inizio della perturbazione sismica, tempo che chiameremo "durata di fase"), con $t_2=t_1+\frac{2l}{a}$, $t_3=t_1+\frac{4l}{a}$, $t_4=t_1+\frac{6l}{a}$, ... istanti contenuti in quelle che chiameremo successive fasi di contraccolpo ognuna della durata $\frac{2l}{a}$, istanti distanziati fra loro precisamente della durata di fase, e chiamando

$$F_1,\ F_2,\ F_3,\ F_4,\ \ldots,$$
$$f_1, f_2, f_3, f_4,\ \ldots,$$
$$P_1,\ P_2,\ P_3,\ P_4,\ \ldots,$$
$$V_1,\ V_2,\ V_3,\ V_4,\ \ldots,$$

i valori delle funzioni F, f, p, v, per $z=0$, per $t=t_1$, $t=t_2$, $t=t_3$, $t=t_4, \ldots$, si ottiene dalle equazioni (2):

$$(11)\qquad \begin{cases} P_1=F_1,\\ P_2=F_2-f_2,\\ P_3=F_3-f_3,\\ P_4=F_4-f_4,\\ \ldots\ldots\ldots \end{cases}$$

$$(12)\qquad \begin{cases} V_1=-\dfrac{1}{\sqrt{E\rho}}F_1,\\ V_2=-\dfrac{1}{\sqrt{E\rho}}\{F_2+f_2\},\\ V_3=-\dfrac{1}{\sqrt{E\rho}}\{F_3+f_3\},\\ V_4=-\dfrac{1}{\sqrt{E\rho}}\{F_4+f_4\},\\ \ldots\ldots\ldots\ldots\ldots\ldots \end{cases}$$

D'altra parte le equazioni (8^{II}), (9^{II}), (10^{II}) per $t=t_2$ prendono la forma

$$(8^{III})\qquad F_1=f_2,$$

$$(9^{III})\qquad \frac{M}{A\sqrt{E\rho}}\left\{\frac{\partial F_1}{\partial t}+\frac{\partial f_2}{\partial t}\right\}=F_1-f_2,$$

$$(10^{III})\qquad \frac{\partial F_1}{\partial t}-\frac{\partial f_2}{\partial t}=\frac{K}{A\sqrt{E\rho}}\{F_1+f_2\};$$

e per $t=t_3$:

$$(8^{IV})\qquad F_2=f_3,$$

$$(9^{IV})\qquad \frac{M}{A\sqrt{E\rho}}\left\{\frac{\partial F_2}{\partial t}+\frac{\partial f_3}{\partial t}\right\}=F_2-f_3,$$

$$(10^{IV}) \qquad \frac{\partial F_2}{\partial t} - \frac{\partial f_3}{\partial t} = \frac{K}{A\sqrt{E\rho}} \{F_2 + f_3\} ;$$

e per $t = t_4$:

$$(8^{V}) \qquad F_3 = f_4,$$

$$(9^{V}) \qquad \frac{M}{A\sqrt{E\rho}} \left\{ \frac{\partial F_3}{\partial t} + \frac{\partial f_4}{\partial t} \right\} = F_3 - f_4,$$

$$(10^{V}) \qquad \frac{\partial F_3}{\partial t} - \frac{\partial f_4}{\partial t} = \frac{K}{A\sqrt{E\rho}} \{F_3 + f_4\} ;$$

e così via.

Sostituendo le (8^{III}), (8^{IV}), (8^{V}), . . . nelle (11), (12), si ottiene:

$$(11^{I}) \quad \begin{cases} P_1 = F_1, \\ P_2 = F_2 - F_1, \\ P_3 = F_3 - F_2, \\ P_4 = F_4 - F_3, \\ \dots\dots\dots\dots \end{cases} \qquad (12^{I}) \quad \begin{cases} V_1 = -\dfrac{1}{\sqrt{E\rho}} F_1, \\ V_2 = -\dfrac{1}{\sqrt{E\rho}} \{F_2 + F_1\}, \\ V_3 = -\dfrac{1}{\sqrt{E\rho}} \{F_3 + F_2\}, \\ V_4 = -\dfrac{1}{\sqrt{E\rho}} \{F_4 + F_3\}, \\ \dots\dots\dots\dots\dots\dots \end{cases}$$

equazioni valide per il caso di carico e di vincolo nulli in sommità.

Risolvendo le (9^{III}), (9^{IV}), (9^{V}), . . ., rispetto a f_2, f_3, f_4, . . ., si otterranno funzioni $f_2 = \phi_2(F_1)$, $f_3 = \phi_3(F_2)$, $f_4 = \phi_4(F_3)$ . . ., le quali sostituite nelle (11), (12) danno luogo a

$$(11^{II}) \quad \begin{cases} P_1 = F_1, \\ P_2 = F_2 - \phi_2(F_1), \\ P_3 = F_3 - \phi_3(F_2), \\ P_4 = F_4 - \phi_4(F_3), \\ \dots\dots\dots\dots, \end{cases} \qquad (12^{II}) \quad \begin{cases} V_1 = -\dfrac{1}{\sqrt{E\rho}} F_1, \\ V_2 = -\dfrac{1}{\sqrt{E\rho}} \{F_2 + \phi_2(F_1)\}, \\ V_3 = -\dfrac{1}{\sqrt{E\rho}} \{F_3 + \phi_3(F_2)\}, \\ V_4 = -\dfrac{1}{\sqrt{E\rho}} \{F_4 + \phi_4(F_3)\}, \\ \dots\dots\dots\dots\dots\dots \end{cases}$$

valide nel caso di carico concentrato in sommità.

Risolvendo le (10^{III}), (10^{IV}), (10^{V}), . . ., rispetto a f_2, f_3, f_4, . . ., si otterranno funzioni $f_2 = \psi_2(F_1)$, $f_3 = \psi_3(F_2)$, $f_4 = \psi_4(F_3)$, . . ., le quali sostituite nelle (11), (12) danno luogo a

$$(11^{III})\quad \begin{cases} P_1 = F_1, \\ P_2 = F_2 - \psi_2(F_1), \\ P_3 = F_3 - \psi_3(F_2), \\ P_4 = F_4 - \psi_4(F_3), \\ \dots\dots\dots\dots \end{cases} \qquad (12^{III})\quad \begin{cases} V_1 = -\dfrac{1}{\sqrt{E\rho}}\,F_1, \\ V_2 = -\dfrac{1}{\sqrt{E\rho}}\,\{F_2 + \psi_2(F_1)\}, \\ V_3 = -\dfrac{1}{\sqrt{E\rho}}\,\{F_3 + \psi_3(F_2)\}, \\ V_4 = -\dfrac{1}{\sqrt{E\rho}}\,\{F_4 + \psi_4(F_3)\}' \\ \dots\dots\dots\dots\dots\dots\dots\dots \end{cases}$$

valide pel caso di vincolo elastico in sommità.

Dalle (11^{I}), (12^{I}), eliminate le F, si ottengono le P (pressioni alla base del montante) espresse per le V, cioè espresse a mezzo della legge del moto sussultorio del suolo, nel caso di montante senza carichi e senza vincoli elastici in sommità.

Ricavate le funzioni ϕ dalle (9^{III}), (9^{IV}), (9^{V}) . . ., dalle (11^{II}), (12^{II}), eliminate le F, si ottengono le P espresse per le V nel caso di carico concentrato in sommità.

Ricavate le funzioni ψ dalle (10^{III}), (10^{IV}), (10^{V}) . . ., dalle (11), (12), eliminate le F, si ottengono le P espresse per le V nel caso di vincolo elastico in sommità.

10. Le formule (11), (12), (11^{I}), (12^{I}), (11^{II}), (12^{III}) risolvono il problema tecnico che stiamo studiando, quando bene inteso, si sappiano integrare le equazioni differenziali (9^{III}), (10^{III}) e le analoghe, e sia assegnato il moto del suolo.

Prima di passare alla considerazione specifica di una ipotesi particolare del moto del suolo, ritengo opportuno di osservare come l'impostazione del problema delle pressioni interne in un montante per un moto sismico sussultorio e la risoluzione di esso almeno fino alle equazioni (11), (12) si svolgano in modo formalmente analogo a quanto accade di problemi in altri campi della Fisica, come per esempio per la propagazione di perturbazioni elettro-magnetiche in lunghe linee e per la propagazione di perturbazioni idrodinamiche in tubi pieni d'acqua. Il quale ultimo problema, noto col nome di problema del colpo di ariete, ha avuto particolare tributo di studio da Lorenzo Allievi.

11. Applicheremo le formule (11^{I}), (12^{I}), (11^{II}), (12^{II}), (11^{III}), (12^{III}) a un caso interessante, quello in cui la velocità del moto sussultorio del suolo aumenti, almeno per le prime fasi, proporzionalmente al tempo, cioè al caso in cui sia $V = \alpha t$, essendo α l'accelerazione constante del moto del suolo e t il tempo valutato a partire dall'istante iniziale del moto.

Si ha allora, per la fase di colpo sismico diretto, dalle prime delle (11^{I}), (12^{I}), (11^{II}), (12^{II}), (11^{III}), (12^{III}):

$$(13)\qquad P_1 = -\sqrt{E\rho}\,.\,\alpha\, t_1,$$

tanto nel caso di carico e di vincolo elastico nulli in sommità, quanto nel caso di carico concentrato in sommità e nel caso di vincolo elastico in sommità.

12. Veniamo alla prima fase di contraccolpo.

Fra le due seconde equazioni delle (11^{I}), (12^{I}) eliminiamo F_2 dopo di aver fatto $F_1 = -\sqrt{E\rho}\,.\,V_1$, $V_1 = at_1$, $V_2 = at_2$.

Si ottiene, colla posizione $\lambda = \sqrt{E\rho}\,.\,a$:

$$P_2 = +\lambda\left(t_2 - \frac{4l}{a}\right). \tag{14^{I}}$$

Integriamo la (9^{III}), dopo aver posto $F_1 = -\lambda t_1$, colla condizione che per $t = \dfrac{2l}{a}$ deve essere ancora $f_2 = 0$, e utilizzando le seconde equazioni delle (11^{II}), (12^{II}) sempre con $V_1 = at_1$, $V_2 = at_2$. Si ottiene, colla posizione $\dfrac{A\sqrt{E\rho}}{M} = \mu$:

$$P_2 = \lambda\left(t_2 - \frac{4l}{a}\right) + 4\frac{\lambda}{\mu}\left(e^{\mu\left(\frac{2l}{a} - t_2\right)} - 1\right), \tag{14^{II}}$$

con e base dei logaritmi iperbolici. La (14^{II}) vale per carico in sommità di massa M.

Analogamente: integriamo la (10^{III}), dopo aver posto $F_1 = -\lambda t_1$, $V_1 = at_1$, $V_2 = at_2$, colla condizione che, per $t = \dfrac{2l}{a}$, deve essere ancora $f_2 = 0$, e utilizzando le seconde equazioni delle ($11'''$), ($12'''$). Si ottiene, colla posizione $\dfrac{K}{A\sqrt{E\rho}} = \nu$:

$$P_2 = -\lambda\left(3t_2 - \frac{4l}{a}\right) - \frac{4\lambda}{\nu}\left(e^{\nu\left(\frac{2l}{a} - t_2\right)} - 1\right). \tag{14^{III}}$$

La (14^{III}) vale per vincolo elastico in sommità.

13. Con analoghi processi si ottiene nella seconda fase di contraccolpo:

Per carico e vincolo nulli in sommità:

$$P_3 = -\lambda\left(t_3 - \frac{4l}{a}\right); \tag{15^{I}}$$

per carico in sommità:

$$P_3 = -\lambda\left(t_3 - \frac{4l}{a}\right) - \frac{4\lambda}{\mu}\left\{2e^{\mu\left(\frac{4l}{a} - t_3\right)} - e^{\mu\left(\frac{2l}{a} - t_3\right)} - 2e^{\mu\left(\frac{4l}{a} - t_3\right)}\,.\,\mu\left(\frac{4l}{a} - t_3\right) - 1\right\}; \tag{15^{II}}$$

per vincolo elastico in sommità:

$$P_3 = -\lambda\left(5t_3 - \frac{12l}{a}\right) - \frac{4\lambda}{\nu}\left\{2e^{\nu\left(\frac{4l}{a} - t_3\right)} + e^{\nu\left(\frac{2l}{a} - t_3\right)} - 2e^{\nu\left(\frac{4l}{a} - t_3\right)}\,.\,\nu\left(\frac{4l}{a} - t_3\right) - 3\right\}. \tag{15^{III}}$$

Con analogo processo si studia la terza fase di contraccolpo.

14. *E'* facile verificare che le formule (14^{II}), (15^{II}), per $M = 0$, cioè per $\mu = \infty$, si riducono, come bene deve essere, alle (14^{I}), (15^{I}); e così le (14^{III}), (15^{III}), per $k = 0$, cioè per $\nu = 0$, si riducono pure alle (14^{I}), (15^{I}).

A GYRO COMPASS INCORPORATING TWO GYROSCOPES

By Professor Sir James B. Henderson,
Adviser on Gyroscopic Equipment to the Admiralty, London, England.

The gyroscopic compass is a scientific instrument of the most sensitive kind and is an excellent example of the application of theory in practice. It can be so designed that the motion of the compass under all conditions can be calculated to a high degree of accuracy so that before the instrument is constructed its behaviour under any circumstances can be foretold, and when the design has materialized the problem then becomes one of locating and eliminating the causes of any differences between the actual and the theoretical behaviour.

During the last ten years the most important problem in gyro compasses has been the elimination of the deviation of the compass introduced by the oscillations of the ship due to wave motion, and when the history of this period comes to be written an important phase of development will be that in which the compass comprised two gyroscopes. One such compass made by the Sperry Gyroscope Company had one of the gyroscopes meridian-seeking and the other the reverse, but the latter was compelled by discontinuous torques to follow the former. In another two-gyro compass made by the author both gyroscopes were meridian-seeking, all torques being continuous functions and the interaction of the two gyroscopes providing mutual damping. The mathematical analysis of the motion of this latter compass, the conditions for its stability and the investigation of the damping coefficients of both periods form the subject of this paper.

Description of the Compass

Figure 1 represents diagramatically a front elevation of the compass without its supports or following motor, and Figure 2 represents a side elevation of the sensitive element alone.

The compass contains two gyroscopes A and A' supported respectively upon horizontal trunnion axes B and B' in vertical gimbal rings C and C' pivoted on the vertical trunnions V and V' in the frame D. This frame is pivoted on the vertical trunnion U in the follower Z. The follower is supported in a binnacle on a large ball race by the disc E which is toothed on its periphery and is driven in azimuth by a small motor controlled by a roller contact F carried by the frame D of the sensitive element and engaging with a two-part commutator G fixed to the follower Z. The following motor therefore keeps the follower Z in phase in azimuth with the roller contact F. The frame D is suspended from the bridge R carried by the plate E by the filar suspension S, the ring C

is suspended from the frame D by the torsion element H and the ring C' is supported from the frame D by the torsion element H' which is in thrust.

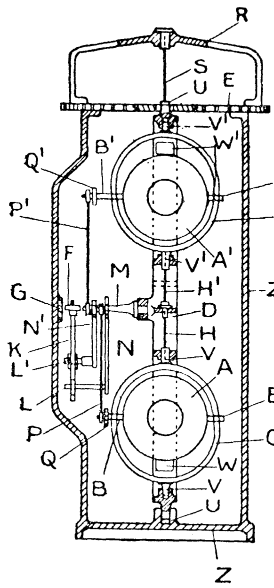


Fig. 1

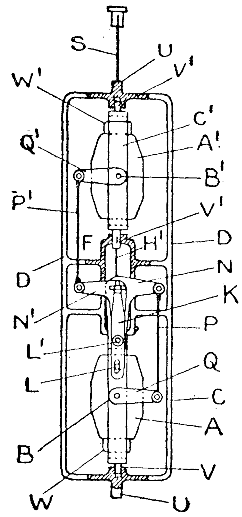


Fig. 2

The gyroscope A is put in gravitational stability by the weight W and the gyroscope A' is put in gravitational instability by the equal weight W'. The main idea in the design is to get an arrangement in which the gravitational control is practically neutral for periodic accelerations of short period due to waves, which is the case if the gyroscope cases A and A' carrying the weights W and W' are connected together by a viscous element. The details of this arrangement are omitted in the diagram in order to prevent complication. The gyroscope A', if acting alone, oscillates about the meridian with its axis of angular momentum pointing south and in order to damp these oscillations energy has to be supplied to the gyroscope by couples about the vertical axis *assisting* the precession. The gyroscope A, if acting alone, oscillates about the meridian with its axis of angular momentum pointing north and to damp these oscillations energy has to be taken from the gyroscope by couples acting about the vertical axis *opposing* the precession. Hence a mutual torque between the two gyroscopes due to the torsion elements H and H' will assist in damping the oscillations of both gyroscopes if the stable gyroscope A precesses faster than the unstable one A', *i.e.*, if A has smaller angular momentum than A'.

With such a combination there would evidently be a stable state with A and A' in the same azimuth and therefore producing no twist in the filar $H-H'$ and with A' tilted more than A so that both precessed in azimuth at the same rate, in which condition they might precess together continuously round and round in azimuth. In order to prevent such a stable state from arising, an external torque with a vertical axis is applied to both gyroscopes proportional to the difference between the tilts of the two rotor axes and in such a direction as to tend to keep the axes of A and A' in phase as regards tilt. This torque is produced by displacing the roller contact F in azimuth by an amount proportional to the difference in tilt between the two gyros. Horizontal cranks Q and Q' are fixed to the trunnions B and B' of the gyros A and A' (Fig. 2) and these are connected by connecting rods P and P' with two bell cranks N and N' pivoted on a spindle M carried by the frame D, so that the bell cranks N and N' tilt in phase with the two gyros. The roller contact F is mounted on a lever K which is pivoted on a crank pin L' fixed to the bell crank N', and a crank pin L on the bell crank N engages with a slot in the end of the lever K. It will be evident that this mechanism displaces the roller F in azimuth by an amount proportional to the difference of tilt of the two gyros and by actuating the following motor twists the suspension S by an amount proportional to the difference in tilt of the two gyroscopes in the direction required to reduce that difference.

Equations of Motion

Let θ_1 be the azimuthal deviation of gyro A reckoned +ive from N. to W.

θ_2 be the azimuthal deviation of gyro A' reckoned +ive from N. to W.

ϕ_1 the tilt of gyro A reckoned +ive with N. end high.

ϕ_2 the tilt of gyro A′ reckoned + ive with N. end high.

$G\phi_1$ the gravitational torque on gyro A due to tilt ϕ_1.

$G\phi_2$ the gravitational torque on gyro A' due to tilt ϕ_2.

$h(\theta_1-\theta_2)$ mutual torque due to torsion element $H-H'$.

$n(\phi_1-\phi_2)$ torque on both gyros applied externally through suspension.

$I_1\Omega_1$ angular momentum of gyro A.

$I_2\Omega_2$ angular momentum of gyro A'.

ω_1 angular velocity of precession of gyro A due to unit couple.

ω_2 angular velocity of precession of gyro A' due to unit couple.

ω angular velocity of the earth.

λ latitude.

ω_1 and ω_2 are introduced to avoid reciprocals in the equations:

$$\omega_1 = \frac{\text{unit couple}}{I_1\Omega_1}$$

and

$$\omega_2 = \frac{\text{unit couple}}{I_2\Omega_2}.$$

It is most convenient to refer the motion of the gyroscopes to axes which move with the rings C and C'. Thus in Figures 3 and 4 which refer respectively to the gyroscopes A and A', OB represents the horizontal trunnion axis B, OV represents the vertical trunnion axis V and OX is perpendicular to both and represents the azimuth of the rotor axis OA. OM represents the meridian and OP is a vector parallel to the axis of the earth and represents the angular velocity of the earth. The angle MOX is θ_1, AOX is ϕ_1 and POM is λ the latitude.

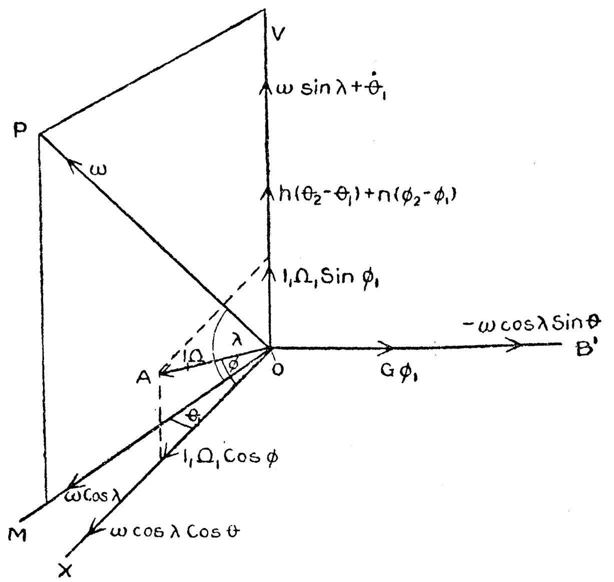


Fig. 3

Writing down on each axis the component angular momentum, the component couple, and the component precessional velocity and equating the couples to the rate of increase of the component angular momentum, assuming θ and ϕ to be small angles, we get from Figure 3:

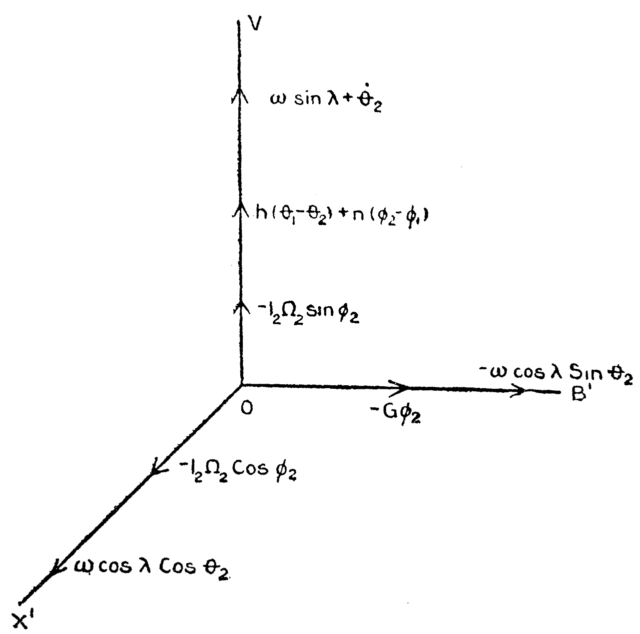


Fig. 4

$$G\phi_1 = I_1\Omega_1(\omega \sin \lambda + \theta_1') - I_1\Omega_1\phi_1\omega \cos \lambda,$$
$$h(\theta_2-\theta_1) + n(\phi_2-\phi_1) = I_1\Omega_1\phi_1' + I_1\Omega_1\omega \cos \lambda . \theta_1,$$

and from Figure 4:

$$-G\phi_2 = -I_2\Omega_2(\omega \sin \lambda + \theta_2') + I_2\Omega_2\phi_2\omega \cos \lambda,$$
$$h(\theta_1-\theta_2) + n(\phi_2-\phi_1) = -I_2\Omega_2\phi_2' - I_2\Omega_2\omega \cos \lambda . \theta_2.$$

Whence

$$(1)\quad \begin{cases} \theta_1' = \phi_1[G\omega_1 + \omega \cos \lambda] - \omega \sin \lambda, \\ \theta_2' = \phi_2'[G\omega_2 + \omega \cos \lambda] - \omega \sin \lambda, \\ \phi_1' = -\omega \cos \lambda . \theta_1 - h\omega_1(\theta_1-\theta_2) - n\omega_1(\phi_1-\phi_2), \\ \phi_2' = -\omega \cos \lambda . \theta_2 - h\omega_2(\theta_1-\theta_2) + n\omega_2(\phi_1-\phi_2). \end{cases}$$

Let $\theta_1+\theta_2=\alpha : \theta_1-\theta_2=\beta : \phi_1+\phi_2=\gamma$ and $\phi_1-\phi_2=\delta$.
Let $\omega_1+\omega_2=s$ and $\omega_1-\omega_2=d$, then by substitution we get

$$(2)\quad \begin{cases} \alpha' = G\left(\frac{\gamma}{2}s + \frac{\delta}{2}d\right) - 2\omega \sin \lambda, \\ \beta' = G\left(\frac{\gamma}{2}d + \frac{\delta}{2}s\right), \\ \gamma' = -\omega \cos \lambda . \alpha - hs\beta - nd\delta, \\ \delta' = -\omega \cos \lambda . \beta - hd\beta - ns\delta. \end{cases}$$

The term $\omega \cos \lambda$ has been omitted in comparison with $G\omega_1$. In the ordinary single-gyro compass $G\omega_1$ is $\frac{2\pi}{3 \text{ minutes}}$ whereas $\omega \cos \lambda$ is $\frac{2\pi}{48 \text{ hours}}$ at $\lambda = 60°$ and the latter is negligible in comparison with the former. α is of course double the mean deviation of the two gyros, β is their relative motion in azimuth, γ is double the mean tilt and δ is their relative tilt.

To find the settling position let $\alpha' = \beta' = \gamma' = \delta' = 0$ and denote zero value by the suffix $_0$, or more directly in equations (1) let $\theta_1' = \theta_2' = \phi_1' = \phi_2' = 0$, then

$${}_0\phi_2 = \frac{\omega \sin \lambda}{G\omega_2}, \quad {}_0\phi_1 = \frac{\omega \sin \lambda}{G\omega_1},$$

$${}_0\theta_1(h\omega_1 + \omega \cos \lambda) = {}_0\theta_2 . h\omega_1 - n\omega_1\delta_0 = h\omega_1 . {}_0\theta_2 + \frac{\omega \sin \lambda . nd}{G\omega_2},$$

and

$${}_0\theta_2 = \frac{-2n \tan \lambda}{G}.$$

All these values are due to $\omega \sin \lambda$ and by putting a weight on the north side of the stable gyro and one on the south side of the unstable gyro, all can be reduced to zero as is done in the single-gyro compass.

Periodic Motion

Let $\alpha = \alpha_0 + Ae^{pt}$, $\beta = \beta_0 + \beta e^{pt}$, $\gamma = \gamma_0 + Ce^{pt}$ and $\delta = \delta_0 + De^{pt}$.
A, B, C and D are determined by the initial conditions and we have to

eliminate these to determine the values of p. Substituting in equations (2) and eliminating A, B, C and D, we get the determinant

$$\begin{vmatrix} p & 0 & -\frac{Gs}{2} & -\frac{Gd}{2} \\ 0 & p & -\frac{Gd}{2} & -\frac{Gs}{2} \\ \omega\cos\lambda & hs & p & nd \\ 0 & (hd+\omega\cos\lambda) & 0 & (p+ns) \end{vmatrix} = 0. \tag{3}$$

Expanding this we get

$$p^4+p^3ns+p^2.Gs(hd+\omega\cos\lambda)+pnG\frac{s^2-d^2}{2}(hd+\omega\cos\lambda)$$

$$+\omega\cos\lambda(hd+\omega\cos\lambda).G^2\frac{s^2-d^2}{4}=0$$

or

$$p^4+p^3ns+p^2Gs(hd+\omega\cos\lambda)+2pnG\omega_1\omega_2(hd+\omega\cos\lambda)+\omega\cos\lambda.hdG^2\omega_1\omega_2=0,$$

or, if $\omega\cos\lambda$ is small compared with hd, we get

$$p^4+p^3ns+p^2Gshd+2pnGh\omega_1\omega_2d+\omega\cos\lambda.hdG^2\omega_1\omega_2=0. \tag{4}$$

In order that the instrument may be a compass this equation must represent two damped oscillations, one being the oscillation of the two gyros together and the other their relative oscillation. The former is the period of the compass and the latter a much shorter period.

The first condition in order that this equation shall represent two damped oscillations is that all coefficients of p shall be positive*. That is satisfied if d is positive, *i.e.*, if $\omega_1>\omega_2$. That is, the stable gyro has smaller angular momentum (or greater gravity control) than the unstable one. The equation in p has to be resolved into two factors as follows:

$$(p^2+2A_1p+A_1^2+B_1^2)(p^2+2A_2p+A_2^2+B_2^2)=0$$

in which A_1 and A_2 are the damping coefficients of the two oscillations and B_1 and B_2 are the two frequencies or periodicities, *i.e.*, $\frac{2\pi}{\text{period}}$.

When the damping is zero let p_1 and p_2 be the periodicities, then

$$(p^2+p_1^2)(p^2+p_2^2)=0,$$

or

$$p^4+p^2(p_1^2+p_2^2)+p_1^2p_2^2=0.$$

*The conditions that an equation $p^4+ap^3+bp^2+cp+d=0$ shall represent two damped oscillations have been stated by Routh as (1) all coefficients must be positive and (2) $abc-c^2-a^2d>0$. This last condition leaves a very large range of choice of the coefficients which must therefore be decided by other considerations.

If in equation (4) we make $n=0$, the damping disappears and $p_1^2+p_2^2=Gshd$, $p_1^2p_2^2=\omega\cos\lambda\,.\,hdG^2\omega_1\omega_2$, and if $p_1=10p_2$, then $p_1^2\doteqdot Gshd$ and $p_2^2\doteqdot\dfrac{\omega\cos\lambda\,.\,G\omega_1\omega_2}{s}$.

The second condition that equation (4) shall represent two damped oscillations is

$$p_1^2>\frac{\text{coeff. of } p}{\text{coeff. of } p^5}>p_2^2,$$

or

$$Gshd>2Ghd\,.\,\frac{\omega_1\omega_2}{s}>\omega\cos\lambda\,.\,G\omega_1\omega_2\frac{1}{s},$$

or

$$(\omega_1+\omega_2)^2>\omega_1\omega_2,\ \text{and}\ 2h(\omega_1-\omega_2)>\omega\cos\lambda.$$

The first is necessarily true and the second can be arranged by suitably choosing the torsion element.

The periodicity in the ordinary single-gyro compass is given by

$$p^2=G\omega_1\,.\,\omega\cos\lambda=\left(\frac{2\pi}{85\text{ minutes}}\right)^2,$$

hence we see that if the gravity control of each gyro in the two-gyro compass is to be approximately the same as in the one-gyro compass—which must be the case if the ballistic deflection is to take the compass on to the virtual meridian, —then the periodicity of the two-gyro compass must be less than that of the single-gyro compass in the ratio of $1:\sqrt{2}$, *i.e.*, the period of the compass is increased in the ratio $\sqrt{2}:1$, or instead of being 85 minutes the period will be 120 minutes.

Assume arbitrarily the ratio $p_1:p_2=10$ and also assume a ratio of $\omega_1:\omega_2$ say 1.1, that is, make the stable gyro with 0.9 of the angular momentum of the unstable gyro,

$$p_1^2=100p_2^2,\quad Gshd\doteqdot\frac{100\,G\omega_1\omega_2\omega\cos\lambda}{\omega_1+\omega_2},$$

or

$$h(\omega_1-\omega_2)\doteqdot\frac{100\omega_1\omega_2}{(\omega_1+\omega_2)^2}\omega\cos\lambda,\ \text{and}\ \omega_1=1.1\omega_2;$$

therefore

$$h\omega_2\doteqdot 250\omega\cos\lambda$$

or, at $\lambda=51°$,

$$h\omega_2=\frac{2\pi}{9\text{ minutes}}.$$

Hence a twist of 1 radian in the torsion element connecting the gyros would make the unstable gyro precess with a speed equal to one revolution in 9 minutes. This determines the stiffness of the torsion element.

It still remains to determine the stiffness of the suspension through which the damping torques are transmitted and to calculate the damping coefficients A_1 and A_2.

If we write equation (4) as

$$p^4+ap^3+bp^2+cp+d'=0,$$

the values of A_1 and A_2 the damping coefficients, are given by

$$A_1=\frac{a\left(p_1^2-\frac{c}{a}\right)}{2(p_1^2-p_2^2)}, \qquad A_2=\frac{a\left(\frac{c}{a}-p_2^2\right)}{2(p_1^2-p_2^2)}.$$

Then

$$\frac{\frac{c}{a}}{p_1^2}=2\frac{Ghd\omega_1\omega_2}{Gs^2hd}=\frac{2\omega_1\omega_2}{(\omega_1+\omega_2)^2}=\frac{2\times1.1}{(2.1)^2}=\frac{1}{2},$$

therefore

$$A_1=\frac{a\left(1-\frac{1}{2}\right)}{2\left(1-\frac{1}{100}\right)}=\frac{a}{3.96}, \quad A_2=\frac{a\left(\frac{1}{2}-\frac{1}{100}\right)}{2\left(1-\frac{1}{100}\right)}=\frac{a}{4.05}.$$

A_2 is the damping coefficient of the period of the compass. Suppose that we damp out the oscillation in one period, say, to 0.2%, *i.e.* $e^{-2\pi}$; then

$$A_2=\frac{2\pi}{120 \text{ minutes}},$$

and $a=ns=\frac{2\pi}{30}$, therefore $n\omega_2=\frac{2\pi}{63 \text{ minutes}}$. Hence the twist in the suspension due to one radian difference of tilt will impart a precessional velocity to the unstable gyro equal to one revolution in 30 minutes;

$$A_1=\frac{2\pi}{30\times3.96}=\frac{2\pi}{119 \text{ minutes}}.$$

Hence the short period is practically damped out in the same time as the long period.

It remains only to note that the torque in the suspension *assists* the precession of the unstable gyro when its tilt exceeds that of the stable one, otherwise it resists this precession. That is, if the north end of the unstable gyro is tilted above the north end of the stable one, the twist in the suspension forces the north end of the unstable gyro to the west.

The particular arrangement shown in Figures 1 and 2 has been selected because it makes the mathematical equations symmetrical for the two gyroscopes and introduces the minimum number of terms. The introduction of the viscous connection between the gyroscopes adds more terms as also does any other arrangement of the roller contact. This arrangement, however, serves to typify the mathematical methods.

In this connection I would like to express my indebtedness to my mathematical colleague Professor Wm. Burnside, F.R.S., for valuable assistance in this and many other problems.

Summary

The following is a summary of the results of the above discussion.

It is shown how the two periods of oscillation of a structure carrying two gyroscopes, such as a gyro-compass, may be calculated, also the conditions for their mutual damping.

The undamped periods of the compass described are 120 and 12 minutes in latitude 51°.

Gravity control is such that $G\omega_1 = \frac{2\pi}{3.14 \text{ minutes}}$, *i.e.*, if the gyro is tilted 30° from horizontal its axis will describe a cone round the vertical at 60° angle in $3.14 \times \sqrt{3}$ minutes $= 5.4$ minutes.

The unstable gyro has 10% greater angular momentum than the stable one.

The torsion element connecting the two gyros about the vertical is such that a twist of 1 radian would give to the unstable gyro a velocity of precession of 1 revolution in 9 minutes.

The suspension and the linkage to the switch producing the torque in the suspension are such that a tilt of one gyro through 1° relatively to the other produces a rate of precession of the unstable gyro of 0.64 degrees per minute.

The long period oscillation is damped to 0.2% in 120 minutes.

The short period oscillation is damped to 0.2% in 120 minutes.

The compass settles on the meridian with both rotor axes horizontal provided that weights are placed on the north end of the stable and south end of the unstable gyros sufficient to produce the precession $\omega \sin \lambda$, the velocity of the meridian about the vertical.

SUR L'ÉTAT ACTUEL DE LA BALISTIQUE EXTÉRIEURE THÉORIQUE

PAR M. P. CHARBONNIER,
Ingénieur Général de l'Artillerie Navale, Ministère de la Marine, Paris, France.

I

1. Avant la guerre, comme on sait, beaucoup d'artilleurs tenaient la Balistique et les balisticiens en piètre estime. Leur arme, du type du 75 français par exemple, était une machine de précision, à grand rendement et d'un maniement sûr. Mais l'emploi sur le champ de bataille en était basé sur une hypothèse—qui d'ailleurs est une erreur courante dans toutes les branches de l'art militaire—c'est que l'ennemi se plierait docilement à toutes les règles du jeu: le *tir de plein fouet* aux faibles distances pour lequel toute la machine était conçue et combinée et d'où toute science balistique était systématiquement exclue.

Avec le tir à longue portée, qui devint la règle sur le front, avec l'emploi de canons de tous modèles et de tous calibres, avec l'adoption de projectiles de formes améliorées, avec la nécessité d'un réglage permanent, sujet aux variations dues à de multiples causes perturbatrices, avec le tir contre aéronefs enfin qui posait des problèmes pratiques que l'empirisme ne pouvait même pas aborder, tout changea. Jusque sur le champ de bataille, la nécessité des méthodes scientifiques de la Balistique rationnelle réhabilitée s'imposa.

En France, l'Artillerie Navale, dans son polygone d'expériences de Gâvre était restée, de tout temps, fidèle aux Sciences de la Balistique: habituée à traiter rationnellement toutes les questions de tir de l'artillerie de bord, elle se trouvait en face des mêmes problèmes qui se posaient maintenant à l'artillerie de terre. C'était donc un terrain qui lui était très familier et pour résoudre ces problèmes, elle avait des méthodes d'expérience et des procédés de calcul sûrs et éprouvés. Ainsi, par une évolution logique et sans saut brusque, la Balistique d'avant-guerre put-elle se transformer en une Balistique, nouvelle si l'on veut, à certains égards, mais qui ne diffère de l'ancienne que par l'extension de méthodes connues aux conditions actuelles du tir. C'est le sort de toutes les sciences appliquées, dont tel ou tel chapître se développe plus ou moins suivant les besoins du moment: l'essentiel c'est que la science soit assez bien assise, dans ses principes, pour pouvoir suivre aisément la contrainte de la pratique et la demande de sa clientèle.

2. Les balisticiens, d'ailleurs, qui travaillaient avec acharnement et méthode dans leurs polygones et dans leurs bureaux d'études furent parfois assez maltraités. Lors du tir des Berthas sur Paris, les journaux quotidiens et de graves revues, se gaussèrent de ces artilleurs naïfs, prétentieux et ignorants, qui ne

connaissaient pas la loi de la diminution de la densité de l'air avec l'altitude, circonstance qui expliquait si évidemment la réussite du coup de maître des Allemands.

Et des savants qualifiés, dans des circonstances solennelles—que je ne veux pas rappeler ici—appuyèrent ces dires de l'autorité de leurs paroles. Que n'ont-ils, à cette occasion, mis à profit l' *a b c* de la méthode scientifique, qui est de recourir aux sources ou tout au moins de s'informer auprès des spécialistes! Ils auraient appris que la question avait déjà été traitée par les balisticiens du XVIII^e siècle, que de Saint-Robert l'avait longuement étudiée vers 1860, qu'en France, pour me borner à ce pays, depuis 1887, la méthode de calcul des trajectoires par arcs successifs, mise en pratique courante pour l'établissement de toutes les tables de tir de la Marine, adoptait la loi de résistance, fonction de l'altitude, qu'on emploie encore aujourd'hui. Ils auraient pu constater que, vers 1892, lorsque la Commission de Gâvre obtînt, avec le canon de 16 cm de 90 calibres des vitesses de l'ordre de 1200 m.s., on avait calculé, avec cette loi de Δ variable avec l'altitude, des trajectoires à longue portée qui franchissaient aisément le Mont-Blanc ou le Pas-de-Calais, et qui présentaient le phénomène, prévu par la théorie, d'un point de vitesse *maximum* sur la branche descendante, après le point classique de vitesse *minimum*.

Le tir des Berthas, au point de vue de la Balistique Extérieure ne présentait donc absolument rien d'imprévu, sitôt qu'on admettait la possibilité de réaliser des vitesses initiales de l'ordre de 1600 m.s. C'est en cette réalisation de Balistique Intérieure que la réussite des Allemands offre un véritable intérêt et fait aborder à l'artillerie de l'avenir un domaine nouveau. Pleine de difficultés pratiques et d'imprévu mécanique sera la mise en oeuvre et en service de ces dures conditions de tir, qui entraîneront des modifications profondes pour le matériel d'artillerie qui s'étudie et se prépare actuellement. Mais ni la Balistique Extérieure, ni la Balistique Intérieure ne se trouvent débordées par ces nouvelles conditions de tir: les bases rationnelles et expérimentales de l'une et de l'autre de ces sciences sont assez sûres pour que les extrapolations, en vue des applications nouvelles, nécessitent autre chose que quelques expériences de vérification et d'extension des constantes qui figurent dans leurs équations.

3. La Balistique Extérieure, au tableau de laquelle nous nous bornons dans cette note, est souvent dite la soeur terrestre de la Mécanique céleste; c'est une soeur aînée, puisque Galilée la créa sous sa forme moderne, bien avant que Newton n'en généralisât la loi fondamentale en l'étendant au monde des planètes. La parenté des deux sciences à leur origine est si étroite que le second volume des «Principes de la Philosophie naturelle» de Newton n'est en réalité qu'un pur traité de Balistique.

Les deux sciences se séparèrent dans la suite: l'une qui s'occupe d'objets immuables, éternels et parfaits, se développe dans le pur domaine des mathématiques et des lois absolues; l'autre, qui participe davantage des imperfections terrestres et est utilisée à des fins très peu spéculatives, marche terre à terre et pas à pas, hésite, revient en arrière, renie parfois son essence rationnelle pour ne plus vouloir être qu'une pure discipline expérimentale; comme science d'application, elle est obligée de se plier aux exigences, aux idées, aux faits, aux modes du moment.

L'historique de la Balistique Extérieure est d'abord intimement liée à l'histoire de la Mécanique rationnelle et de la découverte des lois fondamentales de la Dynamique. Quant elle s'est constituée comme science distincte, son histoire parait se borner à l'étude, à travers les temps, d'une simple équation différentielle qui, au point de vue mathématique, est sans grand intérêt pour les progrès de la science pure. Mais si on considère son histoire au point de vue relatif de ses applications et des efforts des balisticiens pour adapter leur outillage scientifique aux besoins changeants de la pratique et de l'emploi, elle prend une vie et un intérêt de premier ordre*.

II

4. Les équations différentielles de la Balistique Extérieure sont très faciles à écrire et se ramèneraient à des quadratures si l'on savait intégrer l'une d'elles, l'hodographe $\frac{du}{d\tau} = \frac{cvF(v)}{g}$. De la non intégrabilité, en général, de cette équation, découlent toutes les misères et tous les travaux des balisticiens: au point de vue mathématique, cela les a conduit nécessairement et inconsciemment parfois, à l'emploi exclusif des développements en série; toute l'histoire de la Balistique Extérieure n'est que celle des essais tentés pour exploiter l'une ou l'autre des séries possibles, pour les réduire à un nombre minimum de termes de calcul aisé, pour évaluer les rayons de convergence de ces séries, pour déterminer à quelles conditions initiales chacune s'applique légitimement.

On peut distinguer quatre groupes de séries balistiques:

1° les séries qui sont issues d'un des cas-limites, quelle que soit la fonction $F(v)$ et qui admettent la solution finie de ce cas-limite comme premier terme:

$c=0$ balistique du vide,

$g=0$ balistique horizontale,

$\tau=\pm\frac{\pi}{2}$ balistique verticale,

$\tau\sim 0$ tir de plein fouet.

2° les séries où $F(v)$ est voisin d'une fonction qui rend intégrable l'hodographe. En pratique, $F(v)=B_n v^n+\epsilon\psi(v)$ [avec $\epsilon\psi(v)$ petit, produisant une correction dite *erreur balistique*]. C'est la méthode de la résistance monome.

3° les séries qui n'ont point de caractère balistique proprement dit, séries de Taylor, de Maclaurin, d'Euler-Maclaurin, etc. . . .

*Pour l'historique de la Balistique Extérieure, on pourra consulter: (*a*) *Historique de la Balistique Extérieure à la Commission de Gâvre*, par le Ct. P. Charbonnier. Revue Maritime, 1906; traduit en anglais dans le Journal of the U.S. Artillery sous le titre *The Gâvre Commission. An historical sketch of the progress of Exterior Ballistics* (1907) et (*b*) *Essais sur l'histoire de la Balistique.* Mém. Art. Fanç. (1927).

4° les séries qui ne sont valables que dans la voisinage d'un point particulier de la trajectoire (points de vitesse ou de courbure minimum, etc. . . .).

5. Je ne dirai rien de la Balistique du vide ($c=0$), tant ce chapitre de la mécanique élémentaire est classique et complet; ni de la Balistique horizontale ($g=0$) où quelques propriétés qui étonnaient les balisticiens d'autrefois, tel Varignon, sur l'origine ou la fin du mouvement ($v=0$, $v=\infty$) sont maintenant familières aux analystes; ni de la Balistique verticale $\tau=\pm\frac{\pi}{2}$ où la discussion ne présente rien qui ne soit connu depuis longtemps et amené au degré de perfection désirable.

6. L'étude des propriétés générales des trajectoires d'après leurs équations différentielles constitue un chapitre important où la Balistique moderne a pu ajouter quelques résultats intéressants.

Une série de théorèmes dont les principaux sont dus à de Saint-Robert et à Siacci peuvent être liés aujourd'hui par une discussion complète (existence ou non existence du point de vitesse minimum suivant la nature de la fonction $F(v)$, point de courbure minimum, étude de l'extrémité des branches ascendante et descendante de la trajectoire, etc. . . .).

En particulier, on a pu reconnaître l'existence de certaines trajectoires sans *sommet*, qui peuvent se présenter dans des cas particuliers du tir des canons d'avion.

Les tentatives qui ont été faites par les mathématiciens pour intégrer l'équation de la Balistique forment un ensemble important. Bernoulli, d'Alembert, Siacci, le Colonel Jacob, M. Appel, M. Ouivet, M. Esclangon ont, par diverses transformations, montré les formes canoniques, en Analyse, auxquelles les équations sont réductibles.

Le problème de l'intégration de l'hodographe, impossible dans le cas général [$F(v)$ quelconque], a été transformé lorsqu'on s'est demandé, inversement, pour quelles formes de la fonction de résistance cette intégration était possible. Siacci avait fait connaître une dizaine de fonctions donnant cette satisfaction au balisticien, mais son analyse où éclataient de remarquables dons d'intuition mathématique, avait plutôt suggéré la solution que donné une théorie d'ensemble du sujet.

Cette méthode générale a été trouvée récemment et M. Drach, par une application extrêmement intéressante de la théorie des fonctions analytiques, a pu déterminer les fonctions (dont celles indiquées par Siacci ne sont que des cas particuliers), où cette intégration est possible. M. Denjoy a, sur certains points, poussé plus avant le problème résolu par M. Drach.

Ajoutons que cette conquête, qui fait honneur aux mathématiciens, ne semble pas susceptible d'offrir aux balisticiens, jusqu'à nouvel ordre, des ressources nouvelles pour les applications. C'est que les calculs sont très compliqués et souvent presque uniquement symboliques ou présentés sous forme de fonctions de séries; il y a incompatibilité générale des formes trouvées pour $F(v)$ avec la loi réelle; de plus, on se heurte à l'impossibilité d'aller plus loin que l'hodographe intégrée, en laissant subsister, pour les autres éléments, toutes les difficultés de quadrature même tabulaire; et, après cet effort, la vieille loi monome $F(v)=B_n v^n$ subsiste seule comme acceptable pour le développement des théories balistiques.

D'intéressants travaux ont porté sur les *intégromètres balistiques* qui permettent de résoudre graphiquement le problème de la trajectoire, tels que les *intégromètres à lame coupante* du Colonel Filloux et du Colonel Jacob. Une très curieuse *machine à calculer les trajectoires*, inventée par le Capitaine Perrin, permet de résoudre, au moins théoriquement, le problème numérique du calcul d'une trajectoire quelconque.

Le problème *balistique inverse*, qui part des équations finies supposées connues de la trajectoire pour en tirer les lois élémentaires du mouvement, c'est-à-dire les hypothèses impliquées dans la solution donnée, a été posé par Lagrange et résolu par des formules qu'il a établies. Ce problème a une grande importance pratique, puisqu'il permet au théoricien de mesurer, en quelque sorte, l'approximation des formules empiriques ou semi-empiriques qu'on lui propose. Il permet de discuter par exemple l'hyperbole de Newton ou la trajectoire cubique de Piton-Bressant, les facteurs de la trajectoire de Siacci, ceux de M. Sugot, etc. . . .*

III

7. Le cas d'une résistance monôme $F(v)=B_n v^n$, hypothèse où l'hodographe est intégrable très simplement, a constitué toute la Balistique connue jusque vers le milieu du siècle dernier. Innombrables sont donc les travaux des balisticiens sur ce problème, que les modernes n'ont d'ailleurs pas cessé d'étudier. Il n'a point perdu tout son intérêt puisqu'on peut prendre, comme formules de départ, dans les calculs balistiques actuels et comme premier terme de la série, les résultats obtenus dans sa solution.

On sait que la théorie générale introduit une fonction $\xi_n(\tau)=\int_o^\tau \frac{d\tau}{\cos^{n+1}\tau}$ dont les balisticiens ont dressé des tables pour toutes les valeurs utiles de n et que la solution complète du problème balistique a conduit à la présenter sous forme de tables numériques (Euler, Otto, Bashforth, Zabouski, Siacci, etc. . . .).

De très nombreux travaux ont été consacrés à l'étude particulière de la courbe balistique pour des valeurs spécifiées de n: $n=0, 1, 2, 3, 4$. La première, $n=0$, qui représente la trajectoire d'un point pesant avec frottement sur un plan incliné est résoluble par les procédés élémentaires de l'analyse. La seconde, $n=1$, que Huyghens trouva, et qui exprime tous ses éléments en termes finis a donné lieu à des théorèmes intéressants de Varignon, Newton, Lambert, Greenhill, etc. . . . Elle a pu servir avantageusement de point de départ à certains calculs modernes de trajectoires par arcs.

*Pour tout ce qui concerne les théorèmes généraux de la Balistique Extérieure on pourra consulter le *Traité de Balistique Extérieure,* Tome I, 1921, de M. l'Ingénieur Général Charbonnier.

Le cas de $n=2$, c'est-à-dire d'une résistance quadratique, a été de tous le plus étudié. On sait que les éléments de l'arc, à part la longueur de l'arc s, c'est-à-dire x, y, t, ne sont pas intégrables. D'illustres auteurs, Newton, Borda, Legendre, Français, M. de Sparre, M. Bassani ont présenté des solutions approximatives extrêmement ingénieuses, pour pousser jusqu'au bout l'intégration et fournir le premier terme de séries par la substitution d'un élément qui favorise l'intégration à l'élément rebelle. Il est possible de généraliser toutes les méthodes de ces auteurs en montrant que toutes reviennent à remplacer dans l'hodographe $\xi(\tau)$ par une fonction trinome $a_0+a_1 \text{ tang } \tau+a_2 \text{ tang }^2\tau$ où le choix des coefficients a_0, a_1, a_2 caractérise chacune des méthodes proposées.

Le cas de $n=3$ a fait l'objet de travaux mathématiques importants et il présente l'intérêt d'être soluble à l'aide des fonctions elliptiques. M. Greenhill est l'auteur de cette théorie que M. de Sparre a étendue. Le Commandant Demogue, à Gâvre, a ajouté à cette théorie, outre des développements en série intéressants, un très beau théorème qui résoud d'une manière générale la question de la convergence de la série en x.

Le cas de $n=4$ a été aussi fort étudié par Piton-Bressant, le Général Zabouski, M. de Sparre. On peut ramener encore à l'unité les résultats acquis par une transformation de $\xi_n(\tau)$ en un polynome en tang τ, comme nous l'avons vu pour $n=2$.

Si l'on examine enfin l'ensemble des recherches des géomètres sur le problème de l'intégration dans le cas d'une résistance monome, on peut remarquer qu'ils se sont toujours efforcés, comme on l'a dit, de résoudre le problème en posant:

$$\xi_n(\tau)=a_0+a_1 \text{ tang } \tau+a_2 \text{ tang }^2\tau$$

le nombre de terme de second membre étant égal à 3 au plus. Or, il n'est résulté de ce point de départ que des calculs particuliers, des formules très compliquées, des méthodes sans aucune généralité et sans application pratique possible. Mais on peut prendre le problème autrement en posant:

$$\text{tang } \tau=m_0+m_1\xi_n(\tau)+m_2\xi_n^2(\tau)+\ldots$$

le second membre étant un polynome en $\xi_n(\tau)$, d'un nombre de termes aussi grand qu'on veut suivant l'approximation désirée. On peut constater alors que toutes les méthodes s'unifient, que les procédés de calcul deviennent généraux, que la précision est aussi grande que l'on veut et que la solution est obtenue par des fonctions monomes les plus simples de l'analyse et aisément calculables par le praticien.

C'est donc la solution définitive d'un problème posé depuis les premiers temps de la Balistique.

On sait que la préoccupation des balisticiens d'autrefois était essentiellement d'obtenir l'équation finie $y=f(x)$ de la trajectoire. Les développements en série de Maclaurin avaient été poussés très loin, surtout par Français, et on s'efforçait, dans le cas le plus intéressant ($n=2$), de trouver la loi de récurrence des termes successifs de cette série dont l'équation de Borda donnait le premier terme. Le Général Otto, en 1842, avait bien donné une formule de récurrence, mais sans démonstration.

Le problème fut posé à Gâvre, à un savant analyste, le Commandant Demogue, qui non-seulement le résolut dans le cas particulier d'Otto, mais encore dans le cas général d'une résistance de degré n, et pour tous les éléments de la trajectoire.

On voit ainsi, par ce qui précède, que les balisticiens modernes, tout en estimant sans doute, que, dans l'état actuel des nécessités pratiques de la Balistique, ces problèmes sont aujourd'hui plutôt de l'ordre de la spéculation et de la curiosité, ont su combler les lacunes laissées par leurs devanciers et constituer, pour la Balistique, un corps cohérent de doctrine où rien n'est laissé dans l'ombre et auquel l'avenir saura peut-être faire appel pour le développement même des applications pratiques*.

IV

8. Le *tir de plein fouet*, qui suppose que la partie utile de la trajectoire entoure le sommet ($\tau=0$), introduit, comme on sait, les fonctions balistiques de Siacci, où se trouve conservée la forme même de la fonction $F(v)$. On connaît les applications extrêmement nombreuses et intéressantes de ces théories, qui ont donné lieu à d'innombrables travaux des balisticiens modernes. Tous ces travaux peuvent maintenant être réunis par le fil d'une interprétation générale et logique.

Siacci avait procédé par une méthode semi-empirique d'approximation et, en cherchant le second terme de la série, il avait donné une approximation qui supposait cette série ordonnée suivant les puissances du coefficient balistique c. Or, c'était là une hypothèse nouvelle et particulière. En réalité, le développement, dont les formules de Siacci sont le premier terme, résulte naturellement du développement en série, du cosinus dans l'hodographe qui s'écrit:

$$\frac{d\tau}{\cos^2\tau}=\frac{g}{c}\,\frac{du}{uF}-\frac{g}{2c}\left(\frac{uF'}{F}-1\right)\operatorname{tang}^2\tau\frac{du}{uF}$$

et qu'on intègre par termes successifs; il s'introduit ainsi, pour le 2^e terme, de nouvelles fonctions balistiques, où la fonction $F(v)$ figure sans passer par sa dérivée. Des tables de ces fonctions existent.

Sans qu'il soit besoin d'insister sur les détails, on peut dire que la théorie du tir de plein fouet, non-seulement dans sa généralité, mais dans toutes les formes où la spécification ou la restriction de la loi de la résistance l'ont déguisée, se trouve maintenant bien définitivement assise.

On peut jeter une grande clarté sur toutes les variantes proposées pour la méthode de Siacci ou antérieures à ce savant, en remarquant que le procédé d'intégration de l'auteur italien revient, au fond, si on le généralise, à remplacer dans le second membre de l'hodographe écrit:

$$\frac{d\tau}{\cos^2\tau}=\frac{g}{c}\,\frac{du}{u\cos\tau\,F\left(\dfrac{u}{\cos\tau}\right)}$$

*Voir *Traité de Balistique Extérieure*, T. II, 1927.

les deux cosinus, l'un par une constante λ, l'autre par une constante μ, dont on ne connaît, a priori, que les valeurs limites.

On peut remarquer que, après la Balistique quadratique dont le règne finit vers 1840, tous les balisticiens tels que Piton-Bressant, Hélie, Didion, Duchesne, de Saint-Robert et enfin Siacci n'ont au fond employé que la méthode du tir de plein fouet, et qu'il suffit de choisir d'une manière ou d'une autre, les deux constantes λ et μ pour obtenir l'infinité de variantes présentées par les auteurs, et que le manque de cette clef rendait de la plus extrême confusion.

La théorie du tir de plein fouet présente de nombreuses et intéressantes solutions de problèmes de tir, des applications pratiques au choix balistique des armes en projet dues au Commandant Batailler, de beaux théorèmes géométriques sur l'affinité des trajectoires dûs au Colonel Emery et au Capitaine Garbasso, un important chapitre sur les propriétés et l'usage des fonctions secondaires etc., . . . etc. . . .

Ajoutons que la Commission de Gâvre s'était constituée à la veille de la guerre, d'après les nouvelles théories du tir de plein fouet, en tenant compte des deux premiers termes de la série, et avec la loi générale de la résistance de l'air $F(v)$ (compte tenu de la variation de la densité de l'air avec l'altitude), un outillage de calcul complet (trajectoire finie et coefficients différentiels) allant jusqu'à un angle de projection de 20°. Cet outillage répondait complètement et admirablement au problème pratique de l'établissement des tables de tir de l'artillerie de bord.

9. Les autres séries balistiques ont moins d'importance pratique, au moins jusqu'ici. Disons-en seulement quelques mots.

Autour d'un point, qui n'est pas le sommet, on trouve des formules intéressantes dont le premier terme est analogue aux fonctions de Siacci, mais dont le second est fort différent des fonctions du second terme du tir de plein fouet.

Le tir *quasi-vertical* $\tau = \pm \dfrac{\pi}{2}$ introduit de nouvelles fonctions balistiques mais à deux arguments, telles que $\log I_{w_0}^{w} = -\displaystyle\int_{w_0}^{w} \frac{g\,dw}{w(g+cF)}$. Le second terme des séries est connu. Les formules anciennes de Poisson ($n=2$), celles plus récentes de M. Vito Volterra, d'autres dues au Colonel Bianchi, l'une et l'autre applicables à un arc d'inclinaison quelconque de la trajectoire, rentrent dans cette classification.

Le *tir tendu à grande vitesse* (c très grand), le *tir courbe à faible vitesse* (c très petit) donnent au balisticien des séries qui ont leur application dans certains cas. (C'est de cette dernière que Siacci avait tiré la valeur de son coefficient β du tir de plein fouet.)

Non seulement la théorie trouve son intérêt dans ces solutions et ces développements, mais la pratique même y puise des ressources utilisables dans certains cas. En outre, on voit se classer des idées et des formules qu'on rencontre dans les écrits des balisticiens et les rapports des Commissions d'expériences, qui, si on ne considérait pas l'ensemble maintenant solidement et logiquement établi, ne pourraient paraître qu'aberrantes et suspectes.

V

10. La Balistique d'avant-guerre était exclusivement, peut-on dire, la Balistique du *point de chute*, seul point important, avec son voisinage immédiat, pour l'emploi des armes d'alors (tir percutant et tir fusant) et tout le développement historique de cette Balistique, dans ses divers aspects théoriques, s'explique par cette considération.

La Balistique née pendant la guerre est la Balistique de la *trajectoire totale*. Aucun point n'est plus privilégié ni plus intéressant que l'autre, soit qu'on considère le tir contre aéronefs ou le but se meut dans un espace à trois dimensions, et le projectile dans un espace à quatre dimensions (car la loi spéciale de combustion de la fusée introduit une variable de plus); soit que, dans le tir à longue distance contre objectifs terrestres, il s'agisse de corriger la portée des éléments météorologiques variables d'une couche à l'autre et que l'observation fait de mieux en mieux connaître.

Au point de vue théorique, le problème nouveau, bien plus général, paraît infiniment plus compliqué que l'ancien dont l'énoncé était restreint: or, il n'en est rien. Plus, en effet, n'est besoin de cette course à l'équation finie de la trajectoire qui préoccupait presque exclusivement les balisticiens d'autrefois. Eut-on cette trajectoire finie, aussi rigoureusement qu'on a l'équation du vide, pour fournir aux combattants les données nécessaires au réglage de leur tir, il serait nécessaire de la découper ensuite en arcs suffisamment petits.

De là aussi résulte une grande simplification des calculs élémentaires. Nous avons vu que toutes les théories balistiques élaborées par trois siècles de balisticiens, se ramenaient, en réalité, à des développements en série de nature et d'arguments divers suivant les cas. Or, pour le calcul des petits arcs, actuellement fin essentielle de la Balistique, tous ces travaux divers prennent leur importance réelle, acquièrent leur zone d'applicabilité légitime et la convergence de ces séries, faible, douteuse ou nulle quand on les applique à la trajectoire entière devient rapide et sûre dans les limites actuellement utiles.

Comme la variété des développements possibles et de leurs variantes est presque illimitée, le nombre possible des procédés de calcul des trajectoires par arcs successifs sera lui-même très considérable, puisque chacun des développements connus pourra fournir la base de ces calculs modernes.

L'arsenal du balisticien se trouve donc très bien fourni—méthodes de Taylor, d'Euler-Maclaurin; résistance monome; méthode de Siacci—méthode du tir avec fonctions à 2 variables de Bianchi, issues du tir nadiral ou du tir zénithal, etc. . . .

11. Je dirai quelques mots du développement, en France, de cette Balistique moderne. Nous avons vu qu'à la veille de la guerre, la Commission de Gâvre s'était constituée, sur la base des théories du tir de plein fouet, étendu au 2^{e} terme de la série et au terme complémentaire tenant compte de la variation de densité de l'air, un outillage balistique parfait, aussi satisfaisant au point de vue théorique qu'au point de vue pratique et embrassant toute la Balistique des canons jusqu'à l'angle de projection de 20°.

Mais la guerre fit soudain sortir le tir des limites étroites du tir de plein fouet qui était presque seul envisagé jusqu'alors par la Guerre et par la Marine.

Le tir courbe devint la règle; le tir contre avions posait de plus en plus aux balisticiens, avec une généralité inattendue, des problèmes tout nouveaux et extraordinairement difficiles. D'autre part, après une période de scepticisme contre les méthodes savantes de préparation du tir qu'avait toujours préconisées Gâvre et que la Marine appliquait, les combattants, au front, éprouvaient leur valeur pratique; ils les considéraient bientôt comme transformant l'art du tir et ils réclamaient aux balisticiens de l'arrière une précision de plus en plus grande et des données balistiques de plus en plus nombreuses et délicates.

A Gâvre, un travail balistique acharné, toujours guidé par les mêmes méthodes scientifiques, se poursuivait: les tables de tir se multipliaient à l'infini; elles s'étendaient en portée; les coefficients différentiels multiples nécessaires pour des corrections de tout genre se calculaient; des méthodes nouvelles se créaient. Le Président de la Commission de Gâvre avait pu grouper autour de lui, avec quelques ingénieurs ou Officiers d'artillerie spécialistes de la Balistique, tels que MM. Garnier, Sugot, Anne, Demogue, Lyon, etc., . . . un certain nombre d'universitaires distingués, tels que MM. Esclangon, Haag, Valiron, Marcus, Châtelet, Janet, P. Lévy, Fort, Denjoy, Kampé de Fériet, etc. . . .

Un organe scientifico-technique fut créé sous le nom de «Mission balistique des tirs aériens» (M.B.T.A.). Ce travail en commun de savants et d'Officiers tous attachés avec une même passion à leur tâche et cherchant par leur science et par leur zèle à contribuer au salut du pays, fut très fécond. C'est à Gâvre que se constituèrent, en particulier, les méthodes balistiques du tir contre avions et que se calculèrent tous les abaques principaux et correctifs que nécessite ce genre de tir si difficile, où on doit prendre en considération le vent variable avec l'altitude, la pression atmosphérique variant suivant une loi quelconque en altitude, le coefficient balistique du projectile variable sur toute l'étendue de la trajectoire, les altérations du point d'éclatement dues aux multiples causes inhérentes au canon, à la fusée ou à l'atmosphère, etc. . . . etc. . . . La caractéristique de la méthode employée pour ce travail de haute science appliquée à un problème du plus grand intérêt pratique fut que les théoriciens restèrent toujours en contact étroit et permanent avec les utilisateurs du matériel contre avions: le Colonel Pagezy, le metteur sur pied bien connu de la défense anti-aérienne, et le réalisateur du matériel approprié, fut au point de vue balistique, le client constant et satisfait des balisticiens de Gâvre.

12. Nous avons dit que le choix du point de départ des méthodes de calcul des trajectoires par arcs, parmi les nombreux développements en série possibles et connus, était à peu près indifférent. Tout naturellement, les premiers efforts de la Commission de Gâvre s'appliquèrent au perfectionnement de l'ancienne méthode qu'elle avait utilisée pendant de nombreuses années: la méthode d'Euler-Otto, adaptée par Hélie à une fonction quelconque de résistance et introduite à Gâvre par le Capitaine Gossot en 1887 (cas d'une résistance quadratique). Mais, tel qu'il se présentait à l'origine, ce mode de calcul était fort incomplet et assez peu sûr. Les recherches des Officiers et des savants de Gâvre furent orientées dans les voies suivantes, où tout était à faire:

1° comment doit-on opérer le fractionnement d'une trajectoire par arcs pour avoir une précision donnée d'avance?

2° quelle méthode de calcul approximatif peut conduire à la solution pratique la plus rapide et la plus sure?

3° quels sont les procédés qui permettent d'obtenir, en chaque point, les coefficients différentiels tout le long de la trajectoire?

Si au point de vue mathématique, le calcul des trajectoires par arcs successifs ne présente, en général, aucune difficulté et si l'exposé de chacune des méthodes peut, le plus souvent, être réduit à quelques pages, il ne nécessite pas moins du balisticien qui tient à faire passer ses méthodes et procédés de calcul dans la pratique, un effort considérable. La préparation d'une feuille de calcul correcte est un travail long et ardu. Une méthode quelconque, et elles sont nombreuses celles qui en théorie sont équivalentes, ne vaudra que si elle guide le calculateur pas à pas, sans ambiguité, sans écueil caché, si, automatiquement, le calculateur est mené par la main pour l'argument à prendre comme entrée de ses tables numériques auxiliaires, pour l'amplitude des arcs à choisir, pour le nombre de décimales à conserver, pour les vérifications à observer, etc. . . .

Donc, à ce point de vue, on peut faire une classification très nette des méthodes de calcul de trajectoires par arcs; d'une part, celles qu'on peut dire en *puissance*, qui sont en nombre presque infini, qu'on propose chaque jour, et qui, au point de vue théorique sont d'ailleurs satisfaisantes; d'autre part, celles qui sont *réalisées* complètement et ont tout prêt l'outillage énorme qu'elles exigent pour répondre à tous les desiderata du praticien.

La méthode de Gâvre, qui profite de l'intégrabilité de l'hodographe et de celle de l'arc s (dans le cas de $n=2$), fut mise au point conformément à ces directives; le terme principal obtenu par intégration fut corrigé de deux erreurs dite l'une *balistique*, l'autre *géométrique* qui y subsistent; les rayons de convergence furent calculés; la méthode des variations permit le calcul des coefficients différentiels. De sorte que toutes les questions posées se trouvèrent, au bout de peu de temps, résolues, et que la méthode, qui porte le nom de méthode G. H. (Garnier, Haag), développement logique de la méthode de Gâvre, fut mise en service et répondit à tous les besoins pratiques.

Mais des recherches, dans une voie différente, étaient menées parallèlement: on chercha une méthode en quelque sorte plus directe, quoique peut-être plus terre à terre, mais qui éliminerait les corrections introduites par les erreurs inhérentes à la méthode G. H. et éviterait en même temps les tâtonnements inévitables que cette méthode introduisait dans le cours des calculs et qui avait fait donner par les practiciens à ce procédé, le nom de «calcul à la gomme».

L'emploi systématique et discuté à fond de la série de Maclaurin, ou de celle de Maclaurin-Euler permit d'arriver au résultat cherché. La résistance de l'air est prise maintenant sous sa forme la plus générale $f=ce^{-hy}\ F(v)$ (où le second membre peut être d'ailleurs une fonction quelconque des coordonnées du point), et tout le calcul est basé sur une évaluation précise et bien définie du développement de cette fonction, dans des limites autorisées par la théorie. Toute la feuille de calcul acquiert ainsi une clarté et une symétrie parfaites et peut être confiée à un calculateur quelconque, qui est guidé automatiquement dans ses calculs et dans l'amplitude de l'arc permis pour obtenir une précision donnée.

Ce procédé s'appelle méthode G. H. M. (Garnier, Haag, Marcus), il a supplanté la méthode G. H. et est actuellement réglementaire en France.

Il a donné lieu à un exposé complet dans un livre de M. l'Ingénieur en Chef Garnier, avec toutes les variantes nécessaires, les exemples détaillés d'application numériques, le calcul de tous les coefficients différentiels, et de leurs raccordements théoriques, enfin l'examen de tous les cas où la méthode générale se trouve en défaut (tir zénithal, tir nadiral), etc. . . .

13. Dans l'historique de ces recherches—que nous ne présentons ici que très schématiquement*—on verra sans doute avec plaisir un exemple très réussi de ce qu'une collaboration, systématiquement dirigée, des techniciens et des savants peut produire dans un domaine où la science est l'auxiliaire indispensable de la technique.

Cette réunion à Gâvre de savants distingués pour qui la Balistique était une science nouvelle, mais qui les passionna bien vite, eut pour autre résultat des progrès théoriques intéressants dans les divers chapitres de cette science.

Mais les savants qui ont travaillé à Gâvre à des calculs très longs, très minutieux et très arides s'accordaient à reconnaître qu'au point de vue de la science pure, qui était leur domaine jusqu'ici, cette incursion dans les régions pratiques du calcul numérique, a été loin de leur être inutile. Ils ont reconnus que la valeur d'une formule rigoureuse était chose toute relative, qu'elle disparaissait souvent quand il fallait l'appliquer à un cas concret; qu'une formule approximative était souvent d'un meilleur usage; que le calcul numérique poussé à fond non-seulement réalisait l'épreuve de la théorie, mais qu'il constituait même un instrument de recherches susceptible de susciter des idées théoriques et que c'était là qu'on pouvait souvent trouver matière à découverte; le calcul numérique est ainsi comme le domaine expérimental des mathématiques.

14. Chez tous les belligérants, le même problème balistique se posa et reçut des solutions diverses: cette diversité est la conséquence naturelle du nombre considérable de points de départ possibles. La Commission de Gâvre se tint naturellement, autant qu'elle le put, au courant de ces travaux étrangers, et en particulier, il y eut entre nous et nos alliés Anglais et Américains, principalement, échange cordial de documents et collaboration étroite et directe entre balisticiens. En général, Anglais et Américains emploient comme argument le petit espace de temps Dt caractérisant un arc au lieu qu'en France on préfère

*Consulter sur le même sujet les documents suivants:

(1°) *Sur le calcul des trajectoires et de leurs altérations*, par M. Haag, Professeur à la Faculté de Clermont-Ferrand, Journal de l'École Polytechnique; 21[e] Cahier, 1921.

(2°) *Calcul des trajectoires par arcs succesifs*, par M. M. Garnier, Ingénieur en Chef d'Artillerie Navale, Gauthier-Villars, 1921, avec une préface de M. l'Ingénieur Général P. Charbonnier.

(3°) *Conférence sur la Balistique Extérieure*, par M. M. Garnier, Ingénieur en Chef d'Art. Navale,Mém. Art. Franç., Tome 1, 1922.

(4°) *Note sur l'état actuel de la Balistique Extérieure appliquée*, par M. l'Ingénieur Général P. Charbonnier de l'Artillerie Navale, Mém. Art. Franç., 1924.

(5°) *Traité de Balistique Extérieure*, Tome II, 1927, de M. l'Ingénieur Général P. Charbonnier.

l'amplitude angulaire D_τ de l'arc. Ils ne paraissent pas s'être soucié beaucoup des questions de limitation des amplitudes et de convergence, et où nous employons 12 arcs, ils en emploient volontiers une cinquantaine. Nous autres, Français, nous semblons plus sensibles que d'autres, à une certaine esthétique dans la science, qui nous porte à chercher avant tout la clarté dans des règles logiques, absolues et générales et la satisfaction d'une certaine économie dans l'effort minimum à réaliser pour un but à atteindre.

VI

15. Les problèmes *balistiques secondaires* se rapportent comme on sait aux perturbations que subit la trajectoire du point matériel du fait de la mise en contact de trois objets: *l'atmosphère*, la *terre* et le *projectile* dont les réactions mutuelles ne figurent que par leur terme principal dans les équations différentielles du mouvement.

Tous ces problèmes d'altération sont, au point de vue des corrections numériques à faire subir à la trajectoire, sous la dépendance d'une théorie générale des *perturbations*, analogue à celle des planètes, tout au moins dans ses principes mathématiques. Elle peut affecter deux formes, équivalentes dans leur résultat: l'une, partant de la variation des constantes dans les équations finies (méthode de différentiation), l'autre, de la variation des constantes dans les équations différentielles (méthode d'intégration). Les théories sont aujourd'hui bien assises et le calcul des coefficients différentiels nécessaires est entré dans la pratique courante.

Quant aux lois à faire entrer dans l'expression de ces forces perturbatrices, l'étude de chaque cause de trouble en amène naturellement l'évaluation.

16. Pour *l'atmosphère*, la variation de Δ avec l'altitude est une perturbation minime, quand on considère des trajectoires de plein fouet, pour lesquelles on peut calculer 8 fonctions balistiques venant affecter le terme principal; on en possède des tables numériques. Mais, pour les trajectoires très élevées, cette influence de la variation de Δ avec l'altitude ne peut plus être traitée comme une faible correction. Aussi, les balisticiens d'aujourd'hui, dans leurs calculs par arcs, prennent-ils la résistance sous la forme générale $ce^{-hy}F(v)$, sans que cette complication alourdisse sensiblement leurs calculs.

Mais la variation de la densité de l'air avec l'altitude modifie notablement, surtout vers ses extrémités, la forme de la trajectoire dans l'air. Les théorèmes généraux de la Balistique classique doivent donc être revisés: de nouveaux théorèmes ont pu être donnés: ainsi un théorème sur le point de vitesse maximum, sur la branche descendante; d'autres sur les conditions d'existence des trajectoires sans sommet, etc. . . .

De son côté, à Gâvre, M. le Professeur Esclangon est arrivé à des propriétés analogues par une discussion du même problème. Un savant italien, le Général Cavalli a montré que l'intégration rigoureuse de toutes les équations du mouvement était possible, dans le cas de $n=1$, même pour Δ variant avec l'altitude. C'est un curieux résultat, qui force le balisticien à se souvenir que, dans le cas de Δ constant, ce fut là justement le premier cas qu'aborda la Balistique théorique et reçut avec Huyghens une solution rigoureuse. C'est peut-être avec raison que le Général Cavalli a intitulé son travail *La Balistique de l'avenir**.

Le problème de l'action du *vent atmosphérique* sur les projectiles a donné lieu à des travaux nombreux depuis Borda jusqu'à Didion et Siacci. C'est une très belle illustration des théories du mouvement relatif dont l'application est facile, si on se borne aux trajectoires ordinaires avec vent constant. Le problème devient bien plus difficile pour les trajectoires presque verticales; pour les trajectoires élevées, il nécessite une profonde analyse dans le cas où le vent varie avec l'altitude suivant une loi arbitraire. Le problème est complètement résolu dans les calculs modernes et on donne, dans les tables de tir, les coefficients du vent dit *par couche*.

17. Aux réactions du projectile sur l'atmosphère sont dus des phénomènes physiques, qui, sous le nom *d'onde balistique*, sont venus ajouter à l'acoustique un nouveau chapitre des plus intéressants.

La guerre a été la révélation, pour beaucoup d'artilleurs et de savants, de ces phénomènes que certains mêmes ont cru découvrir. C'était pourtant pour les balisticiens un domaine bien connu, bien exploré et bien utilisé. Faut-il rappeler l'histoire de cette découverte qui se développa presque entièrement à Gâvre: signalée pour la première fois en 1884 à Gâvre par le Capitaine Jacob†, retrouvée à Châlons avec les fusils modèle 1886 par le Capitaine Journée qui donna à la Société de Physique une explication inexacte en partie; les phénomènes sonores des projectiles reçurent leur véritable et complète explication du Capitaine de Labouret en 1886‡. Puis, presque immédiatement, une très belle application en fut faite à Gâvre par le Capitaine Gossot§ à la mesure des vitesses des projectiles, méthode qui, depuis plus de 30 ans, est employée couramment sur le polygone de Gâvre.

La théorie en fut poursuivie par des artilleurs; nous donnâmes, dans un mémoire intitulé «le Champ acoustique||,» une théorie géométrique de l'ensemble des ondes engendrées par le projectile dans son trajet dans l'air et qui comprennent, parmi les principales, 1° l'onde *neutre* $\text{tang}\,\psi = \frac{a}{v}$; 2° l'onde *balistique* $\sin\phi = \frac{a}{v}$; 3° onde sphérique d'arrière.

C'est à Gâvre également pendant la guerre, que M. le Professeur Esclangon a étudié à nouveau ce problème, avec les ressources d'un puissant mathématicien

*Voir sur ce sujet:

Ingénieur Général P. Charbonnier: *Les théorèmes généraux de la Balistique généralisée*, Mém. Art. Franç., Tome II, 1923, p. 421.

Major Général Cavalli: *La Balistique de l'avenir*, Mém. Art. Franç., Tome II, 1923, p. 459.

†Capitaine Jacob: *Les phénomènes sonores des projectiles*, Mém. Art. Marine, Tome XX, 1892.

‡Capitaine de Labouret: *Propagation du son pendant le tir*, Tome XVI, 1888.

§Capitaine Gossot: *Détermination des vitesses des projectiles au moyen des phénomènes sonores*, Tome XIX, 1891.

||Ann. de Phys. et Chim., 1906.

et d'un éminent observateur. Il a apporté une intéressante contribution à l'étude des propriétés géométriques et physiques de ces ondes qui constituent ce qu'il nomme l'*«acoustique du canon et du projectile»**.

18. L'influence de la terre sur la Balistique a donné lieu à d'intéressants théorèmes assez peu connus qui se rapportent au cas où le projectile est considéré comme une planète. C'est l'union de la Balistique et de l'astronomie dans une science unique.

Les petites perturbations dues à la variation de la *gravité* avec *l'altitude* ou la *latitude* s'introduisent aisément dans les calculs des trajectoires de l'artillerie. La *rotation* de la terre a donné lieu à un mémoire célèbre de Poisson, interprété géométriquement par de St-Robert. Dans l'air, la question a été plus récemment abordée par les balisticiens et a conduit à l'expression de forces perturbatrices qui donnent, dans les tirs modernes, des termes correctifs dont la grandeur est loin d'être négligeable.

VII

19. La Balistique du projectile oblong animé d'un rapide mouvement de rotation, pose un problème des plus difficile de la mécanique rationnelle: il a fait pendant ces dernières années des progrès extrêmement importants.

M. de Sparre avait, vers 1880, présenté une solution très originale qu'il n'a cessé de développer depuis. Mais certains phénomènes expérimentaux de stabilisation observés dans les polygones ont conduit les savants à reprendre l'étude du problème. Avant la guerre, en collaboration avec M. l'Ingénieur en Chef Garnier, nous nous étions proposés à Gâvre, une revision de la question. Tout d'abord cette théorie nouvelle a pris une base plus générale que les théories antérieures; elle admet, en effet, l'existence d'une force perturbatrice de direction quelconque et non plus seulement dirigée dans le plan de résistance (axe figure et tangente). Elle peut introduire ainsi des forces dues à la rotation ou au frottement de l'air. Les équations différentielles s'établissent assez aisément par l'application des principes de la Mécanique rationnelle. Mais une circonstance existe qui différentie complètement le mouvement du projectile de celui du gyroscope, qui est souvent pris comme modèle et type: c'est que, du fait du déplacement du centre de gravité ou du plan de projection, des forces perturbatrices ont pris naissance, et elles ont, sur le mouvement de l'axe de figure, la même importance que celles envisagées antérieurement. C'est un terme de ce genre qui n'avait pas été pris en considération par M. de Sparre.

Les équations nouvelles peuvent être discutées par un développement en série suivant les puissances inverses de la vitesse de rotation supposée très grande du projectile. Le premier terme s'appellera la *précession* et le second la *nutation*.

La précession peut être discutée complètement: la nouvelle théorie introduit un couple stabilisateur qui explique les phénomènes expérimentaux constatés. On peut, de plus, par diverses méthodes, arriver à serrer le problème de plus

*Le Mémoire de M. le Professeur Esclangon est paru au Mém. Art. Franç., T. V, 1926.

près, à distinguer dans le mouvement de l'axe une élongation principale et des spires enroulantes dans certains cas, déroulantes dans d'autres.

Les conditions de stabilité se précisent alors, et on arrive jusqu'aux formules pouvant être utilisées dans la pratique.

L'étude du mouvement du centre de gravité dans l'espace peut alors être entreprise. Le mouvement spiral de ce centre de gravité autour de la trajectoire moyenne est démontré et de nombreux théorèmes fixent les lois de ce mouvement et les divers degrés de stabilité dont un projectile est susceptible par le rapprochement plus ou moins progressif de l'axe et de la tangente.

La *nutation* qui est le terme du second ordre de la série peut alors être étudiée à fond en appliquant des méthodes dont le principe est encore dû à M. de Sparre. Des applications au cas du départ normal du projectile, au cas des perturbations initiales et au cas des défauts dans la forme ou la répartition de la masse du projectile se trouvent maintenant accessibles au calcul.

A Gâvre également, le savant Professeur, M. Esclangon, s'est attaqué au même problème. Introduisant plus particulièrement l'action du frottement de l'air sur l'ogive du projectile, qui donne lieu à une force normale du plan de résistance, M. Esclangon arrive à des équations différentielles qui ne diffèrent pas essentiellement de celles de la théorie précédente. Mais la discussion des équations et leur intégration sont faites par une méthode différente, qui présente le phénomène sous un aspect un peu autre, quoique les deux théories paraissent au fond réductibles à une seule*.

En Angleterre, une très notable contribution à la même question a été apportée par la discussion d'expériences de passage des projectiles à travers des écrans successifs. Un mémoire très savant dû à MM. Fowler, Gallop, Lock, Richmond†, a interprété ces résultats dont une théorie fort savante et intéressante a été établie par les auteurs.

En résumé, la théorie se trouve maintenant avancée à un point qu'elle était loin d'avoir atteint jusqu'ici, et elle conduit à la possibilité de calculs numériques qui élucideront complètement les particularités du mouvement réel du projectile, le phénomène de la dérivation et la question de la stabilité.

Il semble que le moment soit prochain où une synthèse générale de tous les travaux modernes pourra présenter la théorie sous une forme entièrement satisfaisante, et constituer ainsi sur une base sûre et définitive un des plus beaux chapitres de la Mécanique rationnelle.

VIII

20. La Balistique théorique se présente donc à l'heure actuelle, comme une science rationnelle susceptible de donner avec exactitude et facilité tout ce

*On trouve l'exposé sommaire de ces deux méthodes dans le *Cours de Balistique* de M. l'Ingénieur en Chef Sugot. Voir aussi un ensemble de Mémoires sur la même question dans Mém. Art. Franç., T. VI, 1927.

†*Aérodynamique d'un projectile tournant*, par MM. Fowler, Gallop, Lock, Richmond: Mémorial de l'Artillerie Française, 1922. Traduit des Trans. Royal Soc., Série A, 221 and 222.

dont l'artilleur peut avoir besoin pour préparer, régler, transporter son tir, pour construire, utiliser ses appareils et machines de réglage; pour combiner, discuter et améliorer ses méthodes de tir. La précision et la perfection de la théorie sont nettement en avance sur celles de l'expérience. C'est donc de celle-ci que peuvent venir les progrès futurs, avec une précision accrue dans les données et caractéristiques initiales (mesure des vitesses initiales et restantes, des durées de trajet, des angles de chute, etc. . . .). L'outillage des polygones modernes se perfectionne d'ailleurs de jour en jour et la caractéristique principale de ces progrès paraît être l'emploi de plus en plus courant et précis des procédés photographiques (projectiles en mouvement, points d'éclatement des obus, photographies même de jour, emploi du cinéma, etc. . . .).

L'autre donnée que les balisticiens désireraient posséder avec plus de précision, non pour perfectionner leurs méthodes de calcul, mais pour serrer de plus près les faits balistiques dans leur confrontation avec l'expérience, est une connaissance plus approfondie de la résistance de l'air.

Des études expérimentales sont en cours actuellement sur ce sujet.—Dans une théorie physique et mathématique de la résistance de l'air, le balisticien désirerait trouver les lois générales ou du moins leur forme, de sorte que l'expérience n'aurait plus qu'à déterminer quelques coefficients numériques particuliers. Malheureusement, les mathématiciens laissent peu d'espoir qu'une telle solution soit prochaine. On doit renoncer, semble-t-il, à une loi de résistance unique permettant de passer par proportionnalité d'une forme du projectile à une autre forme. Il y aura des groupes de lois pour des projectiles d'égal affinement à l'ogive et au culot. Peut-être des termes secondaires seront-ils suffisants pour tenir compte des variations de forme.

21. Je terminerai cette rapide vue d'ensemble en disant que les progrès de la Balistique Extérieure dans l'avenir ne peuvent qu'être favorisés par un fait important: c'est qu'à l'exemple de leurs plus illustres devanciers, savants et mathématiciens des 17^e et 18^e siècles, nos savants modernes ont tourné leurs regards vers cette science, l'ont cultivée et déjà enrichie. Travaillant pendant la guerre en étroite communauté avec les balisticiens professionnels ils ont, dans tous les pays, apporté d'importantes contributions à la théorie et aux applications; souhaitons qu'une telle collaboration, qui n'a peut-être pas été utile aux seuls artilleurs, mais à eux-mêmes, continue dans l'avenir.

Mais les conditions modernes des progrès d'une science exigent un effort continu et cohérent, une documentation fidèle et complète, une connaissance rapide et largement répandue des travaux théoriques et expérimentaux de tous ceux qui s'intéressent à cette science et en cultivent les différentes branches. A cette nécessité du travail efficace doit donc correspondre un organe de liaison; il existe sous la forme d'une publication trimestrielle: «le Mémorial de l'Artillerie Française» consacré exclusivement aux hautes sciences de l'Artillerie: les deux Balistiques, la théorie des explosifs, la construction des bouches à feu et du matériel, le tir, les probabilités, etc. . . . et toutes les branches des sciences théoriques et expérimentales se rattachant à l'artillerie.

Le Mémorial de l'Artillerie Française ouvre largement ses colonnes à tous les balisticiens et techniciens d'artillerie du monde, et en particulier sera heureux de publier les mémoires et travaux que les membres du congrès International de Mathématiques de Toronto voudront bien lui adresser.

Le Mémorial de l'Artillerie Française dans ses cinq premières années 1922 à 1926 a publié les travaux suivants sur la Balistique Extérieure et sur la Balistique Intérieure et Expérimentale:

I—BALISTIQUE EXTÉRIEURE

Articles généraux de revues

J. Hadamard, Membre de l'Institut: *Rapport sur les travaux examinés et retenus par la Commission de balistique pendant la durée de la guerre*, (T. 1, 1922, 1[er] fascicule, p. 11).

M. Garnier, Ingénieur en Chef de l'Artillerie Navale: *Conférences sur la Balistique Extérieure*, (T. 1, 1922, 1[er] fascicule, p. 109).

Résistance de l'air

D. Riabouchinski, Docteur ès Sciences, fondateur et directeur de l'Institut aérodynamique de Koutchino: *Aperçus sur quelques recherches de balistique*, (T. 2, 1923, 3[e] fascicule, p. 687).

Andreau, Capitaine d'Infanterie Coloniale: *Application de l'aérodynamique à la balistique*, (T. 1, 1922, 3[e] fascicule, p. 484).

L. Lamothe, Ingénieur Principal de l'Artillerie Navale: *Note sur la loi de résistance de l'air aux projectiles en mouvement*, (T. 2, 1923, 2[e] fascicule, p. 347).

L. Dupuis, Ingénieur de l'Artillerie Navale: *Note au sujet des études à entreprendre sur les lois de la résistance de l'air*, (T. 2, 1923, 4[e] fascicule, p. 963).

L. Bairstow, R. H. Fowler et D. R. Hartree: *Distribution des pressions sur la tête d'un projectile se déplaçant à grande vitesse*, (T. 2, 1923, 4[e] fascicule, p. 977).

E. Vessiot, Professeur à la Faculté des Sciences de Paris, Sous-Directeur de l'École Normale Supérieure: *Remarques sur l'application du principe de l'homogénéité à la Balistique Extérieure*, (T. 4, 1925, 1[er] fascicule, p. 3).

M. Roy, Ingénieur au Corps des Mines, Docteur ès-Sciences: *Note sur l'état actuel de nos connaissances théoriques concernant la résistance de l'air*, (T. 4, 1925, 1[er] fascicule, p. 23).

M. Paul, Ingénieur d'Artillerie Navale: *Sur l'extrapolation de $F(v)$ aux grandes vitesses. Essai de classification rationnelle des lois de résistance de l'air au mouvement des projectiles aux vitesses supérieures à celle du son*, (T. 4, 1925, 1[er] fascicule, p. 67).

Aérologie Balistique

P. Wehrlé, Capitaine d'Artillerie: *Sur l'augmentation de vitesse du vent avec l'altitude*, (T. 2, 1923, 4[e] fascicule, p. 821).

P. Vie, Capitaine d'Artillerie Coloniale: *L'aérologie et ses applications en balistique*, (T. 5; 1926, 3[e] fascicule, p. 677).

Acoustique d'Artillerie

M. Roy, Ingénieur au Corps des Mines: *Sur l'onde balistique et le mouvement d'un fluide compressible au voisinage des projectiles à grande vitesse*, (T. 4, 1925, 1[er] fascicule, p. 43).

E. Burlot, Ingénieur Principal des Poudres: *Compte-rendu sommaire des expériences de la Courtine et des essais qui les ont précédés*, (T. 4, 1925, 2 fascicule, p. 479).

E. A. Milne, B.A., Membre du Trinity College, Cambridge: *Les ondes sonores dans l'atmosphère*, (T. 4, 1925, 2[e] fascicule, p. 527).

C. E. Munroe: *Distribution des ondes de l'explosion de l'Alun Chine*, (T. 4, 1925, 2[e] fascicule, p. 545).

P. Charbonnier, Ingénieur Général de l'Artillerie Navale: *Étude cinématique du champ acoustique d'un projectile*, (T. 4, 1925, 3[e] fascicule, p. 565).

E. Esclangon, Directeur de l'Observatoire de Strasbourg, Professeur à la Faculté des Sciences: *L'acoustique des canons et des projectiles*, (T. 4, 1925, 3[e] fascicule, p. 639).

J. Chapelon, Ingénieur en Chef au Corps des Mines, Répétiteur à l'École Polytechnique: *Les ellipses de probabilité dans le repérage par le son*, (T. 4, 1925, 1er fascicule, p. 123).

E. Burlot, Ingénieur Principal des Poudres: *Étude expérimentale des caractéristiques des ondes de choc*, (T. 5, 1926, 3e fascicule, p. 935).

Théories Balistiques

A. de Stefano, Général de Division de l'Armée italienne: *Sur la trajectoire du centre de gravité d'un projectile dans un milieu de densité variable*, (T. 3, 1924, 4e fascicule, p. 1187).

M. Desprez, Capitaine d'Artillerie: *Abaque de balistique extérieure*, (T. 1, 1922, 1er fascicule, p. 225).

A. Perrin, Capitaine: *Sur la résolution mécanique du problème principal de la balistique extérieure*, (T. 1, 1922, 2e fascicule, p. 337).

Risser, Chef d'Escadron d'Artillerie: *Essai de recherches de formules représentatives des trajectoires dans les tirs à grande et faible vitesse*, (T. 1, 1922, 3e fascicule, p. 565).

Bory, Ingénieur de 1ère classe de réserve de l'Artillerie Navale: *Essai sur la balistique de la fusée*, (T. 1, 1922, 3e fascicule, p. 683).

P. Charbonnier, Ingénieur Général de l'Artillerie Navale: *Les théorèmes généraux de la balistique généralisée*, (T. 2, 1923, 2e fascicule, p. 421).

E. Cavalli, Major Général de l'Armée italienne: *Le problème balistique de l'avenir*, (T. 2, 1923, 2e fascicule, p. 459).

K. Popoff, Professeur à l'Université de Sofia: *Sur une propriété de la trajectoire et son application à l'intégration des équations de la balistique extérieure*, (T. 2, 1923, 4e fascicule, p. 993).

G. Fubini, Professeur à l'Université de Turin: *Quelques recherches de Balistique Extérieure*, (T. 4, 1925, 2e fascicule, p. 381).

K. Popoff, Professeur à l'Université de Sofia: *Les méthodes d'intégration de Poincaré et le problème général de la Balistique extérieure. Leçons faites à la Sorbonne pendant l'année scolaire 1924-1925*, (T. 5, 1926, 1er fascicule, p. 3).

Atmosphère (vent, densité de l'air, etc. . . .)

G. Darrieus, Capitaine: *Sur une nouvelle cause de variation de portée due à la température*, (T. 1, 1922, 2e fascicule, p. 241).

P. Langevin, Professeur au Collège de France: *Note sur la loi de résistance de l'air et sur la correction d'élasticité proposées par M. le Capitaine G. Darrieus*, (T. 1, 1922, 2e fascicule, p. 253).

E. Jouguet, Ingénieur en Chef des Mines: *Remarques sur les lois de la résistance des fluides*, (T. 1, 1922, 2e fascicule, p. 267).

M. Garnier, Ingénieur en Chef de l'Artillerie Navale: *Sur les corrections d'élasticité et de densité de l'air*, (T. 1, 1922, 2e fascicule, p. 271).

M. Valiron, Professeur à l'Université de Strasbourg: *Remarques sur le problème de l'ogive de moindre résistance*, (T. 1, 1922, 2e fascicule, p. 283).

M. Garnier, Ingénieur en Chef de l'Artillerie Navale: *Sur les altérations de portée dues aux perturbations atmosphériques*, (T. 1, 1922, 2e fascicule, p. 299).

Étude de l'action du vent dans les hautes altitudes:

- Note I: M. Garnier, Ingénieur en Chef de l'Artillerie Navale: *Action balistique du vent dans les hautes altitudes*, (T. 2, 1923, 4e fascicule, p. 817).
- Note II: P. Wehrlé, Capitaine d'Artillerie: *Sur l'augmentation de vitesse du vent avec l'altitude*, (T. 2, 1923, 4e fascicule, p. 821).
- Note III: M. Garnier, Ingénieur en Chef de l'Artillerie Navale: *Grandeur comparée des corrections de vent*, (T. 2, 1923, 4e fascicule, p. 835).

P. Boutroue, Capitaine d'Artillerie: *Note sur le vent balistique*, (T. IV, 1925, 1er fascicule, p. 73).

Terre (Coubure, Rotation)

P. Boutroue, Capitaine d'Artillerie: *Note sur la correction de portée due à la rotation de la terre*, (T. 2, 1923, 4e fascicule, p. 853).

Projectile (poids, forme, mouvement gyroscopique)

R. H. Fowler, E. G. Gallop, C. N. H. Lock et H. V. Richmond, F.R.S.: *Aérodynamique d'un projectile tournant*, (T. 1, 1922, 2e fascicule, p. 379).

M. Duchêne, Chef de Bataillon du Génie: *Notions sur l'équilibre et la stabilité des projectiles empennés*, (T. 1, 1922, 4e fascicule, p. 910).

Andreau, Capitaine d'Infanterie Coloniale: *Les projectiles empennés et la précision*, (T. 1, 1922, 4e fascicule, p. 929).

Chronographes Balistiques—Instruments de Polygone

M. Roy, Ingénieur au Corps des Mines, Docteur ès Sciences: *Sur une utilisation des rupteurs électro-acoustiques pour la détermination des angles de chute et des vitesses restantes*, (T. 2, 1923, 3e fascicule, p. 653).

J. Loiseau, Ingénieur d'Artillerie Navale (C.R.): *Mesure du temps*, (T. 4, 1925, 2e fascicule, p. 359).

A. W. Ford: *Le chronographe à solénoide d'Aberdeen*, (T. 5, 1926, 4e fascicule, p. 1293).

Photographie et cinématographie Balistique

J. Kampé de Fériet, Maître de conférences à la Faculté des Sciences de Lille, Membre assistant de la Commission de Gâvre: *La mesure de la vitesse des projectiles par enregistrement photographique sur plaque mobile*, (T. 4, 1925, 2e fascicule, p. 289).

M. Paul, Ingénieur Principal de l'Artillerie Navale: *Sur quelques nouveaux appareils de Balistique extérieure expérimentale*, (T. 4, 1925, 2e fascicule, p. 305).

L. Dunoyer: *Détermination chronophotographique des trajectoires*, (T. 1, 1922, 3e fascicule, p. 533).

A. Lamothe, Ingénieur Principal de l'Artillerie Navale: *Cinématographie des projectiles*, (T. 3, 1924, 3e fascicule, p. 919).

C. Cranz et Bames: *La cinématographie à haute fréquence des filets gazeux et son emploi pour l'étude des explosions et autres phénomènes d'allure très rapide*, (T. 3, 1924, 3e fascicule, p. 951).

Établissement des tables de tir

P. Charbonnier, Ingénieur Général de l'Artillerie Navale: *Note sur l'état actuel de la Balistique extérieure appliquée*, (T. 3, 1924, 1er fascicule, p. 25).

Garnier, Ingénieur en Chef de l'Artillerie Navale: *Probabilités du tir fusant de 75 contre avions*, (T. 2, 1923, 2e fascicule, p. 253). *Moyennes balistiques dans les tirs de polygone*, (T. 2, 1923, 2e fascicule, p. 353).

Whittet et D. A. Gurney: *L'angle de relèvement*, (T. 2, 1923, 4e fascicule, p. 1009).

M. Garnier, Ingénieur en Chef de l'Artillerie Navale: *Méthode générale d'établissement des abaques de tir aérien*, (T. 3, 1924, 3e fascicule, p. 895).

II—TIR

Théorie générale des probabilités

E. Lhoste, Chef d'Escadron d'Artillerie Coloniale: *Étude d'un procédé rationnel permettant de vérifier la loi de Gauss*, (T. 4, 1925, 1er fascicule, p. 245).

M. Dumas, Ingénieur d'Artillerie Navale: *Sur une interprétation des conditions de recette*, (T. 4, 1925, 2e fascicule, p. 395).

J. Haag, Professeur à la Faculté des Sciences de Clermont-Ferrand: *Procédé rationnel permettant de vérifier la loi de Gauss (Problème du Commandant Lhoste)*, (T. 4, 1925, 3fe ascicule, p. 1027).

J. Haag, Professeur à la Faculté des Sciences de Clermont-Ferrand: *Sur l'application de la loi de Gauss à la position probable d'un point dans le plan ou dans l'espace*, (T. 2, 1923, 2e fascicule, p. 217).

E. Boutroue, Capitaine d'Artillerie: *La dispersion et sa généralisation à deux ou trois dimensions*, (T. 2, 1923, 2e fascicule, p. 315).

Probabilités du tir

J. Haag, Professeur à la Faculté des Sciences de Clermont-Ferrand: *Vérification expérimentale de la loi de Gauss en artillerie*, (T. V, 1926, 2[e] fascicule, p. 449).

P. Henry, Chef d'Escadron d'Artillerie: *Probabilité du tir* (*Cours professé en 1894 à l'École d'Application de l'Artillerie et du Génie*), (T. V, 1926, 3[e] fascicule, p. 295).

Observation, Réglage, Transport du tir

Garnier, Ingénieur en Chef de l'Artillerie Navale: *Sur les méthodes de réglage et de repérage par coups fusants hauts*, (T. 2, 1923, 4[e] fascicule, p. 707).

Préparation, Réglage, Conduite du tir

H. Béghin, Docteur ès Sciences, maître de conférences à la Faculté des Sciences de Montpellier, et P. Mongraix, Ingénieur Principal de l'Artillerie Navale: *Description et étude théorique d'un compas isogyrostatique*, (T. 2, 1923, 3[e] fascicule, p. 599).

Metin, Ingénieur en Chef de l'Artillerie Navale: *Exposé sommaire des divers procédés de pointage des canons dans le tir à bord*, (T. 2, 1923, 2[e] fascicule, p. 403).

Monfraix, Ingénieur Principal de l'Artillerie Navale: *De l'emploi en artillerie des systèmes gyroscopiques; contribution à leur étude*, (T. 1, 1922, 3[e] fascicule, p. 613).

Tir et conduite du tir

M. Garnier, Ingénieur en Chef de l'Artillerie Navale: *Sur les corrections de vent dans le tir contre aéronefs*, (T. 2, 1923, 1[er] fascicule, p. 109).

M. Garnier, Ingénieur en Chef de l'Artillerie Navale: *Tir de bord contre aéronefs* (*correction de vent*), (T. 3, 1924, 3[e] fascicule, p. 103).

L. Rougeul, Colonel d'Artillerie: *Étude théorique sur le tir contre avions*, (T. V, 1926, 3[e] fascicule, p. 867).

L. Plumejeaud, Lieutenant de Vaisseau: *Le tir contre avion*, (T. V, 1926, 1[er] fascicule, p. 131.

Bombardement par avions

H. Dubois, Lieutenant de Vaisseau: *Considérations théoriques sur le bombardement aérien en dérive contre bâtiment à la mer*, (T. V, 1926, 1[er] fascicule, p. 141).

III—BALISTIQUE INTÉRIEURE

Articles généraux de revues

M. Mussel, Lieutenant-Colonel d'Artillerie breveté: *La méthode des approximations successives de M. Picard et ses applications possibles en balistique*, (T. 1, 1922, 1[er] fascicule, p. 127).

P. Charbonnier, Ingénieur Général de l'Artillerie Navale: *La Balistique intérieure depuis B. de Morogues* (*1737*) *jusqu'à nos jours*, (T. V, 1926, 4[e] fascicule, p. 1293).

Pyrostatique

M. Letang, Ingénieur, ancien élève de l'École Polytechnique: *Mémoire sur la combustion des poudres en vase clos*, (T. 1, 1922, 4[e] fascicule, p. 955).

H. Muraour, Ingénieur en Chef des Poudres: *Notes de Balistique intérieure*, (T. 4, 1925, 2[e] fascicule, p. 455).

Pyrodynamique

M. Desmazières, Chef d'Escadron d'Artillerie: *Premier mémoire sur la Balistique intérieure*, (T. 1, 1922, 4[e] fascicule, p. 19).

P. Charbonnier, Ingénieur Général de l'Artillerie Navale: *La veine gazeuse*, (T. 1, 1922, 4[e] fascicule, p. 10001).

Sir J. B. Henderson, D.Sc. et le Professeur H. R. Hassé, D.Sc.: *Contribution à la théorie thermodynamique des explosions*, (T. 1, 1922, 4[e] fascicule, p. 1079).

P. Langevin, Professeur au Collège de France: *Note sur les effets balistiques des gaz de la poudre dans une tuyère convergente-divergente*, (T. 2, 1923, 1^er^ fascicule, p. 1).

E. Emery, Général de Brigade: *Note sur les relations d'affinité ou de similitude en balistique intérieure*, (T. 2, 1923, 1^er^ fascicule, p. 21).

A. E. H. Love, F.R.S., Professeur de Physique à l'Université d'Oxford, et F. B. Pidduck, M.A. Agrégé du Queen's College, Oxford: *Le problème balistique de Lagrange*, (T. 3, 1924, 2^e^ fascicule, p. 701).

G. Sugot, Ingénieur en Chef de l'Artillerie Navale: *Précis de Balistique intérieure*, (T. 3, 1924, 4^e^ fascicule, p. 979).

F. Desmazières, Chef d'Escadron d'Artillerie: *Examen des principales études théoriques de balistique intérieure publiées à l'étranger de 1913 à 1922*, (T. 3, 1924, 4^e^ fascicule, p. 1009).

M. Voituriez, Capitaine d'Artillerie: *Le problème de la Balistique intérieure et les résistances passives*, (T. 4, 1925, 1^er^ fascicule, p. 131).

G. Leplat, Ingénieur d'Artillerie Navale: *Note sur les calculs de Balistique intérieure au moyen des formules Charbonnier-Sugot et sur les divers abaques employés pour faciliter ces calculs*, (T. 4, 1925, 1^er^ fascicule, p. 151).

G. Sugot, Ingénieur en Chef de l'Artillerie Navale: *Balistique intérieure théorique* (*Leçons professées à l'École d'application d'Artillerie Navale*), (T. 5, 1926, 4^e^ fascicule, p. 1131).

F. Desmazières, Chef d'Escadron d'Artillerie: *Les effets balistiques des poudres.* (*Conférence faite au centre d'études tactiques de Metz*), (T. 5, 1926, 4^e^ fascicule, p. 1261).

C. Rinderknech, Capitaine d'Artillerie: *Essai d'un exposé rationnel des théories de similitude balistique*, (T. 5, 1926, 4^e^ fascicule, p. 1275).

Pyromanomètres, Appareils de mesure

J. de Fosseux, Ingénieur Principal de l'Artillerie Navale: *Expériences sur le tarage des crushers au moyen d'une bombe à piston libre*, (T. 4, 1925, 1^er^ fascicule, p. 197).

H. Muraour, Ingénieur en Chef des Poudres: *Note sur la théorie des crushers*, (T. 4, 1925, 1^er^ fascicule, p. 229).

J. C. Karcher, Physicien: *Une méthode piézo-électrique pour la mesure instantanée des hautes pressions*, (T. 4, 1925, 1^er^ fascicule, p. 257).

D. A. Keys: *L'oscillographe cathodique et son application à la mesure exacte des pressions d'explosion, des variations de potentiel dans les tubes à vide et les magnétos à haute tension*, (T. 4, 1925, 1^er^ fascicule, p. 265).

D. A. Keys, M.A., Ph.D.: *Une méthode piézo-électrique pour la mesure des pressions explosives* (*analyse du professeur J. J. Thomson, O.M., F.R.S.*), (T. 4, 1925, 1^er^ fascicule, p. 277).

M. Lamothe, Ingénieur Principal de l'Artillerie Navale: *Examen des principaux travaux sur la théorie des crushers en cuivre*, (T. 1, 1922, 4^e^ fascicule, p. 827).

M. Malaval, Ingénieur en Chef de l'Artillerie Navale: *Au sujet de la note de M. l'Ingénieur Principal Lamothe: «Examen des principaux travaux sur la théorie des crushers en cuivre»*, (T. 1. 1922, 4^e^ fascicule, p. 905).

Desmazières, Chef d'Escadron d'Artillerie: *Mesure des pressions absolues au moyen de crushers*, (T. 1, 1922, 4^e^ fascicule, p. 777).

IV—POUDRES ET EXPLOSIFS

Propriétés mécaniques des explosifs

H. Muraour, Ingénieur en Chef des Poudres: *Note sur la théorie des explosifs*, (T. 5, 1924, 1^er^ fascicule, p. 329).

Propriétés physiques et chimiques des explosifs

H. Muraour, Ingénieur en Chef des Poudres: *Comparaison entre les températures d'explosion calculées à partir des pressions explosives et les températures d'explosions calculées à partir des chaleurs spécifiques*, (T. 3, 1922, 1^ea^ fascicule, p. 349).

H. Muraour, Ingénieur en Chef des Poudres: *Note sur la chaleur spécifique des gaz à haute température* (*comparaison des expériences de Berthelot et Vieille avec les expériences effectuées au Laboratoire de Nernst*), (T. 3, 1922, 1^er^ fascicule, p. 373).

O. Poppenberg, Professeur: *Sur la décomposition des substances explosives*, (T. 3, 1922, 1er fascicule, p. 383).

H. Muraour, Ingénieur en Chef des Poudres: *Note sur l'influence du refroidissement dans la mesure des pressions explosives*, (T. 3, 1922, 1er fascicule, p. 339).

Matières premières minérals (azote, acides sulfurique et nitrique, ammoniaque, etc. . . .)

R. Schwarz, Docteur: *Le coton et ses succédanés pour la fabrication de la nitrocellulose*, (T. 1, 1922, 1er fascicule, p. 229).

Explosifs

H. Muraour, Ingénieur en Chef des Poudres: *Note sur les différents types d'explosifs utilisés en Allemagne avant et pendant la guerre*, (T. 2, 1923, 4e fascicule, p. 801).

Poudres balistiques

H. Muraour, Ingénieur en Chef des Poudres: *Note sur les différents types de poudre utilisés en Allemagne avant et pendant la guerre*, (T. 2, 1923, 2e fascicule, p. 503).

Fulminate et explosifs d'amorçage

R. Muller, Ingénieur d'Artillerie Navale: *Le fulminate de mercure*, (T. 4, 1925, 4e fascicule, p. 1033).

Emploi des explosifs pour les destructions

H. Muraour, Ingénieur en Chef des Poudres: *La nouvelle réglementation des explosifs de mine en Allemagne*, (T. 2, 1923, 4e fascicule, p. 763).

Stabilité et épreuves des poudres et des explosifs

A. Apard, Ingénieur-chimiste au laboratoire des poudres de la Marine: *Propriétés hygroscopiques des poudres B*, (T. 4, 1925, 2e fascicule, p. 439).

Accidents et sécurité

P. Breteau, Professeur au Val-de-Grace: *Sur le danger d'intoxication par un tir intensif sous abri*, (T. 2, 1923, 4e fascicule, p. 837).

F. Febvret, Ingénieur de réserve de l'Artillerie Navale: *Les dangers d'incendie par inflammation spontanée des chiffons*, (T. 2, 1923, 4e fascicule, p. 847).

F. Taradoire, Ingénieur-chimiste: *Inflammation spontanée des chiffons imbibés de substances diverses*, (T. 4, 1925, 4e fascicule, p. 1067 et T. 5, 1926, 2e fascicule, p. 571).

F. H. Miles: *L'incendie de Old Hickory*, (T. 5, 1926, 1er fascicule, p. 245).

C. G. Storm, Chief of the Explosives Section of the Ammunition Division, Ordnance Office: *Détonation par influence*, (T. 5, 1926, 1er fascicule, p. 251).

Articles généraux de revues

A. Stettbacher: *Les explosifs de 1914 à 1923*, (T. 5, 1926, 2e fascicule, p. 603).

C. L. Reese, Docteur en Philosophie, Docteur ès Sciences, membre de l'Institut Franklin: *Les progrès des explosifs dans ces vingt-cinq dernières années*, (T. 5, 1926, 2e fascicule, p. 655).

Reprinted from the Proceedings of the International Mathematical Congress, Toronto, Canada, 1924.

THE CHOICE OF INDEPENDENT VARIABLE IN THE CALCULATION OF TRAJECTORIES BY SMALL ARCS

By Mr. F. R. W. Hunt,
Artillery College, Woolwich, England.

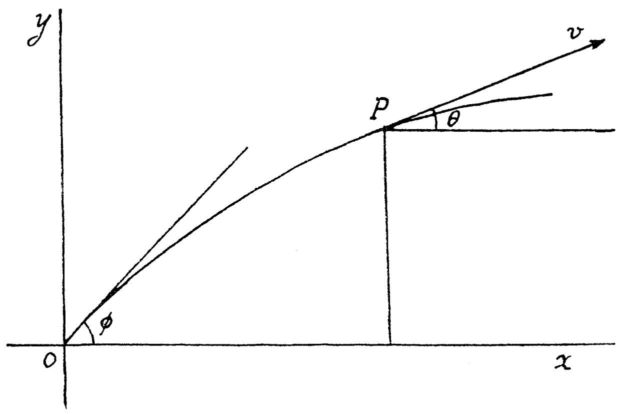
Let us consider briefly the problem of calculating trajectories by small arcs. The projectile is considered as a heavy particle moving in a resisting medium. The retardation due to the air's resistance is $F(v)/Cf(y)$ where $F(v)$ and $f(y)$ are tabulated functions of velocity and height respectively of the particle, and C, the ballistic coefficient,* is a constant for a given projectile.

Then the differential equations of motion are

$$\frac{d(v\cos\theta)}{dt}=\frac{d^2x}{dt^2}=-\frac{F(v)}{Cf(y)}\cos\theta$$

$$\frac{d(v\sin\theta)}{dt}=\frac{d^2y}{dt^2}=-\frac{F(v)}{Cf(y)}\sin\theta-g$$

where x and y are the horizontal and vertical distances of the projectile from the origin at time t and θ is the inclination of the path at that instant.



These equations have to be solved by a step-by-step or small-arc process, the data being C, the ballistic coefficient, and the velocity and inclination at the origin. The solution has to yield values of x, y, and t for a series of points on the trajectory (and sometimes the values of v and θ at the second intersection of the trajectory with the horizon).

*Some writers call the inverse of this quantity the ballistic coefficient.

The first step towards the solution is the choice of the independent variable. This will probably have to be made from the following:

$$x,\ y,\ t,\ s \text{ (length of arc)},\ \theta,\ \tan\theta,\ v,\ v\cos\theta,\ v\sin\theta.$$

Of these, t, θ and $\tan\theta$ have been extensively used; x has been used to some extent. We shall consider the merits of these elements from two points of view:

(1) the calculation of the trajectory;

(2) the application of results.

The calculation of the trajectory.

Two methods of calculation suggest themselves:

(*a*) a formal central difference method;

(*b*) approximate quadrature not based on central differences.

Both these methods are tentative in that certain quantities have first to be estimated, the estimates being subsequently checked or verified. It is therefore evident that an independent variable which has a maximum or minimum on the trajectory is undesirable. Since a small change of the independent variable, when approaching the stationary value, will involve large alterations in the values of the elements, it would clearly be necessary to take extremely small intervals; as the process progresses through the stationary value, moreover, the computing becomes unnecessarily difficult.

The elements y and v both have stationary values; we must therefore remove these from our list of possible independent variables.

The variable $v\sin\theta$ as independent variable leads to rather more complicated equations than the others and has no compensating advantage over them; we shall therefore omit it.

Of the remaining variables t is the most suitable independent variable for the process of central differences; the values of x and y are obtained and the differences checked at once by direct application of the equations given above.

In the process of approximate quadrature the first step is essentially the solution of a differential equation of the first order and degree containing the term $F(v)/Cf(y)$. In every case to be considered an estimate has to be made of the mean value of y in the arc. Since $f(y)$, which is virtually the specific volume of the air at height y, is a function which changes slowly with y, the mean value of y need not be estimated with great accuracy; in fact, quite a bad estimate leads to an inconsiderable error in the solution of the equation. In addition to $\bar{y}$ other estimates may have to be made and we proceed to consider these in each case.

Dealing first with x as independent variable, the differential equation to be solved is

$$d(v\cos\theta) = -\frac{F(v)}{vCf(y)}\,dx.$$

Here the mean value of v has to be estimated in addition to $\bar{y}$ and the estimate cannot be checked until the equation

$$d\theta = -\frac{g}{v^2}\,dx$$

has also been solved with the same estimate of $\bar{v}$.

Secondly, with t as independent variable the differential equation requiring fewest estimates is

$$\frac{d(v\cos\theta)}{v\cos\theta} = -\frac{F(v)}{vCf(y)}\,dt.$$

Here $\bar{v}$ has to be estimated; it cannot be checked until the equation

$$\frac{d\theta}{\cos\theta} = -\frac{g}{v}\,dt$$

has also been solved. These equations lead to logarithmic solutions which, demanding frequent reference to tables, are deprecated by computers. If these logarithmic solutions are to be avoided an additional estimate, *e.g.*, $\bar{\theta}$, has to be made, the equations taking some such form as,

$$d(v\cos\theta) = -\frac{F(v)}{Cf(y)}\cos\theta\,dt,$$

or

$$\frac{d(v\cos\theta)}{v^2\cos^2\theta} = -\frac{F(v)}{v^2Cf(y)}\sec\theta\,dt,$$

and

$$\sec^2\theta d\theta = -\frac{g}{v\cos\theta}\,dt.$$

Dealing next with θ (or $\tan\theta$) the differential equation takes the form

$$d(v\cos\theta) = \frac{vF(v)}{Cf(y)g}\,d\theta$$

or

$$\frac{d(v\cos\theta)}{v^2\cos^2\theta} = \frac{F(v)}{vCf(y)g}\,d\,(\tan\theta).$$

Again $\bar{v}$ has to be estimated, but, unlike the cases just considered, the estimate can be checked without recourse to another equation.

Considering next the horizontal component of the velocity, $v\cos\theta$, as independent variable the differential equation is

$$d\theta = \frac{Cf(y)g}{vF(v)}\,d(v\cos\theta)$$

or, if it is desirable to avoid circular measure,

$$d(\tan\theta)=\frac{vCf(y)g}{F(v)}\,\frac{d(v\cos\theta)}{v^2\cos^2\theta}.$$

Here again $\bar{v}$ has to be estimated and, as with θ and $\tan\theta$, the estimate can be checked without solving another equation.

Unlike $\bar{y}$ the estimate of $\bar{v}$ must be made with considerable accuracy in the cases already considered, since none of the functions of v involved in the quadratures can be said to change slowly with v; in fact, some of them change very rapidly in certain regions of velocity.

By choosing s as independent variable this difficult estimate can be avoided. We may put the differential equation in the form

$$ds=-\frac{vdv}{F(v)}\,Cf(y)$$

and use may be made of one of Siacci's primary ballistic functions, namely,

$$D(v)=\int_k^v\frac{vdv}{F(v)}$$

of which tables exist.

The solution takes the form

$$D(v_1)=D(v_0)-\frac{\delta s}{Cf(\bar{y})}$$

for the interval δs, which is exact but for the easily estimated $\bar{y}$.

This method involves no additional reference to tables—we merely use the table of $D(v)$ instead of $F(v)$—and definitely avoids all difficult estimates. We therefore conclude that s has some appreciable advantage over the other variables when considered solely from this point of view.

Altogether distinct from these considerations is the question of the number of intervals required for a given piece of work. The size of interval is determined by the degree of accuracy required in the resulting values of x, y and possibly t; and, *caeteris paribus*, the larger the interval, the more economical the work.

An exhaustive enquiry into this question would here be out of place; we propose to consider one method of determining which variable leads to the largest intervals for given degree of accuracy.

Let the independent variable be denoted by q. Then in a quadrature of the form

$$\delta x=\tfrac{1}{2}(x_1'+x_0')\,\delta q$$

wherein

$$x'=dx/dq,$$

there is an error

$$\epsilon(\delta x)^3 = \frac{(\delta x)^3}{12} \frac{d^3x}{dq^3}\left(\frac{dx}{dq}\right)^{-3} + O(\delta x)^4,$$

the values of the differential coefficients being taken at the beginning of the arc. The independent variable yielding the largest intervals, and hence the most economical, is that which yields the smallest value of ϵ.

To simplify the work we assume that the true values of x_0' and x_1' have been obtained and that $\bar{y}$ has been estimated accurately for the arc.

Writing

$$F = \frac{F(v)}{Cgf(y)}, \quad n = \frac{v}{F}\frac{dF}{dv},$$

we obtain for the quantity $12v^4 \cos^2\theta|\epsilon|/g^{-2}$ the following expressions:

with t, $nF^2 + (n-1)F \sin\theta$,

" θ, $2(n+2)F^2 + 2(n+5)F\sin\theta + 2(1+2\sin^2\theta)$,

" $\tan\theta$, $2(n+2)F^2 + 2(n-1)F\sin\theta$,

" s, $2F\sin\theta + 4\sin^2\theta - 1$,

" $v\cos\theta$, $\{n-1\}\ \{nF^2 + (3n-1)F\sin\theta + 2n\sin^2\theta - 1\}$.

(In deducing the last it was assumed that the change in n in the interval could be neglected; this is generally a reasonable assumption.)

From these results we conclude that, from the origin to some point beyond the vertex of the trajectory, $\tan\theta$ is better than θ, and t is better than both. There appears to be little to choose between t and s when C is large or when the velocity is not great; for moderate and small C and high velocity s is better than t. Both t and s appear to be better than $v\cos\theta$. For the remaining portion of the trajectory no definite conclusion can be drawn from these results. As, however, the arcs of smallest size occur at the beginning of the trajectory, the consideration of the last part of the trajectory is not of great importance.

Similar analysis on the solution for y yields the result that x may or may not have an advantage over s in the first few arcs; for the rest of the trajectory s is better than x.

This argument leads to the conclusion that s and t, as independent variables, generally require fewer steps than the others for a given piece of work.

There is one further consideration in connection with the calculation of trajectories. Values of x, y and t are required finally; it may therefore be desirable to have values of one of these elements in even intervals. Since y is definitely ruled out, either x or t would have to be chosen as independent variable for this purpose.

The application of results.

We shall consider the more important applications of the calculated trajectory, namely:

(a) variations,

(b) the fuze-setting of time fuzes.

Variations.

The variations of a trajectory for given changes of conditions are always required at constant height. In anti-aircraft gunnery height is generally accepted as an independent variable, while atmospheric variations generally are always considered in terms of height. It is therefore very unfortunate that y is unsuitable as independent variable in the calculation of trajectories.

It is clearly most economical to calculate the variations for the same intervals as those in which the trajectory is calculated, since the results of the trajectory calculation are immediately available; it is therefore simplest to use the same independent variable for the variations as is used in the trajectory calculations.

With x or t only two variations have to be calculated, namely, δy and δt or δy and δx respectively. With any other of the variables three variations have to be calculated, namely, δx, δy and δt. By choosing x or t a direct saving in labour would result.

Time fuzes.

Finally we come to the calculation of fuze-setting for time fuzes. The setting required to give a burst at a given point is either a function of time alone or else a function of height and time. Trajectories calculated in even intervals of time therefore have the necessary information immediately available for the calculation of the fuze-setting, and a knowledge of the variation of time of flight at constant height makes the calculation of fuze variations a simple matter.

ON THE CONES OF STEADY COMPRESSION FOR A FLYING BULLET

By Sir Joseph Larmor,

Lucasian Professor of Mathematics, University of Cambridge, Cambridge, England

When a bullet or shot is flying through the atmosphere, at a speed greater than will allow the air to get out of its way by waves carrying off the compression, the natural expectation would perhaps be that it can progress only by smashing its way through. Long ago Th. Young propounded the correlative idea that when a hammer-blow communicates to the surface of a hard solid a speed greater than that of any possible relieving waves travelling into the interior, the structure of the superficial parts must be shattered.

But a striking discovery was reached photographically more than a quarter of a century ago by the experimental acumen of C. V. Boys, that, in fact, no such smashing through is involved. The bullet flies through the air perfectly smoothly, therefore without dissipation of energy, except that a rather narrow wake of fine-grained turbulence is formed in its rear, which contracts with smaller terminal cross-section, being, so to say, a sort of misfit in the closing up of the disturbance. This very remarkable type of permanent motion in elastic fluids, as striking in its way (if it be accepted) as Helmholtz's discovery of the permanence of vortex rings in perfect fluid, does not seem to have attracted the full attention that it deserves: yet, if the matter has not been misconceived, the general lines of an explanation are not remote.

The original photographs by Boys were the subject of correspondence between him and Stokes: (see the two volumes of "Scientific Correspondence of Sir George Stokes"). More recent photographs by Professor Dayton C. Miller and J. C. Quayle are in Jour. Franklin Inst. 1922; some of them are reproduced by L. Thompson in Proc. U.S. Nat. Acad. Sci., June 1924, p. 280, and in fact suggested the present note. The concentration of the resistance to a resultant suction located on the rear of the bullet that they indicate, must have some influence in stabilizing its lengthways direction even when there is no spin.

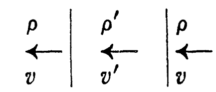
The most significant feature in the photographs is a fine thread-like curve, surrounding the head of the bullet a little way in front and ranging back at the sides so as to take on a hyperbolic form. Another curve precisely parallel strikes off on both sides from the rear of the bullet, or rather converges to a point some distance behind it so that the whole of the bullet is included between, while exhibited close to it is the local misfit above referred to. Such curves are an optical indication of abrupt change of density of the air. It is now suggested that the flat cone of changing air between the two blunt conical boundaries, thus shown in section as curves, is maintained at a definite steady compression determined by the velocity of the bullet, air being compressed into it in front and expanding out of it behind as it travels onwards as a form across the stationary atmosphere. Thus in each position of the shadow-form, so to say, there is within it a local condensation of the air: if this contraction of volume is just sufficient to make room for the bullet, there will be no need for the latter to push air aside in mass as it travels onward: so that this shielding conical boundary of a condensed region, which advances steadily in front of it, will keep the whole motion steady.

In contrast, for velocities of the bullet less than sound, it will appear that the same relations hold but give rise to local expansion instead of contraction, so that no room is provided for the bullet and the effects cannot be merely local. Resistance would then be expected to be produced, mainly, as in the case of liquids, by throwing off vortex rings, in addition to sound-waves, both of which travel away with energy that must have been abstracted from that of the projectile. The features of the motion seem also to be less definite and regular. [See however postscript].

To test this order of ideas one begins naturally with the case of a flat slab-form of augmented density travelling across a still atmosphere, being sustained at the steady excess of density by condensation in front and expansion behind as it flits across the air. The conditions are simplified by impressing a uniform velocity on the whole system, which by relativity has no dynamical effect, so as to bring the form of condensation to rest. Then the air moves up to a stationary region of augmented density with velocity v, through it with v', and away on the other side with v. If the density within this slab is ρ' and outside it is ρ, we have

Steady

$$\rho v = \rho' v' = M,$$



when M is the flux across per unit area.

Moreover a momentum Mv delivered at the front boundary per unit time diminishes to Mv' on crossing it: the rest of it must provide a surface force producing change of pressure given by

$$p' - p = Mv - Mv' = \rho v^2 - \rho' v'^2.$$

This involves the change of density given by the characteristic equation of air, which is with sufficient approximation

$$\delta p = c^2 \delta\rho; \text{ or } p' - p = c^2\left(\frac{M}{v'} - \frac{M}{v}\right),$$

where c is the velocity of sound, assumed the same throughout, as the process can be isothermal.

On equating these values of $p' - p$, a simple relation emerges in the form

$$vv' = c^2.$$

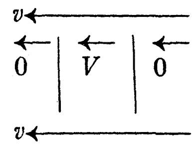
It is to be noted that there is no trouble arising from failure of conservation of energy, for what is imparted to the slab in front is abstracted in the rear.

Impose now a velocity opposite to v on the whole system. The air outside is now at rest in front and rear, while the slab-form progresses steadily through it with this new velocity v, but the air at each instant within the slab also moves forward in mass with a different velocity $(V=v-v')$ in the same direction as v. The relation which must be satisfied to make this motion possible is

$$vv'=c^2,$$

which now becomes

$$v(v-V)=c^2$$



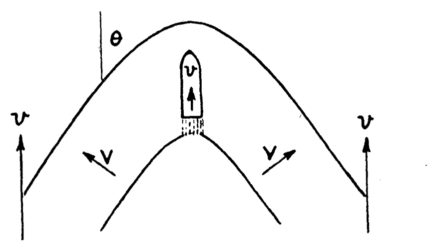
If v, the velocity with which the slab-form advances through the stagnant air, were less than c, the velocity of sound, V would be negative, that is, the air in the progressing slab would then move backward: but as above remarked, the density of the air within the slab is now diminished, which is the reason why the solution then fails for the case of a bullet. For the steady density of the air within the slab is given by

$$\rho'=\rho\frac{v}{v'}=\rho\frac{v^2}{c^2}.$$

If v were slightly different from the velocity of sound c, the change of density and the velocity V of the air would both be small: exact equality would obliterate the slab altogether, which is not really paradoxical.

The argument holds also when the compression-form advances obliquely to itself, provided the transverse component of its velocity $v \sin \theta$ takes the place of v. The air inside would now have a tangential velocity relative to the moving slab, but no such velocity relative to the atmosphere outside, so that there will be no slip.

The same argument is applicable to each part, practically flat, of the bent conical surface bounding the cone of condensed air that accompanies the flying bullet. But now each element of the layer of compression moves outward from



the axis with the component of this large transverse velocity V during the time the compression takes to pass over it. The permanent shift thus produced involves a rarefaction that could not proceed to all distances. Hence V must tend to vanish, and the compression in the layer to disappear, at a distance from the axis. Indeed it is remarkable that the dividing line can be traced so far in the photographs. Anyhow it would be this necessity that fixes the asymptotic value α of the slope of the cone to satisfy $v \sin \alpha = c$ so that the distant parts of the cone travel transversely with the same velocity as sound.

The experiments, if explained in this way, by sharp transition to a region in compression, point to a definite steady state of movement of the air while it is within the slab-form. That state would be determined at each instant, along with the changes of density which make it possible, by the velocity prescribed along its moving conical boundaries, which is V transverse where, at slope θ, as above,

$$V = v \sin \theta - \frac{c^2}{v \sin \theta}\text{; whence } v \sin \theta \, [v \sin \theta - V] = c^2.$$

Also at a distance from the axis of the cone θ is to tend to the asymptotic value which makes V vanish, and further the total local shrinkage by compression of the air must, for steadiness, be equal to the volume of the bullet. The problem thus furnished—one of a new type in hydrodynamic analysis—is whether these data determine the form of the boundary blunt cone, and, for example, whether that form is sensibly independent of the form of the bullet, depending only on its velocity which prescribes the asymptotic angle. The answer to the former question can hardly fail to be in the affirmative, a surface of abrupt transition can exist and is definite: as regards the latter, experiment can probably provide still further crucial indications. The hydrodynamic problem now confronting us as regards the motion of the air in a condensed layer can be elucidated by an analogy derived from electrostatics. The velocity potential of the air within the moving form is the analogue of the electric potential of a charged sheet, as it is constant along the boundary cone of the layer. It could, of course, be determined whatever the form of this sheet subject to the limiting slope α: but the density of charge is also specified everywhere in terms of the slope θ, being proportional to V, and this restricts the sheet for given α to one standard form. When the steady potential is determined, the compression at each point follows. It would appear, then, if and so far as these ideas are confirmed, that the form of the boundary cone for given velocity is a definite universal one, determined intrinsically without reference to the form of the bullet, except that its volume would fix the thickness of the condensed layer. Thus the cones belonging to two fragments cross each other instead of fusing together. In these directions and others, experiment can still provide further tests and indications for theory.

It is to be noted that this compression and adjusting motion in the slab would be sensible only where the conical boundary is curved: thus as the photographs show it could, at high speeds, be localized near the nose of the bullet and, presumably, along it.

If the interpretation here sketched has substance, it would seem that it is largely in the rear end of the bullet or shell that further ballistic improvement can come. The smooth motion in front, at speeds greater than that of sound, can involve no dissipation of energy, and does not seem to interact much, at any rate overtly, with the communicated local turbulence set up from the rear end. But it is easy to overlook essential features; experiment alone can control, and the existing results are not now accessible to the writer. To what degree does the rumbling sound heard sideways, other than the sharp ping whose origin seems clear, proceed solely from the wake? To what extent is the length of the bullet really involved in the form of the cones? Where would the rearward conical boundary strike off if the bullet were pointed at both ends?

The problem is formulated here as that of the motion of a sheet of air between two boundaries: but the boundaries could hardly act quite independently near the apex, as suggested, except when their distance apart is large compared with their radii of curvature and the breadth of the bullet, and experiment alone can test the degree of discrepancy.

Postscript, Sept. 1, 1926:—It appears that this interesting theory of a uniform convected slab of excess density between the cones is not wide enough for the facts, as one is reminded by a Woolwich Ordnance report [RD 63] just published. Both the Boys' cones are shown by the photographs to be advancing compressions; thus a slab will have to be compressed in front and rarefied in the rear. It can be readily proved that this cannot be unless that difference is somehow actively maintained. A permanent set of wavelets of compression may, however, advance from the rapidly intruding front of the bullet, and another set, also probably compression (as infra) from the rapidly receding rear, especially as both travel far above the speed of sound: the Boys' cones would be Huygenian composite fronts, envelopes of the wavelets, bounding the progress of their fields. Indeed, analogy to the bow and stern wave-ridges of a ship lies at hand, which the photographs powerfully confirm: only the speed of the surface waves on the wake, which feed them, there depends on their lengths. If the bow and stern waves are both cones of elevated water the analogy must be nearly complete. But the very remarkable adjustments which produce so sharply defined cones and apparently smooth motion, except in the narrow wake of the bullet, remain for further explanation. There must be frictional retardation generated along the bullet, the analogue of ship friction: it would seem that the friction may be eased in each case by a thin sheet of vortex rollers which pass into the wake, so that instead of direct rubbing it is mainly transferred turbulence,—in the case of the bullet there is no screw to intensify the wake. These rollers running into the irregularities on the surface of the bullet might send off the observed subsidiary cones of compression, and by running off at the end might originate the tail cone. It may readily be verified that the compression at a Boys' front of the order of one-tenth will maintain a kinetic pressure of one atmosphere on the front if it is advancing at a speed of one km/sec. But this need not be a retarding pressure on the bullet—that would seem to come, as supra, mainly from the suction at the wake. For at a surface surrounding the system the pressure in the still air is presumably of normal amount except where it crosses the wake; so if the wake were absent it would have no resultant. The excess pressure in a compression travelling free is in fact compensated internally by the inertia of the motion. Moreover, if the pressure, as usually estimated, near the apex, operated directly on the bullet without compensation, it would appear that the retardation might be excessive. An unrifled bullet retarded from the nose would be violently unstable. All the features on the photographs point to steady motion except in the wake. This amounts to convection of an excess density,—as has been seen, not uniform; where the pressure is, for example, 5 atmospheres the velocity of the mass of air as a whole would be of the order of 4/5 of that of the bullet. But any such velocity is negatived for most of the region lying between the cones by the spherical wavelet struck off from the wire-tip visible in it, so that the compressions seem to be confined to the sides near the Boys' cones.

Various other details that appear on the Woolwich photographs also strongly favour the ship-waves analogy. There are these isolated circular wavelets with centre at the point where the bullet touched the rearward one of the wires that made the spark circuit; this shows also that a slight cause can send out a sensible wavelet, and thus confirms the striking off of the secondary cones from indentations on the bullet. In the shadow pictures the region within the rear cone is filled up with circular wavelets, diverging more or less equidistantly from successive positions of the front end of the wake as the bullet travels along leaving them behind; they are tangential to the cone. The same feature occurs in instantaneous patterns of ship-waves. If these wavelets are truly spherical and so circular in section, it is a confirmation that the air, so close to the bullet, through which they travel, is itself quiescent and of density not far from normal; where they cross a dense band of air they would be deformed. Behind a tapering rear, but in front of the wake, conelets break away, of smaller angles; they arise from a more permanent set of sources, on the air stream converging towards the axis from the sides of the bullet and finally breaking into a now narrower wake. It is this breaking place that is the travelling source producing the wavelets forming the main rear-cone; their forms as shadowed show that it is an intermittent source, perhaps due to its origin from a sheet of collapsing vortices. When the bullet lies oblique to its path the wave should start off with a transverse velocity, as experiment indicates.

As to the rear wavelets being also of compression, the theory of collapsing bubbles, as applied by Sir Charles Parsons to the pitting of turbine screws by cavitation, may be recalled; the radial collapse of the air into the vacuous cavity left by the bullet can produce sudden great pressure where the walls come together, enough to start a wavelet of pressure; on the other hand, an expansion can usually be produced only gradually, and does not relieve itself as a wave.

The radii of the isolated spherical wavelets knocked off from where the nose of the bullet struck the wire seem to show that the wave, doubtless very intense near its source, travels rather faster than the bullet; yet it is strange to note the dark radial lines, apparently from flying fragments of the wire, going ahead far beyond the front cone—unless indeed they may be shadows of the fragments in the light of the spark.

ON THE EQUILIBRIUM OF GASES IN THE REACTION OF EXPLOSIVES

By Professor N. Yamaga,
Imperial University of Tokyo, Tokyo, Japan.

I

Introduction

The study of the reaction in explosives such as the combustion of smokeless powders or the detonation of high explosives, is necessarily connected with the examination of the properties of the gases resulting from the explosions, and their thermodynamical equilibrium at the moment of explosion. Otherwise the pressure of explosions, the composition and volume of gaseous products and the force of explosives, etc., would be unknown. These investigations are not only important in regard to blasting powder, but they are especially important in regard to propellant explosives.

The chief development in the theory of explosives occurred towards the end of last century, two of the principal investigators being Berthelot* and Sarrau†. The theories are chiefly based on the application of the first law of thermodynamics to the reactions of explosives, but so far practically no use has been made of the second law, especially in problems dealing with equilibrium and stability. The combustion of propellant explosives is simply the reaction of the gaseous system at high temperature and high pressure, and consequently must be discussed under the laws of equilibrium.

In regard to the time in which the explosive action takes place, it is of course very short and the question of the velocity of reaction arises whether equilibrium is reached within this short interval or not. Boudouard‡, and Haber and Richardt§ have investigated the velocity with which the equilibrium between CO_2, CO, H_2 and H_2O is established at high temperatures. They found the velocity of reaction was fairly rapid at about 1000° C., and therefore at temperatures from 2000° to 4000° which are reached in the reaction of explosives, equilibrium is undoubtedly established instantaneously. Although the velocity of cooling of a gas in a closed vessel or gun is quite rapid, the transition to the

*M. Berthelot: *Sur la force des matières explosives*, 1883.

†E. Sarrau: *Théorie des explosifs*, 1895.

‡O. Boudouard: Bull. Soc. Chim., Paris (3), t. 25, p. 484, 1901.

§F. Haber und F. Richardt: Z. f. anorg. Chem., Bd. 38, S. 5, 1903, also *Thermodynamics of technical gas reactions* (English Translation).

state of equilibrium will proceed rapidly before the temperature is reduced and the region of false equilibrium is reached, *i.e.*, the equilibrium is "frozen".

The composition of gaseous products of explosives taken from a closed vessel after cooling, as determined by gas analysis, is not the same as at the instant of explosion. It is a well known fact that the composition of gases resulting from an explosion varies according to the density of loading, and it is now considered that this variation is caused by the pressure of the explosion*, *i.e.*, density of loading. But I consider that the variation is not caused by the pressure, but by the speed of cooling due to the density of the gases, *i.e.*, the influence of the temperature in cooling on the variation of gaseous equilibrium, which I shall discuss in Section VI.

All previous calculations which have been used in the theory of explosives have been obtained from the results of measuring the gases after cooling. About 1906, Nernst gave a theory explaining the gas reactions at high temperature and at the same time several experiments on the specific heat of gases at fairly high temperatures were made. This enabled me to calculate the equilibrium of gas reactions at temperatures up to 3000° or 4000° C. The relations between the characteristic equations of states, and the specific heat of explosion gases under such high pressure and temperature†, are still unknown, but it is possible to deduce from the theory of corresponding states a relation, such as the characteristic equation, as it is not very complicated owing to the corresponding temperature being quite high near the temperature of explosion.

As the chemical elements in the propellant explosives such as cordite, nitrocellulose powder and ballistite, are chiefly carbon, hydrogen, oxygen and nitrogen, the products of combustion are carbon dioxide, carbon monoxide, hydrogen, water and nitrogen, and the chief reaction is what we call "water-gas reaction".

$$CO+H_2O \rightleftharpoons CO_2+H_2$$

Beside this, three different reactions for the formation of methane may be considered:

$$CO+3H_2 \rightleftharpoons CH_4+H_2O,$$
$$CO_2+4H_2 \rightleftharpoons CH_4+2H_2O,$$
$$2CO+2H_2 \rightleftharpoons CH_4+CO_2.$$

As these reactions decrease the volume, the amount of methane produced tends to increase with the higher pressure. It will be shown in Section VI that the formation of methane occurs only in the cooling stage of the gases and not at the instant of explosion.

The study of the reaction at the moment of explosion is of great importance not only for the reasons just stated but also in regard to the solution of the problems connected with the working capacity of the propellant, the cooling and oxidation of gases, muzzle flame and the detonation of high explosives.

*Brunswig: *Explosives*, English Translation, p. 165.

†N. Yamaga: *On the specific heat of gases resulting from explosives*, Jour. of Soc. of Ordnance and Explosives, vol. 16, 1.

II

Some Thermodynamic Relations

(1) *Thermodynamic relations of one component systems.*

The thermodynamical relations of homogeneous one component systems (internal energy, entropy, free energy and thermodynamic potential, etc.) can be divided into two classes according as temperature and volume or temperature and pressure are taken as independent variables.

Let: N = number of mols. in the system.
U = internal energy of the system.
S = entropy of the system.
ϵ = internal energy of the system under standard state.
y = entropy of the system under standard state.
C_v = molecular (specific) heat at constant volume.
C_p = molecular (specific) heat at constant pressure.
Ψ = free energy of the system.
Φ = thermodynamic potential of the system.
P = pressure of the system.
V = volume of the system.
T = absolute temperature of the system.

Case 1: On taking T and V as the independent variables we have the three well known thermodynamical equations

(2, 1) $$U=\epsilon+N\int C_v dT+T^2\int\frac{\partial}{\partial T}\left(\frac{P}{T}\right)dV,$$

(2, 2) $$S=y+N\int\frac{C_v}{T}dT+\int\frac{\partial P}{\partial T}\,dV,$$

(2, 3) $$\Psi=U-ST=\epsilon-yT-TN\int\frac{\Gamma(T)}{T^2}\,dT-\int PdV.$$

Here $\Gamma(T)=\int C_v dT$ and in what follows $\Gamma'(T)=\int C_p dT$.

We see that

$$\frac{d}{dT}\left[\frac{1}{T}\int C_v dT\right]=-\frac{\int C_v dT}{T^2}+\frac{C_v}{T},$$

therefore

$$\int C_v dT-T\int\frac{C_v}{T}dT=-T\int\frac{\Gamma(T)dT}{T^2}.$$

Case 2: Taking T and P as the independent variables we have

(2, 4) $$U=\epsilon+N\int C_p dT-\int\left(P\frac{\partial V}{\partial P}+T\frac{\partial V}{\partial T}\right)dP-\int P\frac{\partial V}{\partial T}dT,$$

(2, 5) $$S=y+N\int\frac{C_p}{T}dT-\int\left(\frac{\partial V}{\partial T}\right)_p dP,$$

(2, 6) $$\Phi=U-ST+PV=\epsilon-yT-TN\int\frac{\Gamma'(T)}{T^2}dT-\int P\frac{\partial V}{dP}dP-\int P\frac{\partial V}{\partial T}dT+PV,$$

The homogeneous system of two or more components is generally discussed as an ideal gas mixture, but actual gases never fulfil this condition exactly: the deviations, however, are small, especially at high pressures. It is thus necessary to give a brief general discussion of actual gas mixtures or at least of gases to which van der Waals' equation of state may be applicable.

(2) *Thermodynamic relations of an ideal gas mixture.*

Assume at first that the mixture at every instant can be regarded as an ideal gas mixture. Let:

$n_1, n_2, \ldots, n_i,$	be the number of mols. of the ideal gases $G_1, G_2, \ldots, G_i$;
$u_1, u_2, \ldots, u_i,$	the molecular internal energies of the gases in the unmixed state;
$v_1, v_2, \ldots, v_i,$	the molecular volumes of $G_1, G_2, \ldots, G_i$, in the unmixed state;
$\psi_1, \psi_2, \ldots, \psi_i,$	their molecular free energies in the unmixed state;
$C_{v_1}, C_{v_2}, \ldots, C_{v_i},$	their molecular heats at constant volume (unmixed);
$C_{p_1}, C_{p_2}, \ldots, C_{p_i},$	their molecular heats at constant pressure (unmixed).

The molecular free energy ψ_1 of gas G_1 is given by (2, 3),

$$\psi_1=u_1-Ts_1=\epsilon_1-y_1T-T\int\frac{\Gamma_1(T)}{T^2}\,dT-\int Pdv_1.$$

If the total volume of an ideal gas mixture is V,

$$v_1=\frac{V}{n_1}.$$

The free energy of n_1 mols. of gas G_1 by itself (see Equation (2, 3)) is

(2, 7) $$n_1\psi_1=n_1\epsilon_1-n_1y_1T-n_1T\int\frac{\Gamma_1(T)}{T^2}dT-n_1RT\ln v_1$$

and Ψ the free energy of the system, from the definition of an ideal gas mixture, is

$$\Psi=n_1\psi_1+n_2\psi_2+\ldots+n_i\psi_i$$
$$=\Sigma n_1\left[\epsilon_1-y_1T-T\int\frac{\Gamma_1(T)}{T^2}dT-RT\ln\frac{V}{n_1}\right].$$

If c_1, c_2, . . . be the "numerical concentrations" of the gases respectively and p_1, p_2, . . . be their partial pressures then

$$c_1 = \frac{n_1}{\Sigma n_1},$$

$$p_1 = c_1 P, \tag{2, 8}$$

and if ξ_1, ξ_2, . . . be the volumetric molecular concentrations, then

$$\xi_1 = \frac{n_1}{V} = \frac{n_1}{\Sigma(n_1 v_1)}, \tag{2, 9}$$

$$\Psi = T\Sigma n_1 \left[\frac{\epsilon_1}{T} - y_1 - \int \frac{\Gamma_1(T)}{T^2} dT - R \ln \frac{RT}{P} + R \ln c_1 \right] \tag{2, 10}$$

$$= T\Sigma n_1 \left[\frac{\epsilon_1}{T} - y_1 - \int \frac{\Gamma_1(T)}{T^2} dT + R \ln \xi_1 \right].$$

From

$$\Phi = \Psi + PV = \Psi + RT\Sigma n_1$$

we can obtain the thermodynamic potential Φ of the system,

$$\Phi = T\Sigma n_1 \left[\frac{\epsilon_1}{T} - y_1 - \int \frac{\Gamma_1(T)}{T^2} dT - R \ln \frac{RT}{p_1} + R \right] \tag{2, 11}$$

$$= T\Sigma n_1 \left[\frac{\epsilon_1}{T} - y_1 - \int \frac{\Gamma_1'(T)}{T^2} dT - R \ln \frac{RT}{P} + R \ln c_1 + R \right].$$

These expressions are the same as those generally given in ordinary books on thermodynamics*.

(3) *Actual gas mixtures.*

In the case where no chemical reaction takes place on the mixing of the gases (called non-ideal physical mixture), the thermodynamic relations can be treated quite simply if the deviations from the ideal gas mixture are not very large, but the internal energy of the mixture is no longer the sum of the energies of the components as in the case of an ideal gas. Let the difference be ΔU: then

$$U = n_1 u_1 + n_2 u_2 + \ . \ . \ . + n_i u_i - \Delta U. \tag{2, 12}$$

Further the sum of the molecular volumes of the unmixed gases multiplied each by the number of mols. of that gas in the mixture is not the volume of the mixed gas; if the difference is ΔV then

$$V = n_1 v_1 + n_2 v_2 + \ . \ . \ . + n_i v_i - \Delta V; \tag{2, 13}$$

*e.g., see J. R. Partington: *A text-book of thermodynamics*, 1913, p. 270.

thus the heat evolved on mixing the gases, if mixing proceeds under constant volume, is by the first law of thermodynamics,

$$Q_v = \Delta U,$$

and if under constant pressure

$$Q_p = \Delta U + P\Delta V. \tag{2, 14}$$

Case 1: *T and V variable.* In this case U is given from the expressions (2, 12) and (2, 1):

$$U = \sum n_1\left[\epsilon_1 + \int (C_v)_1 dT + T^2 \int \frac{\partial}{\partial T}\left(\frac{P}{T}\right) dv_1\right] - \Delta U, \tag{2, 15}$$

$$PdV = P\sum n_1 dv_1 - Pd\Delta V.$$

If the entropy of the mixture is S, we have

$$TdS = \sum n_1(du_1 + Pdv_1) - d\Delta U - Pd\Delta V. \tag{2, 16}$$

Now if the entropies of the separate gases are $s_1, s_2, \ldots, s_i$,

$$dS = n_1 ds_1 + n_2 ds_2 + \ldots + n_i ds_i - d\Delta S;$$

integrating we have

$$S = n_1 s_1 + n_2 s_2 + \ldots + n_i s_1 + \mathbf{C} - \Delta S \tag{2, 17}$$

where $\mathbf{C}$ is simply a constant depending only on the numbers of mols. $n_1, n_2, \ldots, n_i$ and not on the volume or the temperature.

Further for the free energy Ψ we have

$$\begin{aligned} \Psi &= U - TS \\ &= \sum n_1 u_1 - T\sum n_1 s_1 - T\mathbf{C} - \Delta U + T\Delta S \\ &= \sum n_1 \psi_1 - \Delta U + T\Delta S. \end{aligned} \tag{2, 18}$$

Now put

$$\Delta U - T\Delta S = \sigma_v,$$

$$\frac{\partial^2 \Psi}{\partial T^2} = \sum n_1 \frac{\partial^2 \psi_1}{\partial T^2} - \frac{\partial^2 \sigma_v}{\partial T^2},$$

accordingly

$$C_v = \sum n_1 (C_v)_1 + T\frac{\partial^2 \sigma_v}{\partial T^2}; \tag{2, 19}$$

therefore the specific heat of a gas mixture under constant volume differs by the term $T\dfrac{\partial^2 \sigma_v}{\partial T^2}$ from the sum of the specific heats of the unmixed gases.

Case 2: *T and P variable.*

By a treatment similar to that in the last section we have

$$U = \sum n_1\left[\epsilon_1 + \int (C_p)_1 dT - \int\left(P\frac{\partial v_1}{\partial p} + T\frac{\partial v_1}{\partial T}\right)dP - \int P\frac{\partial v_1}{\partial T} dT\right] - \Delta U, \tag{2, 20}$$

(2, 21) $$S=\Sigma n_1\left[y_1+\int\frac{(C_p)_1}{T}dT-\int\left(\frac{\partial v_1}{\partial T}\right)_p dP\right]-\Delta S+\mathbf{C}',$$

and the thermodynamic potential

(2, 22) $$\begin{aligned}\Phi &= U-ST+PV\\ &=\Sigma n_1[u_1-Ts_1+Pv_1]-T\mathbf{C}'-\Delta U+T\Delta S-P\Delta V\\ &=\Sigma n_1\phi_1-\Delta U+T\Delta S-P\Delta V.\end{aligned}$$

Now put

(2, 23) $$\Delta U-T\Delta S+P\Delta V=\sigma_p,$$

then

$$\frac{\partial^2\Phi}{\partial T^2}=\Sigma n_1\frac{\partial^2\phi_1}{\partial T^2}-\frac{\partial^2\sigma_p}{\partial T^2},$$

(2, 24) $$C_p=\Sigma n_1(C_p)_1+T\frac{\partial^2\sigma_p}{\partial T^2};$$

therefore the specific heat at constant pressure differs by $T\dfrac{\partial^2\sigma_p}{\partial T^2}$ from the sum of the specific heats of the several gases.

For the practical applications of these expressions, the characteristic equation of a gas mixture must be clearly known. In the next chapter I shall go a step forward in this discussion.

III

Characteristic Equations of Gases

(1) *Case of a simple gas.*

The characteristic equation of gases under explosions cannot be found empirically owing to the difficulties of experimenting under such high pressure and high temperature as occur in explosive reactions. The range of temperature within which the experiments were carried out by Amagat* and other investigators, was generally from ordinary temperature to about 260° C: accordingly the characteristic equations of van der Waals and others have only been proved empirically in this range of temperatures. But if the corresponding states of van der Waals be substantially applicable, we can deduce the phenomena of a substance of a considerably higher critical temperature at high temperatures from the behaviour of a gas of very low critical temperature at ordinary temperatures. For example, nitrogen and hydrogen owing to their low critical temperatures may be considered in a state of relatively high temperature compared with substances having high critical points such as carbon dioxide and steam. It is therefore not difficult to assume the behaviour of gases under high pressure and high temperature from the theory of corresponding states.

*E. H. Amagat: *Notes sur la physique et la thermodynamique* 1912; *Mémoires sur l'élasticité et la dilatabilité des fluides jusqu'aux très hautes pressions.* Ann. Chim. Phys. [6], t. 29, p. 68, 1893, etc.

Apart from van der Waals' equation, this applicability of corresponding states has been proved by laborious calculations from the experimental equation* of Kammerlingh Onnes:

$$pv = K_4 t \left\{ 1 + B\frac{K_4}{v} + C\frac{K_4^2}{v^2} + D\frac{K_4^4}{v^4} + E\frac{K_4^6}{v^6} + F\frac{K_4^8}{v^8} \right\} \tag{3, 1}$$

where p, v and t are the corresponding pressure, volume and temperature respectively and B, C, D, E, F are "reduced virial coefficients" which are functions of temperature only,

$$B = b_1 + \frac{b_2}{t} + \frac{b_3}{t^2} + \frac{b_4}{t^4} + \frac{b_5}{t^5}, \text{ etc.} \tag{3, 2}$$

For practical application this equation is too complicated and since the temperature of explosion is so very high compared to the critical temperature of every gas, we can neglect some of the higher powers of $\frac{1}{t}$ for the following reasons:

Gases resulting from explosions are generally Hydrogen, Nitrogen, Carbon dioxide, Carbon monoxide, Methane, Oxygen and Water, and their corresponding temperatures or the ratios of the temperature of explosion to the critical temperature of these gases, can be calculated, assuming the temperature of explosion to be about 3000° C.

At the same time, the corresponding volume $v = \frac{V}{V_k}$, where V_k is the critical volume and V is the volume at 3000° C. and 3000 atmos., can be found from the ideal gas law as shown in Table III.

TABLE III

	T_k	$t = T_0/T_k$	V_k	$v = V_0/V_k$
CO_2	304.1	10	0.00438	1.21
CO	131.9	23	0.00517	1.0
CH_4	177.5	17	0.00467	1.1
O_2	155.0	19	0.00404	1.2
N_2	126.5	24	0.0046	1.1
H_2O	647.0	4.6	0.00386	1.1
H_2	32.0	94	0.00293	1.7

From this table we see that the smallest values are $t=5$ and $v=1$, so that the real volume must be greater than that calculated from the ideal gas law. According to the calculation of Professor Kammerlingh Onnes† it is sufficient in his experimental equation of state to take the terms up to $\frac{1}{v^2}$‡; if $t_a = 1$, $v_a = 10^3$ or $t_a = 2$ and $v_a = 10$.

*H. Kammerlingh Onnes u. W. H. Keesom: *Die Zustandsgleichung*, Enc. Math. Wiss., Bd. v. Teil 1, **10**. Leiden, Comm. Suppl. Nr. 23, 36.

†H. Kammerlingh Onnes, u. W. H. Keesom, *loc. cit. Fussn.* 392.

‡Chwolson: *Traité de Physique*, tome III, p. 852; see also H. Kammerlingh Onnes u. W. H. Keesom, *loc. cit.*

This corresponds to the case of an explosion where $t=5$ and $v=1.0$ and the third term of the reduced virial coefficient is less than the second order, so that his equation of state reduces to

$$Pv=RT\left(1+\frac{b}{v}+\frac{c}{v^2}\right)+\frac{a}{v}. \tag{3, 3}$$

This equation has the same form as van der Waals', H. A. Lorentz's or Brinkman's but it ultimately coincides with van der Waals', when the second virial coefficient is taken as $\frac{1}{v-b}$. [In (3, 3) and in (3, 4) v is the molecular volume, not the 'corresponding volume' of (3, 1)].

E. Sarrau has adopted a formula founded on the equation of Clausius by disregarding the second member since it gradually approaches zero at high temperature. But this formula,

$$P(v+\beta)=RT \tag{3, 4}$$

where β is called the covolume*, had already been investigated by Dupré† in 1869.

This equation can be obtained by omitting the constant "a" from van der Waals' equation.

Noble and Abel's equation of pressure, which is in practically the same form as Dupré's or Sarrau's, is very convenient and useful in the study of Internal Ballistics and for experiments in a closed vessel where only one constant is involved. But these equations are not yet in agreement with experiments. Of the numerous characteristic equations, van der Waals' is very convenient on account of its simplicity and because the values of the constants "a" and "b" have been measured accurately by many experiments; application of this formula to a mixed gas has also been discussed.

(2) *Case of mixed gases.*

The characteristic equations which have already been written, are equally applicable to mixed gases, but the calculation of the characteristic constants, *e.g.*, van der Waal's "a" and "b", from their components requires a very complicated expression even for a binary mixture. For ternary or quaternary mixtures‡ which constitute the gases produced by explosives, this calculation is impossible and the constants must be found experimentally. Van der Waals§ and Lorentz|| have proposed the expressions

$$\begin{cases} a=a_1x^2+2a_{12}x(1-x)+a_2(1-x)^2, \\ b=b_1x^2+2b_{12}x(1-x)+b_2(1-x)^2, \end{cases} \tag{3, 5}$$

*This is really kernvolume and covolume means difference of kern volume and limit volume, H. Kammerlingh Onnes und W. H. Keesom, *Die Zustandsgleichung* Enc. Math. Wiss., Bd. v. Teil 1, 10, *Fussn.* 172.

†Dupré: *Théorie mécanique de la chaleur*, 1869.

‡Schreinemaker: Z. f. phy. Chemie 22, (1897); Roozeboom: *Die heterogenen Gleichgewichte* III, 1; van der Waals: *Die Continuität*; H. Kammerslingh Onnes und W. H. Keesom, *Die Zustandsgleichung*; F. Caubet: *Liquéfaction des Mélanges gazeux*, 1901.

§Die Continuität, p. 79.

||Ann. Phys. u. Chem. Bd. XII (1881), p. 127.

where x and $(1-x)$ are the numerical concentrations of a binary mixture and a_{12} and b_{12} are constants which can be determined by experiment.

Berthelot* assumed

$$a_{12}^2 = a_1 a_2 \tag{3, 6}$$

and van der Waals' equation† gives

$$\sqrt[3]{b_{12}} = \tfrac{1}{2}(\sqrt[3]{b_1} + \sqrt[3]{b_2}) \tag{3, 7}$$

assuming the molecule spherical.

Neither equation is in very good agreement with the experimental results.

If these equations can be generalized so as to include a system of many components and if we let,

$$N = \text{total number of mols.} = \sum n_1,$$
$$n_1,\ n_2,\ \ldots = \text{number of mols. of gases } G_1,\ G_2,\ \ldots$$

then we have the corresponding relations

$$\begin{aligned} N^2 a &= \sum(a_1 n_1^2) + \sum(a_{12} n_1 n_2), \\ N^2 b &= \sum(b_1 n_1^2) + \sum(b_{12} n_1 n_2). \end{aligned} \tag{3, 8}$$

The terms

$$\sum(a_{123} n_1 n_2 n_3) \text{ and } \sum(a_{123\ldots} n_1 n_2 n_3 \ldots)$$

it may be, occur in the equations; but from the Kinetic Theory of Gases it has been proved that the probability of a molecule coming into collision with three or more molecules at the same time is very small and consequently these terms can be neglected‡.

Although a_{pq}, b_{pq} can be determined from experiments, they will not be independent of the concentrations $n_1, n_2, \ldots$ and the relation will be more complicated than in a binary mixture so that it will be better not to use the above assumption, but to take instead the equations

$$\begin{aligned} N_a &= \sum(a_1 n_1), \\ N_b &= \sum(b_1 n_2). \end{aligned} \tag{3, 9}$$

Hereafter I shall chiefly use van der Waals' equation. If Noble and Abel's equation should be required it will be derived by putting

$$a = 0,\ b = \beta.$$

(i) *Calculations of* ΔV *and* ΔP.

The calculation of two terms of deviations,

$$-\Delta V = V - \sum(n_1 v_1),$$

and

$$\Delta P = P - \sum p_1$$

in Section II (2, 13) can be made by means of van der Waals' equation.

*Comptes Rendus Acad. Sciences, t. 126, p. 1858, 1898.

†*loc. cit.*

‡Kuenen: *Verdämpfung und Verflüssigung*, S. 111, etc.; Jeans: *The dynamical theory of gases*, 1916.

By the first approximation of the equation of state we have

$$Pv = RT - \frac{1}{v}(a - bRT)$$

(where v is the molecular volume) and therefore for mixed gases

$$PV = NRT - \frac{N^2}{V}(a - bRT),$$

so that, for the individual gases

$$n_1 P v_1 = n_1 RT - \frac{n_1}{v_1}(a_1 - b_1 RT),$$

$$n_2 P v_2 = n_2 RT - \frac{n_2}{v_2}(a_2 - b_2 RT), \text{ etc.}$$

On combining the above equations, we have

$$P(V - \Sigma n_1 v_1) = -P\Delta V = -\frac{N^2}{V}(a - bRT) + \Sigma\left\{\frac{n_1}{v_1}(a_1 - b_1 RT)\right\},$$

but

$$P\frac{V}{N} = Pv_1 = Pv_2 = \ldots = RT,$$

therefore

$$-\Delta V = \frac{\Sigma n_1 a_1 - Na}{RT} - \{\Sigma(b_1 n_1) - bN\}$$

$$= \frac{N\Sigma n_1 a_1 - \Sigma a_1 n_1^2 - \Sigma a_{12} n_1 n_2}{NRT} - \Sigma(b_1 n_1) + \frac{\Sigma b_1 n_1^2 + \Sigma b_{12} n_1 n_2}{N};$$

this result coincides with that for a binary mixture*.
Similarly for ΔP, we have

$$P = \frac{NRT}{V - bN} - \frac{aN^2}{V^2}, \quad p_1 = \frac{n_1 RT}{V - b_1 n_1} - \frac{a_1 n_1^2}{V^2}, \text{ etc.}$$

but

$$P - \Sigma p_1 = RT\left\{\frac{N}{V - bN} - \Sigma\frac{n_1}{V - b_1 n_1}\right\} - \frac{1}{V^2}(aN^2 - \Sigma a_1 n_1^2)$$

(3, 10)
$$\Delta P = \frac{1}{V^2}\left\{RT\Sigma(b_{12} n_1 n_2) - \Sigma(a_{12} n_1 n_2)\right\}.$$

Now calculate ΔV from the above equation using the approximate value of ΔP; this gives the more convenient form

(3, 11)
$$\Delta V = \frac{V^2}{NRT}\Delta P = \frac{\Sigma(b_{12} n_1 n_2)}{N} - \frac{\Sigma(a_{12} n_1 n_2)}{NRT}.$$

This coincides with the results of experiments†; that is ΔV and ΔP become

*Van der Waals: *Die Continuität II*, p. 54.
†Van der Waals: *Die Continuität II*, p. 58.

greater, when the pressure is higher, and smaller when the temperature is higher.

Accordingly at high temperature as in the case of explosive reactions

$$\Delta V=\frac{\sum(b_{12}n_1n_2)}{N},\quad \Delta P=\frac{RT}{V^2}\sum(b_{12}n_1n_2),$$

i.e., ΔV and ΔP are independent of the terms of mutual attraction a_{12}, etc.

(ii) *Calculation of* ΔU.

The measurement of the increased internal energy due to the mixing of gases is impossible owing to the quantity of heat being so small that it cannot be measured so that the calculations can only be made by the aid of the Clapeyron-Clausius equation.

Let $-\Delta V$ be the volume change due to the mixing of gases,

$$\frac{dP}{dT}=-\frac{1}{T}\frac{W}{\Delta V}.$$

From equation (3, 11)

$$\Delta V=\frac{\sum(b_{12}n_1n_2)}{N}-\frac{\sum(a_{12}n_1n_2)}{NRT}$$

and

$$\frac{dP}{dT}=\frac{R}{v-b},$$

therefore

$$W=-\frac{NRT}{V-Nb}\Delta V;$$

hence

(3, 12) $$\Delta U=\frac{\sum(a_{12}n_1n_2)-RT\sum(b_{12}n_1n_2)}{V}\quad \text{(very closely)}.$$

(iii) *Calculation of* ΔS *and* σ_v *when* T *and* V *are variables.*

(*a*) *Calculation of* ΔS.

From the differences of the following thermodynamic relations,

$$\frac{\partial U}{\partial T}=T\frac{\partial S}{\partial T},\quad \frac{\partial U}{\partial V}+P=T\frac{\partial S}{\partial V},\quad \left(\frac{\partial U}{\partial V}\right)_T=T^2\frac{\partial}{\partial T}\left(\frac{P}{T}\right),$$

we have for deviations Δ,

$$\frac{\partial\Delta U}{\partial T}=T\frac{\partial\Delta S}{\partial T},\quad \frac{\partial\Delta U}{\partial V}-\Delta P=T\frac{\partial\Delta S}{\partial V},\quad \frac{\partial\Delta U}{\partial V}=-T^2\frac{\partial}{\partial T}\left(\frac{\Delta P}{T}\right).$$

Combining these three equations we have

$$T\frac{\partial \Delta S}{\partial V}=\frac{\partial \Delta U}{\partial V}-\Delta P=-T^2\frac{\partial}{\partial T}\left(\frac{\Delta P}{T}\right)-\Delta P=-T\frac{\partial \Delta P}{\partial T}$$

or

$$-\frac{\partial \Delta P}{\partial T}=\frac{\partial \Delta S}{\partial V},$$

hence

$$d\Delta S=\frac{\partial \Delta S}{\partial T}dT+\frac{\partial \Delta S}{\partial V}dV=\frac{1}{T}\frac{\partial \Delta U}{\partial T}dT-\frac{\partial \Delta P}{\partial T}dV;$$

integrating we have

$$\text{(3, 13)}\quad \Delta S=\int\frac{1}{T}\frac{\partial \Delta U}{\partial T}dT-\int\frac{\partial \Delta P}{\partial T}dV=-\frac{R(1-\ln T)}{V}(\Sigma b_{12}n_1n_2).$$

(*b*) *Calculation of* σ_v.

From (2, 18) and previous relations σ_v can be calculated thus,

$$\text{(3, 14)}\quad \sigma_v=\Delta U-T\Delta S=\frac{1}{V}\left[\Sigma(a_{12}n_1n_2)+RT\ln T\,\Sigma(b_{12}n_1n_2)\right].$$

Making use of the relation

$$\frac{\partial^2\sigma_v}{\partial T^2}=\frac{R}{VT}\Sigma(b_{12}n_1n_2)$$

we obtain for the specific heat at constant volume, the formula

$$\text{(3, 15)}\quad C_v=\Sigma n_1(C_v)_1+T\frac{\partial^2\sigma_v}{\partial T^2}=\Sigma n_1(C_v)_1+\frac{R}{V}\Sigma(b_{12}n_1n_2).$$

(iv) *Calculation of* ΔS *and* σ_p *where* T *and* P *are variables.*

(*a*) *Calculation of* ΔS.

Using the thermodynamic relations,

$$\left(\frac{\partial S}{\partial P}\right)_T=\left(-\frac{\partial V}{\partial T}\right)_P,\quad \frac{\partial S}{\partial T}=\frac{1}{T}\frac{\partial U}{\partial T},$$

we have

$$d\Delta S=\frac{\partial \Delta S}{\partial T}dT+\frac{\partial \Delta S}{\partial P}dP=\frac{1}{T}\frac{\partial \Delta U}{\partial T}dT-\frac{\partial \Delta V}{\partial T}dP,$$

whence

$$\text{(3, 16)}\quad \Delta S=\int\frac{1}{T}\frac{\partial \Delta U}{\partial T}dT-\int\frac{\partial \Delta V}{\partial T}dP=\frac{3}{2}\frac{P}{NRT}\Sigma(a_{12}n_1n_2).$$

(*b*) *Calculation of* σ_p.

From (2, 23) and (3, 16) we have

(3, 17) $$\sigma_p = \Delta U - T\Delta S + P\Delta V = -\frac{3}{2}\frac{P}{NRT}\Sigma(a_{12}n_1n_2);$$

similarly for the specific heat at constant pressure

$$T\frac{\partial^2\sigma_p}{\partial T^2} = -\frac{3P}{NRT^2}\Sigma(a_{12}n_1n_2)$$

and therefore

(3, 18) $$C_p = \Sigma n_1(C_p)_1 + T\frac{\partial^2\sigma_p}{\partial T^2} = \Sigma n_1(C_p)_1 - \frac{3P}{NRT^2}\Sigma(a_{12}n_1n_2).$$

(v) *Case when* $a_{12} = b_{12} = 0$.

If the constants "a" and "b" have in the mixture the relations,

$$Na = \Sigma n_1a_1, \quad Nb = \Sigma n_1b_1$$

then on substituting these relations in the previous equations we have

$$\Delta U = 0,\ \Delta V = 0,\ \Delta S = 0,\ \sigma_v = 0,\ \sigma_p = 0,$$
$$C_v = \Sigma n_1(C_v)_1,\ C_p = \Sigma n_1(C_p)_1;$$

these are the same as in the case of an ideal gas mixture.

IV

Chemical Equilibrium

In this Section we shall show how to calculate the chemical equilibrium of actual gas mixtures at high temperature and under high pressure, such as occur in the gases resulting from an explosion, using the relations given in Section II. The discussion will be conveniently given in two cases, taking T, v and T, p as variables.

(1) *T and V variable.*

Consider one mol. of a gas with a single component and calculate

$$\int\frac{\partial}{\partial T}\left(\frac{P}{T}\right)dV,\ \int\frac{\partial P}{\partial T}dV,\ \int PdV$$

in the equations (2, 1), (2, 2), (2, 3) by means of the characteristic equation of van der Waals.

We have

(4, 1) $$U = \epsilon + \int C_v dT - \frac{a}{V},$$

(4, 2) $$S = y + \int\frac{C_v dT}{T} + R\ln(V-b),$$

(4, 3) $$\psi = V - ST = \epsilon - yT - T\int\frac{\Gamma(T)}{T^2}dT - RT\ln(V-b) - \frac{a}{V}.$$

The Free Energy of gaseous mixtures can be similarly expressed, but in this case the volume of the gas is not that of the gases in their separate states but the whole volume of the gas mixture V.

It follows,

$$v_1 = \frac{V-\Delta V}{n_1}, \quad v_1 - b_1 = \frac{V}{n_1}\left(1-\frac{\Delta V}{V}\right) - b_1 = \frac{V}{n_1}\left(1-\frac{\Delta V}{V}\right)\left(1-b_1\frac{n_1}{V}\right).$$

Let $\xi_1, \xi_2, \ldots$ be the volumetric molecular concentrations,

$$\frac{n_1}{V} = \xi_1, \text{ etc.}$$

and from (2, 17)

$$\mathbf{C} = R\Sigma n_1\left[\ln \xi_1 - \ln\left(1-\frac{\Delta V}{V}\right) - \ln(1-b_1\xi_1)\right]. \tag{4, 4}$$

The Free Energy of a gaseous mixture is, by (2, 10),

$$\Psi = \Sigma n_1\psi_1 + RT\Sigma n_1\left[\ln\xi_1 - \ln\left(1-\frac{\Delta V}{V}\right) - \ln(1-b_1\xi_1)\right] - \sigma_v$$

$$= \Sigma n_1\left[\epsilon_1 - y_1 T - T\int\frac{\Gamma_1(T)}{T^2}dT - \frac{a_1}{V} + RT\ln\xi_1 - RT\ln(1-b_1\xi_1)\right]$$

$$-NRT\ln\left(1-\frac{\Delta V}{V}\right) - \sigma_v.$$

Now put

$$\frac{\epsilon_1}{T} - y_1 - \int\frac{\Gamma_1(T)}{T^2}dT = f_1, \text{ etc.} \tag{4, 5}$$

and take the first terms of the expansions of

$$\ln\left(1-\frac{\Delta V}{V}\right) \text{ and } \ln(1-b_1\xi_1),$$

that is to say $-\frac{\Delta V}{V}$ and $-b_1\xi_1$ respectively. We obtain the formulae:

$$\Psi = T\Sigma n_1 f_1 + RT\Sigma n_1[\ln\xi_1 + b_1\xi_1] + NRT\frac{\Delta V}{V} - \Sigma a_1\xi_1 - \sigma_v \tag{4, 6}$$

$$= T\Sigma n_1 f_1 + RT\Sigma n_1 \ln\frac{n_1}{V} + NRT\frac{\Delta V}{V} + RT\frac{\Sigma n_1^2 b_1}{V} - \frac{\Sigma n_1 a_1}{V} - \sigma_v,$$

$$\mu_1 = \frac{\partial\Psi}{\partial n_1} = Tf_1 + RT\ln\frac{n_1}{V} + RT + \frac{NRT}{V}\frac{\partial\Delta V}{\partial n_1} + RT\frac{2n_1b_1}{V^2} - \frac{a_1}{V} - \frac{\partial\sigma_v}{\partial n_1}, \tag{4, 7}$$

where μ_1 is the chemical potential of the component G_1.

Now let the correction be $T\Delta$ so that

$$NRT\frac{\Delta V}{V}+RT\frac{\Sigma n_1^2 b_1}{V}-\frac{\Sigma n_1 a_1}{V}-\sigma_v=T\Delta, \tag{4, 8}$$

then

$$\Psi=T\Sigma n_1 f_1+RT\Sigma n_1 \ln\frac{n_1}{V}+T\Delta \tag{4, 9}$$

and for an ideal gas Δ must be zero.

If any changes of composition have resulted from reversible chemical changes, there will be simultaneously a variation of Ψ which may be assumed to have varied continuously.

If, besides the temperature, the total volume V be maintained constant, the condition of equilibrium is that the free energy must be a minimum:

$$\delta\Psi=0,$$

subject to the conditions:

$$\delta V=0,\ \delta T=0.$$

Thus we have the canonical potential equations at constant volume:

$$\delta\Psi=T\Sigma\left[f_1+R+R\ln\frac{n_1}{V}\right]\delta n_1+T\Sigma\frac{\partial\Delta}{\partial n_1}\delta n_1=0.$$

In general, if the state of equilibrium is represented by the chemical equation with ordinary stoichiometric coefficients $\nu_1, \nu_2, \ldots$

$$\nu_1 A_1+\nu_2 A_2+\ldots \rightleftharpoons \nu_1' A_1'+\nu_2' A_2'+\ldots \tag{4, 10}$$

and

$$\delta n_1=\nu_1\lambda,\ \delta n_2=\nu_2\lambda, \text{ etc.}$$

where the λ's are positive magnitudes independent of the ν's.

Let μ_1 be the molecular chemical potential of the gas G_1 in the mixture, then

$$\mu_1=\left(\frac{\partial\psi}{\partial n_1}\right)_{TV},$$

or

$$\mu_1=T(f_1+R+R\ln\xi_1)+T\frac{\partial\Delta}{\partial n_1};$$

the equation of equilibrium is therefore $\Sigma(\mu_1\nu_1-\mu_1'\nu_1')=0$, which for brevity we shall indicate by

$$\Sigma\mu_1\nu_1=0,$$

or

$$\Sigma\nu_1\ln\xi_1=-\frac{1}{R}\Sigma\nu_1\left(f_1+R+\frac{\partial\Delta}{\partial n_1}\right).$$

Let K be the equilibrium constant,

$$\ln K_v=\Sigma\nu_1\ln\xi_1=-\frac{1}{R}\Sigma\nu_1\left[\frac{\epsilon_1}{T}-\int\frac{\Gamma_1(T)}{T^2}dT-y_1+R+\frac{\partial\Delta}{\partial n_1}\right]; \tag{4, 11}$$

here $\Sigma\nu_1\epsilon_1$ corresponds to the heat of reaction Q_0, referred to a standard condition.

(2) *T and P variable.*

As in the last case, the thermodynamic potential Φ can be expressed by means of van der Waals' equation; but in this case the mathematical process is not so simple as in the last and some abbreviations are necessary.

For one mol. of a single gas, we have from (2, 6)

$$\Phi = U - ST + PV = \epsilon - yT - T\int\frac{\Gamma'(T)}{T^2}dT + RT\ln P - \frac{ae}{(RT)^2}P^{2*}. \tag{4, 12}$$

In the gaseous mixture, $c_1, c_2, \ldots$ are the numerical concentrations of the components $G_1, G_2, \ldots$. We have

$$c_1 = \frac{n_1}{\Sigma n_1} = \frac{p_1}{P+\Delta P}, \quad p_1 = c_1 P\left(1+\frac{\Delta P}{P}\right), \tag{4, 13}$$

then

$$\Phi = \Sigma n_1\left[\epsilon_1 - y_1 T - T\int\frac{\Gamma_1'(T)}{T^2}dT + RT\ln p_1 - \frac{a_1 e_1}{(RT)^2}p_1^2\right] + \sigma_p \tag{4, 14}$$

$$= \Sigma n_1\left[\epsilon_1 - y_1 T - T\int\frac{\Gamma_1'(T)}{T^2}dT + RT\ln c_1 + RT\ln P\right.$$

$$\left. + RT\ln\left(1+\frac{\Delta P}{P}\right) - \frac{a_1 e_1}{(RT)^2}c_1^2 P^2\right] + \sigma_p.$$

Now let

$$g_1(T) = \frac{\epsilon_1}{T} - y_1 - \int\frac{\Gamma_1'(T)}{T^2}dT + R\ln P \tag{4, 15}$$

and

*Expanding van der Waals' equation (in which v is the volume of one mol.):

$$(v-b)P + \frac{a}{v} - \frac{ab}{RT^2}P^2 = RT$$

and differentiating we have:

$$P\frac{\partial v}{\partial P} = -\frac{RTP}{\left(P^2 - \frac{a^2}{v^4}\right)} + \frac{2abP}{RT^2},$$

thence

$$\int P\frac{\partial v}{\partial P}dP = -RT\ln P + \frac{RT}{2}\ln\left(1-\frac{a^2P^2}{(RT)^4}\right) + \frac{abP^2}{(RT)^2}$$

$$= -RT\ln P + \frac{a}{(RT)^2}\left[b - \tfrac{1}{2}\frac{a}{RT}\right]P^2$$

$$= -RT\ln P + \frac{ae}{(RT)^2}P^2, \quad \left(e = b - \tfrac{1}{2}\frac{a}{RT}\right).$$

(4, 16)
$$NR\ln\left(1+\frac{\Delta P}{P}\right)-\frac{P^2}{R^2T^3}\sum(n_1c_1^2a_1e_1)+\frac{\sigma_p}{T}=\nabla;$$

then

(4, 17)
$$\Phi=T\sum n_1[g_1(T)+R\ln c_1]+T\nabla.$$

In the case of an ideal gas ∇ should be zero.

In the state of equilibrium,

$$\delta\Phi=0,$$

the canonical potential equation should be modified by terms containing ∇, thus

$$\sum\left[(g_1(T)+R\ln c_1+\frac{\partial\nabla}{\partial n_1}\right]\delta n_1=0,$$

and the molecular chemical potential will be

(4, 19)
$$\mu_1=\left(\frac{\partial\Phi}{\partial n_1}\right)_{PT}=Tg_1(T)+TR\ln c_1+T\frac{\partial\nabla}{\partial n_1}.$$

The equation of equilibrium is therefore

$$\sum\nu_1\ln c_1=-\frac{1}{R}\sum\nu_1g_1(T)-\frac{1}{R}\sum\nu_1\frac{\partial\Delta}{\partial n_1}.$$

Substituting in (4, 15)

(4, 20)
$$\ln K_p=\sum\nu_1\ln c_1=-\frac{1}{R}\sum\nu_1\left[\frac{\epsilon_1}{T}-y_1-\int\frac{\Gamma_1'(T)}{T^2}dT+R\ln P+\frac{\partial\nabla}{\partial n_1}\right].$$

The equilibrium constant may be expressed in terms of partial pressures instead of numerical concentrations; letting this constant be K_p' we have

(4, 21)
$$\sum\nu_1\ln p_1=\ln K_p'=-\frac{1}{R}\sum\nu_1\left[\frac{\epsilon_1}{T}-y_1-\int\frac{\Gamma_1'(T)}{T^2}dT\right]+\frac{\partial\nabla'}{\partial n_1}$$

where

$$\nabla'=-\frac{P^3}{R^2T^3}\sum\frac{n_1^3}{N^3}a_1e_1+\frac{\sigma_p}{T}.$$

Let Q_p be the heat of reaction at constant pressure, then

(4, 22)
$$\ln K_p'=-\frac{Q_0+\sum\nu_1\int(C_p)_1dT}{RT}+\frac{\sum\nu_1}{R}\int\frac{(C_p)_1}{T}dT+\frac{\partial\nabla'}{\partial n_1}+\frac{\sum\nu_1y_1}{R};$$

here $\frac{y}{R}$ is the term corresponding to the chemical constant.

(3) *Calculation of* Δ *and* ∇.

Calculation of Δ. From (4, 8)

$$\Delta = NR\frac{\Delta V}{V} + R\frac{\Sigma n_1^2 b_1}{V} - \frac{\Sigma n_1 a_1}{VT} - \frac{\sigma_v}{T}$$

$$= \frac{R}{V}\{\Sigma(n_1^2 b_1) + (1-\ln T)\Sigma(b_{12}n_1n_2)\} - \frac{1}{VT}\{\Sigma(n_1a_1) + 2\Sigma(a_{12}n_1n_2)\};$$

hence

$$\frac{\partial\Delta}{\partial n_1} = \frac{R}{V}\left\{2(n_1b_1) + (1-\ln T)\Sigma(b_{12}n_2)\right\} - \frac{1}{VT}\left\{2\Sigma(a_{12}n_2) + a_1\right\}.$$

If $a_{12} = b_{12} = 0$

$$\Delta = \frac{R}{V}\Sigma(n_1^2 b_1) - \frac{1}{VT}\Sigma(n_1a_1),$$

(4, 23)
$$\frac{\partial\Delta}{\partial n_1} = \frac{R}{V}\left(2n_1b_1 - \frac{a_1}{RT}\right).$$

In either case $\Sigma\nu_1\frac{\partial\Delta}{\partial n_1}$ is the difference of the stoichiometric coefficients of the two sides in the chemical equation.

Calculation of ∇. From (4, 16)

$$\nabla = NR\ln\left(1+\frac{\Delta P}{P}\right) - \frac{P^2}{R^2T^3}\Sigma(n_1c_1^2a_1e_1) + \frac{\sigma_p}{T}$$

$$= +\frac{P}{NT}\Sigma(b_{12}n_1n_2) - \frac{5P}{2NRT^2}\Sigma(a_{12}n_1n_2) - \frac{P^2}{R^2T^3}\Sigma(n_1c_1^2a_1e_1);$$

similarly

$$\frac{\partial\nabla}{\partial n_1} = -\frac{5P}{2NRT^2}\Sigma a_{12}n_2 + \frac{P}{NT}\Sigma b_{12}n_2 - \frac{3P^2}{N^2R^2T^3}n_1^2a_1\left(b_1 - \frac{1}{2}\frac{a_1}{RT}\right).$$

If $b_{12} = a_{12} = 0$

$$\nabla = -\frac{P^2}{N^2R^2T^3}\Sigma n_1^3a_1\left(b_1 - \frac{1}{2}\frac{a_1}{RT}\right),$$

$$\frac{\partial\nabla}{\partial n_1} = -\frac{3P^2}{N^2R^2T^3}n_1^2a_1\left(b_1 - \frac{1}{2}\frac{a_1}{RT}\right).$$

These expressions contain the terms ab and a^2 and therefore can be neglected even in the cases where $\left(\frac{P^2}{T^3}\right)$ is large, and therefore

$$\nabla = 0. \tag{4, 24}$$

V

Equilibrium Constant at High Temperature

Modern explosives generally contain as their chemical constituents carbon, hydrogen, nitrogen and insufficient oxygen for complete oxidation. The gases resulting from them are consequently carbon dioxide, carbon monoxide, water, hydrogen and nitrogen. Sometimes small quantities of methane are found in the gases resulting from the explosion, but this is not a principal constituent of the reaction and therefore the convertible components which enter into the discussion of chemical equilibrium at high temperature are CO_2, CO, H_2O and H_2 (nitrogen being an inert gas) that is "water-gas reaction".

Previous experiments by several investigators on the equilibrium of water gas were generally performed at temperatures below 1500° C. The calculations of chemical equilibrium at temperatures from about 3000° to 4000° C. such as occur in an explosion, can easily be made by means of the Nernst-theorem, but the specific heat of gases has to be expressed accurately as a function of the temperature up to such high temperatures.

In the following pages, formulae of water gas reaction as given by several authorities are compared and new calculations are also given by the aid of relatively exact formulae of specific heat at high temperature.

Physico-Chemical Data

The physico-chemical numbers which are used in connection with these calculations, are given in Tables I and II.

(i) *Physico-chemical constants of gases and vapours.*

TABLE I

	Mol. Wt.	Observed density 0° C., 760 mm. air = 1	Mol. Vol. calculated from the observed density	Observed Wt. of 1 litre of gas at 0° C., 760 mm.	Volume of 1 kg. gas in litre at 0° C., 760 mm.
CO_2	44.005	1.52908	22258	1.9768	506
CO	28.005	0.9673	22395	1.2507	800
Air	......	1.0000	.....	1.2927	773
CH_4	16.037	0.5545	22363	0.7168	1395
N_2	28.02	0.96727	22407	1.2505	800
H_2O	18.016	0.6219 (cal.)	22408 (cal.)	0.8040 (cal.)	1243
H_2	2.016	0.06952	22433	0.08987	11125

TABLE II

HEAT OF FORMATION

			Const. Pressure		Const. Volume	
			per mol.	per gram.	per mol.	per gram.
CO_2	C (dia) $+O_2$	15° C.	+94,310	2143	94,310	2143
	C (amorph) $+O_2$	15°	97,600	2220	97,600	2220
CO	C (dia) +O	15°	26,100	932	26,400	943
	C (amorph) +O	15°	29,400	1050	29,700	1060
CO_2	CO +O	15°	68,200	1550	67,900	154
H_2O (gas)	H_2 +O	15°	58,100	3228	57,900	322
H_2O (liq.)	H_2 +O	15°	68,700	3820	67,900	377
CH_4	C (dia) $+2H_2$	15°	18,400*	1150	17,800	1110
	C(amorph) $+2H_2$	15°	21,700	1360	21,100	132

*Taking the heat of combustion of methane as 213,500 at const. pressure.

(ii) *Formulae for the calculation of equilibrium in water gas reactions.*

The equilibrium constants of a homogeneous system can easily be calculated by the Nernst-theorem for an ideal gas.

Let the chemical equation for the reaction be

(5, 1) $$\nu_1 x_1 + \nu_2 x_2 + \ldots = \nu_1' x_1' + \nu_2'' x_2'' + \ldots + Q \text{ cal.}$$

where ν_i is the stoichiometric coefficient of gas G_i, etc., then from (4, 22)

(5, 2) $$\ln K_p' = -\frac{Q_0}{RT} + \frac{1}{R}\int_0^T \frac{dT}{T^2}\int_0^T \Sigma\nu_1 (C_p)_1 dT + \Sigma\nu_1 i_1.$$

Nernst has expressed the sum of the specific heats by

$$\Sigma\nu_1 (C_p)_1 = 3.5\Sigma\nu_1 + 2\Sigma\nu_1\beta T + 3\Sigma\nu_1\alpha_1 T^2 + \ldots$$

in this case,

(5, 3) $$\ln K_p = -\frac{Q_0}{4.571T} + 1.75\Sigma\nu_1 \log T + \frac{\Sigma\nu_1\beta}{4.571}T$$

$$+\frac{\Sigma\nu_1 a}{2\times 4.571}T^2 + \ldots + \Sigma\nu_1 \mathbf{C}_1$$

where $\mathbf{C}_i$ is the chemical constant of the gas G_i.

As the thermochemical equation of water-gas is

(5, 4) $$CO + H_2O \rightleftharpoons CO_2 + H_2 + 10100 \text{ cal.}$$

the equilibrium constant

(5, 5) $$K = \frac{(C_{H_2O})(C_{CO})}{(C_{CO_2})(C_{H_2})}$$

is independent of the pressure since this is the reaction of gases "without condensation".

The following formulae for equilibrium have been obtained by several investigators:

Haber's formula*

$$\log K = -\frac{2116}{T} + 0.783 \log T - 0.00043T. \qquad ((1))$$

Nernst's formula†

$$\log K = -\frac{2225}{T} - 0.0004T + 1.9. \qquad ((2))$$

Haber's second formula‡

$$\log K = -\frac{2170}{T} + 0.979 \log T - 1.082\times10^{-3}T + 1.734\times10^{-7}T^2 - 0.02858. \qquad ((3))$$

Hahn's formula §

$$\log K = -\frac{2232}{T} - 0.08463 \log T - 0.0002203T + 2.5084. \qquad ((4))$$

Schreber's formula||

$$\log K = -\frac{2213}{T} + 0.04061 \log T - 0.000158T + 2.0266. \qquad ((5))$$

Langen's formula||

$$\log K = -\frac{2245}{T} - 0.2783 \log T - 0.0000981T + 2.9653. \qquad ((6))$$

Hahn's second formula¶

$$\log K = -\frac{2226}{T} - 0.0003909T + 2.4506. \qquad ((7))$$

These equations are plotted in Figure 1, the number on the curve corresponding to the number of the formula.

(iii) *Author's calculations of equilibrium.*

The previous equations give values in agreement with the experimental results within certain limits of temperature but not for the higher temperatures as shown in Figure 1. This is due to the uncertainty of the specific heat-temperature function at these temperatures. To correct this we calculated the equilibrium constant by means of the Nernst-theorem, using the results of the specific heat which is applied** up to comparatively high temperatures although the experimental results at high temperatures are only given by the equations of Pier†† and Bjerrum‡‡.

*F. Haber: *Thermodynamics of technical gas reactions* (English Translation), p. 143.
†R. Abegg: *Handbuch der anorganischen Chemie Bd. III*, 2te *Abt.*, S. 198.
‡Z. f. phys. Chem. Bd. 68, S. 731, 1910.
§Z. f. phys. Chem. Bd. 44, S. 513, 1903.
||F. Haber: *Thermodynamics of technical gas reactions* (English Translation), p. 143.
¶Z. f. phys. Chem. Bd. 48, S. 735, 1904.
**N. Yamaga, *loc. cit.*
††Z. f. Elektrochemie Bd. 15, S. 536, 1909; Bd. 16, S. 897, 1910.
‡‡Z. f. Elektrochemie Bd. 17, S. 731, 1911; Bd. 18, S. 101, 1912.

(*a*) *Pier's experimental equations are* for

Hydrogen

(5, 6) $$C_v=4.454+0.0009T,$$

Diatomic gas (H_2 excluded)

(5, 7) $$C_v=4.654+0.0009T,$$

Carbon dioxide

(5, 8) $$C_v=4.780+8.25\times10^{-3}T-3.18\times10^{-6}T^2+0.4\times10^{-9}T^3,$$

Steam

(5, 9) $$C_v=5.776+1.18\times10^{-3}T-0.655\times10^{-6}T^2+0.8\times10^{-9}T^3.$$

(*b*) *Bjerrum has given theoretical formulae for the specific heats by means of the Nernst-Lindemann equation. Let $\overline{C}_v$ be the mean specific heat, then* for

Hydrogen

(5, 10) $$\overline{C}_v=\frac{5}{2}R+R\phi(\lambda=2.0)$$

where

$$\phi=\tfrac{1}{2}\frac{\frac{\beta c}{\lambda T}}{e^{\frac{\beta c}{\lambda T}}-1}+\tfrac{1}{2}\frac{\frac{\beta c}{2\lambda T}}{e^{\frac{\beta c}{2\lambda T}}-1},$$

Diatomic gas

(5, 11) $$\overline{C}_v=\frac{5}{2}R+R\phi(\lambda=2.4),$$

Carbon dioxide

(5, 12) $$\overline{C}_v=R[3\phi(\lambda=8.1)+2\phi(\lambda=5.0)],$$

Steam

(5, 13) $$\overline{C}_v=R\left[3+2\phi(1.3)+\phi(3.6)+\left(\frac{T}{3100}\right)^4\right].$$

(*c*) *Calculation of the equilibrium constant K by means of Pier's results for specific heats.*

The heat evolved by the reaction under constant pressure at the temperature 15° C. is by Table II,

$$CO+H_2O=CO_2+H_2+10100 \text{ cal.}$$

The specific heat of the gaseous mixture is given by the algebraic sum of these four constituents:

(5, 14) $$C_p=\sum\nu_1(C_p)_1=1.196-0.00707T+2.525\times10^{-6}T^2+0.4\times10^{-9}T^3.$$

Hence the heat of reaction at absolute zero is

(5, 15) $$Q_0=Q_{15^\circ}-55=10045 \text{ cal.}$$

If we put these values in the equation of equilibrium (5, 2), we have

$$\ln K=-\frac{10045}{RT}+0.600\ln T-\frac{.00353T}{R}+\frac{0.42\times10^{-6}T^2}{R}+\text{Const.}$$

We multiply each term by 0.4343 to transform from the natural to the common logarithms and finally obtain:

$$\log K = -\frac{2196}{T} + 0.60 \log T - 0.000773T + 0.092\times10^{-6}T^2 + 0.0065\times10^{-9}T^3 + 0.8930. \tag{(8)}$$

The constant 0.8930 is given to make the equation coincide with the experimental results, but this constant is not equal to Nernst's sum of chemical constants 2, 4, because in the latter case

$$i = w - \frac{a+R^*}{R},$$

Nernst having taken for the specific heat

$$C_p = 3.5 + 2\beta_1 T.$$

(*d*) *Calculation of the equilibrium constant by means of Bjerrum's results for specific heat.*

As in the preceding section, the sum of his specific heats is

$$\Sigma\int \nu_1 (C_v)_1 dT = T\overline{C}_v = RT\left[\phi(2.0)+2\phi(1.3)+\phi(3.6)-\phi(2.4)-2\phi(5.0)-\phi(8.1)+\left(\frac{T}{3100}\right)^4\right]; \tag{5, 16}$$

accordingly

$$Q_0 = 10100 - \int_0^{290} \Sigma\nu_1 (C_p)_1 dT = 10100 - 78.7 = 10021 \text{ cal.} \tag{5, 17}$$

The terms in equation (5, 2) are easily integrated and the equation becomes

$$\int \frac{1}{T^2}\int C_p dT = \int \left\{ \frac{1}{2}\,\frac{\frac{\beta c}{\lambda T^2}}{e^{\frac{\beta c}{\lambda T}}-1} + \frac{1}{2}\,\frac{\frac{\beta c}{2\lambda T^2}}{e^{\frac{\beta c}{2\lambda T}}-1} \right\} dT = \frac{3\beta c}{4\lambda T} - \frac{1}{2}\ln\left(e^{\frac{\beta c}{\lambda T}}-1\right) - \frac{1}{2}\ln\left(e^{\frac{\beta c}{2\lambda T}}-1\right). \tag{5, 18}$$

The equation of equilibrium is

$$\ln K = -\frac{10021}{RT} - F(2.0) - 2F(5.0) - F(8.1) + F(2.4) + 2F(1.3) + F(3.6) + \frac{1}{4}\left(\frac{T}{3100}\right)^4 + \Sigma\nu_1 i_1 \tag{5, 19}$$

where

$$F(\lambda) = \frac{1}{2}\left[\frac{3}{2}\,\frac{\beta c}{\lambda T} - \ln\left(e^{\frac{\beta c}{\lambda T}}-1\right) - \ln\left(e^{\frac{\beta c}{2\lambda T}}-1\right)\right].$$

*Partington: *A text-book of thermodynamics*, p. 499.

On transforming natural to common logarithms we obtain

((9))
$$\log K = -\frac{2190.4}{T} + 0.1086\left(\frac{T}{3100}\right)^4$$
$$-0.4343[F(2.0)+2F(5.0)+F(8.1)-F(2.4)$$
$$-2F(1.3)-F(3.6)]+\Sigma\nu_1\mathbf{C}_1;$$

here

$$\Sigma\nu_1\mathbf{C}_1 = 3.5(CO)+3.6(H_2O)-3.2(CO_2)-1.6(H_2)=2.30.$$

This value of the chemical constant is in good agreement with the results of experiments which give in accord with ((9)):

$$\Sigma\nu_1\mathbf{C}_1 = 2.27.$$

(iv) *Comparison of the equations* ((1)) *to* ((9)).

Table III shows the equilibrium constants over a wide range of temperature as calculated from the equations ((1)) to ((9)), and Figure 1 expresses the results graphically.

As we see in the figure, the equations are in good agreement within the limits of the experiments, *i.e.*, from 500° up to 2000° absolute. Above this limit equations ((1)), ((3)), ((4)) and ((7)) give widely divergent values, due to the uncertainty of the specific heat functions at the higher temperatures.

The only equations which can be applied to these high temperatures are Haber's second formula and the author's calculations by means of Pier's and Bjerrum's equations for the specific heats.

Although it is unfortunately impossible to verify the theoretical calculation by experimental researches at temperatures from 2500° to 4000°, yet these three equations give nearly the same values. As the specific heats given by Bjerrum based on Pier's experiments can be considered exact up to high temperature, the equations of equilibrium as obtained from these specific heats will also be correct and the constant of integration coincides with the Nernst chemical constant.

We shall therefore in this paper discuss the explosive gaseous reaction by means of equation ((9)).

(v) *Reactions in Methane formation.*

In the experiments on the explosion of powder in a closed vessel, it is found that methane always exists in the gases evolved and the quantity of this gas is greater, the greater the pressure of explosion, *i.e.*, the greater the density of loading.

The following three equations generally give the reactions in the formation of methane from water-gas:

(5, 20) $$CO + 3H_2 = CH_4 + H_2O,$$
(5, 21) $$CO_2 + 4H_2 = CH_4 + 2H_2O,$$
(5, 22) $$2CO + 2H_2 = CH_4 + CO_2.$$

TABLE III

Temp	Equation	Haber I 1	Nernst 2	Haber II 3	Hahn I 4	Schreben 5	Langen 6	Hahn II 7	Yamaga I 8	Yamaga II 9
300	log. k.	−5.2425	−5.637	−5.6321	−5.248			−5.086	−5.1715	−5.032
	k.	5.723×10^{-6}	2.307×10^{-6}	2.328×10^{-6}	6.237×10^{-6}			8.204×10^{-6}	6.74×10^{-6}	9.290×10^{-6}
500	log. k.	−2.978	−2.750	−2.223	−2.2913	−2.369	−2.3243	−2.197	−2.2506	−2.178
	k.	1.052×10^{-3}	1.778×10^{-3}	5.984×10^{-3}	1.505×10^{-3}	4.276×10^{-3}	4.740×10^{-3}	6.353×10^{-3}	5.61×10^{-3}	6.637×10^{-3}
800	log. k.	−0.716	−1.201	−0.5970	−0.7034			−0.6446	−0.6761	−0.6419
	k.	1.923×10^{-1}	6.295×10^{-2}	2.529×10^{-1}	1.980×10^{-1}			2.267×10^{-1}	2.11×10^{-1}	0.2281
1000	log. k.	−0.197	−0.725	−0.170	−0.1944	−0.2226	−0.2127	−0.1663	−0.1837	−0.170
	k.	6.353×10^{-1}	1.884×10^{-1}	6.761×10^{-1}	6.383×10^{-1}	5.989×10^{-1}	6.128×10^{-1}	6.808×10^{-1}	0.6551	0.6761
1100	log. k.	−0.016	−0.560	−0.0078	−0.0180			−0.0194	−0.0254	
	k.	9.638×10^{-1}	2.75×10^{-1}	9.817×10^{-1}	9.594×10^{-1}			9.550×10^{-1}	0.9432?	
1500	log. k.	+0.4262	−0.190	−0.4013	−0.4213			0.32965	+0.3910	+0.3867
	k.	2.667	6.457×10^{-1}	2.518	2.636			2.133	2.460	2.436
2000	log. k.	+0.6767	−0.0125	0.6467	0.6724	+0.7382	+0.7280	0.5558	+0.6438	+0.6298
	k.	4.742	9.705×10^{-1}	4.426	4.699	5.473	5.346	3.589	4.403	4.264
3000	log. k.	+0.7271	+0.0417	0.9690	0.8092	0.9562	0.9551	0.5360	0.92786	0.8832
	k.	5.333	1.009	9.311	6.399	9.0408	9.018	3.436	8.470	7.642
4000	log. k.	+0.570	+0.2562	1.395	0.7645	0.9877	1.0092	0.3301	1.2946	1.1427
	k.	3.715	1.803	24.83	5.808	9.721	10.21	2.138	19.71	13.89

These are not independent but are connected with one another by the water-gas reaction:

$$CO+H_2O=CO_2+H_2$$

and therefore the equilibrium constants for every equation can be derived from the following relations:

$$(5,\ 23)\qquad \left\{\begin{aligned} &\log\frac{(H_2O)(CO)}{(CO_2)(H_2)}=k_1,\\ &\log\frac{(CO)(H_2)^3}{(CH_4)(H_2O)}=k_2,\\ &\log\frac{(CO_2)(H_2)^4}{(CH_4)(H_2O)^2}=k_3,\\ &\log\frac{(CO)^2(H_2)^2}{(CH_4)(CO_2)}=k_4,\end{aligned}\right.$$

we evidently have

$$k_3=k_2-k_1,\quad k_4=k_2+k_1.$$

Knowing two of the constants k then we can immediately determine the other two.

As in the previous case, Pier's results have been used for the calculation of the specific heats of reacting gases so that they could be used up to high temperatures. Here the specific heat of methane was taken as

$$(5,\ 24)\qquad C_p=9.090+0.00106T^*$$

from the experimental results of carbon, hydrogen and methane equilibrium.

Now taking the equation (5, 20) the heat evolution of the methane reaction at 15° C. is

$$(5,\ 25)\qquad CO+3H_2=CH_4+H_2O+50400\text{ cal.}$$

and the algebraic sum of the specific heats of these gases is

$$(5,\ 26)\qquad \Sigma\nu_1(C_p)_1=9.105+0.00136T+0.655\times10^{-6}T^2-0.8\times10^{-9}T^3.$$

*Assuming the equation of methane-hydrogen equilibrium to be

$$\log\frac{(CH_4)}{(H_2)^2}=\frac{4050}{T}-a\log T-bT+4617$$

the values a and b as calculated from the results of Pring's experiments (Jour. of Chem. Society, vol. 97, p. 509) are

$$a=3.04,\quad b=0.00041;$$

we therefore have

$$Q=Q_0+6.01T+0.001872T^2.$$

Subtracting the specific heats of hydrogen and carbon, for the specific heat of methane, CH_4, we have finally

$$9.090+0.00106T.$$

Therefore the heat evolution at absolute zero is given by

$$Q_0 = 50{,}400 - 2700 = 47{,}700 \text{ cal.}$$

and the equation of equilibrium is

$$\ln K = \frac{1}{R}\left[-\frac{47700}{T} + 9.105\ln T + 0.00068T + 0.109\times10^{-6}T^2 - 0.067\times10^{-9}T^3\right] + \Sigma\nu_1 i_1.$$

Transforming this into common logarithms we have

$$\log K = -\frac{10436}{T} + 4.55\log T + 0.000149T + 0.024\times10^{-6}T^2 - 0.015\times10^{-9}T^3 + \Sigma\nu_1\mathbf{C}_1. \tag{5, 27}$$

The constant **C** can be determined by experiment or by Nernst's theorem.

The equation given by Poltitzer* is

$$\log K = -\frac{10860}{T} + 3.5\ln T + 1.9. \tag{5, 28}$$

We here have for the chemical constant

$$\Sigma\nu_1\mathbf{C}_1 = 3.5\ (CO) + 3\times1.6\ (H_2) - 2.8\ (CH_4) - 3.6\ (H_2O) = 1.9.$$

Owing to the lack of experimental results, it is better to compare this calculation with the experimental values in the following two reactions

$$\begin{aligned} 2CO &= C + CO_2 + 38{,}350, \\ C + 2H_2 &= CH_4 + 21{,}700. \end{aligned} \tag{5, 29}$$

The former equation of equilibrium has been tested between the temperatures of 445° C. and 1473° C by Boudouard†, J. K. Clement‡ and Rhead and Wheeler§ and the latter by von Wartenberg|| and Mayer and Altmayer¶ at temperatures from 600° C. to 900° C. Their equations of equilibrium are respectively

*_Die Berechnung chemischer Affinitäten_, 1912, S. 98.
†O. Boudouard, Ann. Chim. Phys. [7], 24, p. 5, 1901.
‡J. K. Clement, Univ of Illinois, Bull No. 30, 1909; see Chem. Abstr. 4, p. 2747, 1910.
§T. F. E. Rhead and R. V. Wheeler, Jour. Chem. Society 97, p. 2178, 1910.
||H. von Wartenberg, Z. f. Phys. Chem. Bd. 61, S. 370, 1907 and Bd. 63, S. 269 1908.
¶M. Mayer und v. Altmayer, Ber. d. d. Chem. Ges. Bd. 40, S. 2134, 1907.

(5, 30) $$\log \frac{(CO)^2}{(CO_2)} = -\frac{8200}{T} + 1.75 \log T - 0.0006T + 3.8,$$

(5, 31) $$\log \frac{(H_2)^2}{(CH_4)} = -\frac{4480}{T} + 1.75 \log T + 5.5 \times 10^{-4}T + 0.4,$$

Combining these two equations we have a reaction of the form

$$2CO + 2H_2 = CH_4 + CO_2$$

and its equation of equilibrium is

(5, 32) $$\log \frac{(CO)^2(H_2)^2}{(CH_4)(CO_2)} = -\frac{12680}{T} + 3.50 \log T - 0.0005T + 4.2.$$

The value of the equilibrium constant of the equation (5, 27) is simply the difference between this value and the water-gas equilibrium. A comparison of this value with the above equation (5, 28) is given in Table IV_1, which takes **C** $= -1.8$.

TABLE IV_1

T	$\log \frac{(CO)^2(H_2)^2}{(CH_4)(CO_2)}$ Calculated from (5, 32)	k_2 Calculated from $k_4 - k_1$	k_2 Calculated from (5, 28)	k_2 Calculated from (5, 27) taking **C** $= -1.8$
500	−11.739	−9.527	−10.37	−10.315
800	− 2.580	−1.883	− 1.505	− 1.529
1000	+ 1.970	2.174	+ 1.540	+ 1.573
2000	8.314	7.659	+ 8.024	+ 8.267
3000			10.46	10.790
4000			11.79	11.976

From the equation in question we see that the formation of methane occurs at temperatures below 1000° C. and as illustrated later, the methane which is found in explosion gas in a closed vessel is formed during the period of cooling after explosion. The equilibrium constants of the other two equations of (6, 2) are given by the relations

$$\log \frac{(CO_2)(H_2)^4}{(CH_4)(H_2O)^2} = k_3 = k_2 - k_1,$$

$$\log \frac{(CO)^2(H_2)^2}{(CH_4)(CO_2)} = k_4 = k_2 + k_1,$$

and Tables IV_1 and IV_2 and Figure 2 show the results of these calculated values.

TABLE IV_2

Temp. abs.		$\log\frac{(H_2O)(CO)}{(CO_2)(H_2)}$	$\log\frac{(CO)(H_2)^3}{(CH_4)(H_2O)}$	$\log\frac{(CO_2)(H_2)^4}{(CH_4)(H_2O)^2}$	$\log\frac{(CO)^2(H_2)^2}{(CH_4)(CO_2)}$
300	log *k*	−5.032	−25.181	−20.139	−30.213
	k	9.290×10^{-6}	6.592×10^{-26}	7.261×10^{-21}	6.124×10^{-31}
500	log *k*	−2.178	−10.348	−8.170	−12.526
	k	6.637×10^{-3}	4.487×10^{-11}	6.761×10^{-9}	2.979×10^{-13}
800	log *k*	−0.6419	−1.719	−1.0771	−2.3609
	k	0.2281	1.910×10^{-2}	8.373×10^{-2}	4.356×10^{-3}
1000	log *k*	−0.170	1.5809	+1.7509	+1.4109
	k	0.6761	38.10	56.35	25.75
1500	log *k*	+0.3867	5.8721	+5.4854	6.2588
	k	2.436	$7.4490\times10^{+5}$	$3.058\times10^{+5}$	$1.814\times10^{+6}$
2000	log *k*	+0.6298	8.184	7.5542	8.8138
	k	4.264	$1.528\times10^{+8}$	$3.583\times10^{+7}$	$6.513\times10^{+8}$
2500	log *k*	0.77144	9.7593	8.98786	10.53074
	k	5.907	$5.746\times10^{+9}$	$5.725\times10^{+8}$	$3.394\times10^{+10}$
3000	log *k*	0.8832	10.730	9.8468	11.6132
	k	7.642	$5.370\times10^{+10}$	$7.028\times10^{+9}$	$4.104\times10^{+11}$
4000	log *k*	1.1427	12.112	10.9692	13.2547
	k	13.89	$1.294\times10^{+12}$	$9.315\times10^{+10}$	$1.798\times10^{+13}$

(vi) *Equilibrium constants of the resulting gas from explosives.*

The equilibrium constants which have been discussed are for constant pressure but they can be applied as well to the reactions of gases at constant volume, such as explosions in closed vessels, because the value at the point of equilibrium is the same in either case.

However a correction for the high pressure of the gases must be made to these results.

(1) *Correction for high pressure.*

Although the previous equations are satisfactorily applied to an ideal gas mixture, it is necessary to make some corrections for the gas at high pressure, such as the gases resulting from explosives.

(*a*) *Correction due to specific heat.*

The relation of the specific heat to pressure or volume can be derived from the thermodynamic relations

$$\left(\frac{\partial C_v}{\partial V}\right)_T = T\left(\frac{\partial^2 P}{\partial T^2}\right)_V,$$

$$\left(\frac{\partial C_p}{\partial P}\right)_T = T\left(\frac{\partial^2 V}{\partial T^2}\right)_P.$$

The experimental results are very few and the calculations* from van der Waals' equation give

$$C_v=f(T),\ C_p=(C_p)_{P=1}+\frac{2ab(P-1)}{RT^2}-\frac{3ab(P^2-1)}{R^2T^3}.$$

As the temperature increases the 2nd and 3rd terms decrease and for high values of the temperature can be neglected for practical purposes, so that

$$C_p=(C_p)_{P=1}.$$

(*b*) *Specific heat of the gaseous mixture.*

The specific heats of actual gaseous mixtures by (3, 15) and (3, 18) are

$$C_v=\Sigma n_1(C_v)_1+\frac{R}{V}\Sigma(b_{12}n_1n_2),$$

$$C_p=\Sigma n_1(C_v)_1-\frac{3P}{NRT^2}\Sigma(a_{12}n_1n_2).$$

Similarly

$$C_p=\Sigma n_1(C_p)_{P=1}+2\frac{(P-1)}{RT^2}\Sigma n_1a_1b_1-3\frac{P^2-1}{R^2T^3}\Sigma n_1a_1b_1-\frac{3P}{NRT^2}\Sigma(a_{12}n_1n_2),$$

but for high temperature

$$C_p=\Sigma n_1(C_p)_{P=1}.$$

(*c*) *Correction due to* Δ *and* ∇.

These terms Δ and ∇ in the equation of equilibrium make the calculation of the equilibrium constant difficult since they contain the factors $n_1, n_2, \ldots, n_i$, as shown in equations (4, 8) and (4, 16). But it is possible to obtain the constant by neglecting ∇ and Δ at first and calculating K and then inserting $n_1, n_2, \ldots$ in Δ and ∇ and again calculating K.

The terms

$$\Sigma\nu_i\frac{\partial\Delta}{\partial n_i},\quad \Sigma\nu_i\frac{\partial\nabla}{\partial n_i}$$

are the differences of the sums Σn_1b_1 and Σn_1a_1 for CO and H_2O on the one hand and for H_2 and CO_2 on the other hand in water-gas reactions.

*N. Yamaga: Jour. of Soc. of Ordnance and Explosives, vol. 16, p. 1.

Calculating a and b as well as Σb_1 and Σa_1 we have

	a	sum	b	sum
CO	275		1683	
	>	1364	>	3045
H_2O	1089		1362	
N_2	386		977	
	>	1105	>	2889
CO_2	719		1912	

$$\Sigma(a_1)=259,\quad \Sigma(b_1)=156.$$

As regards the mutual coefficients a_{12}, b_{12}, it is better to neglect these ambiguous terms, on account of the lack of experimental values on the one hand and because they always appear as a difference of terms in the chemical equations on the other hand. Ultimately, therefore, these correction terms can be neglected without appreciable error so that

$$\Delta=\nabla=0.$$

In considering the chemical equilibrium of the reaction of explosives we can, therefore, treat it as an ideal gas reaction owing to the high temperature.

VI

On the Equilibrium of Explosive Gases

(1) *Water-gas Reaction.*

A. Noble and F. Abel* as well as E. Sarrau and Vieille† have discovered that the gaseous compositions resulting from explosives undergo changes owing to the density of loading. Later Noble and Abel have made many researches‡ on M. D. Cordite, Cordite M.K. 1. and Nitrocellulose powder.

According to their results, carbon dioxide increases with the density of loading, while carbon monoxide and hydrogen decrease as shown by the curves in Figures 3_1, 3_2 and 3_3.

Water-gas reaction, however, is a reaction without condensation and consequently must be independent of pressure. This contradicts the facts, and therefore there must be some other action causing it.

Take, for another example, the reaction in methane formation; the quantity of methane is great when the density of loading is high. Of course the quantity of methane should be increased with the pressure, as the reaction in methane

*Proc. Roy. Soc., London, Ser. A, vol. 56, p. 205 (1894).
†Mém. Poudres et Salpêtres II, 126, 337 (1884-85).
‡Phil. Trans. Roy. Soc. London, Ser. A, vol. 205, p. 201 (1906); Ser. A, vol. 206, p. 453 (1906).

formation is one of reaction with condensation, but it is very doubtful whether methane gas exists at the moment of explosion because the methane is only stable below the temperature 1500°* as shown in Tables IV, Section V.

Table V shows the equilibrium constant K of water-gas reaction and the temperature corresponding to K for 3 kinds of powder. These calculations have been made from the concentration of gases relating to water-gas reaction from Noble's analytical composition† of the products of explosion under various densities of loading.

TABLE V

	Cordite		M.D. Cordite		Nitrocellulose	
Δ	k.	temp.	k.	temp.	k.	temp.
0.05	2.134	1430	2.25	1450	2.187	1440
0.10	1.664	1310	1.82	1350	1.71	1325
0.15	1.562	1290	1.50	1275	1.64	1310
0.20	1.23	1200	1.33	1230	1.68	1320
0.22	..		..		1.54	1280
0.30	1.03	1140	1.11	1165	1.38	1240
0.40	1.09	1160	1.13	1170	1.35	1235
0.45	..	..	1.12	1167	1.26	1210
0.50	1.02	1140	..	..	..	..

Looking at the table, we see that the temperature under which the gases are in equilibrium is lower when the density of loading is high. This means that the velocity of cooling of the gas is so slow due to its high density that it has sufficient time to change its composition according to the law of equilibrium until a certain temperature is reached.

Take, for another example, Noble's experimental results† for M. D. Cordite. Owing to the difficulties of calculation due to the complexity of the reaction for the larger volume percentage of methane, the equilibrium constant and the corresponding temperature have only been calculated up to a loading density $\Delta = 0.25$. Table VI_1 gives the calculated results, which show that the equilibrium constant is smaller the higher the density of loading; accordingly the equilibrium temperature is lower owing to the slowness of cooling.

TABLE VI_1

Δ	0.05	0.10	0.15	0.20	0.25
K	2.25	1.82	1.50	1.335	1.233
T	1450°	1350°	1275°	1230°	1200°
CO_2	14.85	16.45	18.35	20.30	22.25
CO	34.87	33.95	32.40	30.50	28.45
H_2O	18.15	17.00	16.20	16.00	16.05
H_2	18.95	19.20	19.00	18.00	16.65
N_2	12.89	12.57	12.45	12.50	12.65
CH_4	0.29	0.83	1.60	2.70	3.95

The quantity of methane given in this table has been recalculated by means

*Also see Table IX, Section VI.

†Noble A., *loc. cit.*, also see Poppenberg u. Stephan: Z. f. Schiess u. Spreng. Bd. 4, S. 282, 1906.

of Table IX from the equilibrium constants of the following three equations for methane formation:

$$CO + 3H_2 = CH_4 + H_2O,$$

$$CO_2 + 4H_2 = CH_4 + 2H_2O,$$

$$2CO + 2H_2 = CH_4 + CO_2.$$

Corrections have been made to the gases corresponding to the above three reactions. Table VI_2 and the full line in Figure 4 represent these results.

TABLE VI_2

Δ	0.05	0.10	0.15	0.20	0.25
CO_2	14.74	16.23	17.65	19.02	20.20
CO	35.01	34.15	33.00	31.70	30.40
H_2O	17.78	16.00	14.30	12.94	11.65
H_2	19.68	21.25	22.95	24.40	26.00
N_2	12.79	12.37	12.10	11.94	11.75

Consider the water-gas reaction containing nitrogen as a like constituent with that of the explosives. The composition* of the explosive used in this calculation† is

Carbon $= A_c = 21.50$ mol. per cent.
Oxygen $= A_o = 35.30$ " " " ,
Hydrogen $= 2A_h = 32.20$ " " " ,
Nitrogen $= 2A_n = 11.00$ " " " .

To find out the percentage of each gas, let x be the number of mols. of carbon dioxide, then

$$(6, 1) \qquad \begin{cases} (CO) = A_c - x, \\ (H_2O) = A_o - A_c - x, \\ (H_2) = A_h - A_o + A_c + x \end{cases}$$

and

$$(6, 2) \qquad K = \frac{(CO)(H_2O)}{(CO_2)(H_2)} = \frac{(21.5-x)(13.8-x)}{x(18.4+x)}.$$

The concentration of CO_2, x, can be easily calculated by solving this quadratic equation giving the numerical value to K which is of course definite for a certain temperature. Having found x, the percentage composition of the other gases can be readily calculated.

Table VII and Figure 4 (dotted lines) represent the percentage of each gas in the mixture between the temperatures 1200° and 1500° C. and we can see this line is very close to the curve plotted from Noble's experimental results (Table VI_1).

*Calculated from gas analysis.
†Method of working, see Section VII.

TABLE VII

Δ	1200	1300	1400	1500
CO_2	19.7	16.2	15.1	14.4
CO	31.2	35.0	35.95	36.5
H_2O	13.0	16.4	17.70	18.2
H_2	25.1	21.4	20.25	19.9
N_2	11.0	11.0	11.0	11.0

The composition of the gaseous mixture does not change according to the pressure but varies with the rate of cooling.

Again take, as another example, the author's experiments on calorimetric determination of powder with Berthelot's bomb,

the volume of bomb = 280 c.c.,
the weight of powder, M.D. Cordite = 2 grams.

COMPOSITION OF GASES IN VOLUME PERCENTAGES.

Experiment	Δ	CO_2	CO	H_2	H_2O	CH_4	N_2	K
1	0.007	111.6	334.0	119.8	261.2	1.1	117.0	6.5
2	0.007	114.8	338.6	117.6	225.6	0.8	120.0	5.7

The value of K calculated from these results is about 6.0 and this corresponds to a temperature of equilibrium of about 2500°; that is to say for a small value of the density of loading, K is increased considerably. Poppenberg and Stephan* also confirmed these facts by their experiments.

Thus, the conclusion is that the composition of the gaseous products of explosives does not depend on the density of loading or the pressure of explosion but on the rate of cooling. Or, in other words, the temperature and the equilibrium constant k ought to remain the same if no heat is lost and must be independent of the density of loading.

It is a big mistake to take the constituents given by the gas analysis of the products of explosion after cooling as the constituents at the instant of explosion.

The measurement of the pressure of combustion of powder and its velocity of cooling in the closed vessel of Vieille by means of the oscillograph is now proceeding.

(2) *Methane Formation.*

The methane formation from gaseous products of explosion was first studied by Noble and Abel†. Their conclusive opinion was that the quantity of methane increased with the pressure of explosion. Noble's results‡ show that the volume

*loc. cit.

†Proc. Roy. Soc. London, Ser. A, vol. 56, p. 205 (1894).

‡Phil. Trans. Roy. Soc. London, Ser. A, vol. 205, p. 201 (1906), and Ser. A, vol. 206, p. 453 (1906).

percentage of methane is 0.3 at loading density $\Delta=0.05$ and that it increases up to about 10 per cent. when $\Delta=0.5$*. The reaction of methane formation may be considered from the following three chemical equations as already stated in Section V:

$$CO+3H_2=CH_4+H_2O,$$

$$CO_2+4H_2=CH_4+2H_2O,$$

$$2CO+2H_2=CH_4+CO_2.$$

These three equations are not independent of one anotber but are all related in consequence of the water-gas reaction*. The equilibrium constants K of these three equations are very large at high temperatures and have only small values under 1000° C. as shown in Table IV. On the other hand the velocity of formation of methane becomes small at this lower temperature and only a small quantity of methane is present in the resulting gases. We shall calculate the quantity of methane from the value of K of these three reactions for M. D. Cordite.

The quantity of methane y can be calculated from the concentration of water-gas at temperature T provided the quantity of methane is not very large.

Let the water-gas equilibrium k_1† be, from (7, 12), Section VII,

$$k_1=\frac{(21.6-x)(15-x)}{x(x+0.8)}=\frac{(CO)(H_2O)}{(CO_2)(H_2)};$$

then

$$k_2=\frac{(CO-y)(H_2-y)^3}{y(H_2O-y)},$$

(6, 3)
$$k_3=\frac{(CO_2-y_2)(H_2-4y_2)^4}{y_2(HO+2y_2)},$$

$$k_4=\frac{(CO-2y)^2(H_2-2y)^2}{y(CO_2+y)}.$$

For high temperature y is small and we can simply take

(6, 4)
$$k_2=\frac{(CO)(H)^3}{y(H_2O)}$$

without appreciable error.

Below 1000°, y becomes large and the calculations should be made by (6, 3). The values of the k's thus found should be multiplied by a pressure factor $p^{-\Sigma\nu}=p^{-2}$ for each expression.

Table VIII shows the partial pressure of methane at 2000° absolute in the case of M.D. Cordite.

*See Curve, Fig. 6.
†See (5, 23).

TABLE VIII

COMPOSITION OF GASES AT 2000° ABSOLUTE

	Molecular composition	Partial pressure	
		under 1 atm.	under 3000 atm.
CO_2	5.475	.130	390
CO	16.125	.380	1.140
H_2O	9.525	.225	680
H_2	6.275	.146	440
N_2	5.1	.120	360
Total.	42.50	1.000	3.000

The formation of methane under a pressure of 3000 atmospheres is given by

$$P_{CH_4}=\frac{1140\times(440)^3}{1.5\times10^8\times680}=0.95 \text{ atm.},$$

or

$$\frac{0.95\times100}{3000}=0.03 \text{ per cent.}$$

In the same manner Table IX gives the values of y_1, y_2, and y_3 from which the percentage volume of methane at a pressure of 1 atmosphere as well as 3000 atmospheres at various temperatures from $T=3000°$ to $T=1000°$ can be calculated.

TABLE IX

(VOLUME %)

Temp.	y_1	y_2	y_3	1 atm.	3000 atm.
3000	2.4×10^{-12}	2.3×10^{-12}	2.5×10^{-12}	5.9×10^{-11}	5.0×10^{-4}
2500	2.9×10^{-11}	4.7×10^{-11}	3.0×10^{-11}	1.0×10^{-10}	9.0×10^{-3}
2000	1.5×10^{-9}	1.4×10^{-9}		3.3×10^{-9}	0.03
1500	5.3×10^{-7}	5.1×10^{-7}	5.7×10^{-7}	1.3×10^{-6}	11.0
1250	6.7×10^{-5}	7.2×10^{-5}	9.95×10^{-5}	2.3×10^{-4}	
1000	0.029	0.019	0.043	0.10	

Figure 5 shows the logarithm of the k's and even under 3000 atmospheres pressure the methane formation is negligible at and above 2000° C.

VII

CALCULATION OF THREE ELEMENTS K, x, AND T

Previously the calculation of the fundamental constants of smokeless powder such as force of explosive, temperature and heat of explosion, and the composition and volume of the gases, were generally made from the results of the analysis of gaseous products after cooling, but as I have shown in Section VI the gaseous products of explosion undergo changes in the cooling stage; the results as can be seen from the analysis of gases after cooling give the composition corresponding to the temperatures 1100°-1400° as shown in Table V.

To find out the composition of gases at the moment of explosion, I have made calculations in three steps as follows:

(1) *Relation between x and K.*

Modern smokeless powder generally contains four chemical elements: carbon, oxygen, hydrogen, and nitrogen. Besides these, calcium and sodium salts may be found in very small quantities.

Let A_c, A_o, $2A_n$ and $2A_h$ be the percentage number of atoms carbon, oxygen, nitrogen and hydrogen respectively in unit weight of explosive. The gaseous products of explosion are carbon-dioxide, carbon-monoxide, water, hydrogen and nitrogen and if we let the number of mols. of each gas which were contained in the original explosive be,

$$\begin{array}{ll} CO_2 & x \text{ mols.}, \\ CO & c \quad \text{''}, \\ H_2O & w \quad \text{''}, \\ H_2 & h \quad \text{''}, \\ N_2 & n \quad \text{''}. \end{array} \tag{7, 1}$$

Then we have 4 relations,

$$\begin{aligned} c &= A_c - x, \\ w &= A_o - A_c - x, \\ h &= A_h - (A_o - A_c - x) = A_h - w, \\ n &= A_n; \end{aligned} \tag{7, 2}$$

the equilibrium constant K can be written down in terms of the components:

$$K = \frac{(CO)(H_2O)}{(CO_2)(H_2)} = \frac{cw}{xh} = \frac{(A_c - x)(A_o - A_c - x)}{x[A_h - (A_o - A_c - x)]}. \tag{7, 3}$$

This equation expresses the value of K in terms of x, concentration of carbondioxide. Here K is a function of temperature as shown in equation ((9)), Section V; let it be $K(T)$.

Then

$$\frac{(A_c - x)(A_o - A_c - x)}{x(A_h - A_o + A_c + x)} = K(T); \tag{7, 4}$$

this shows that x is a function of temperature.

(2) *Calculation of heat of reaction, Q.*

Let

q = heat of formation of powder,
Q = heat of explosion of powder,
P = sum of the heats of formation of products of decomposition.

Taking the heats of formation of

carbon-dioxide	q_{CO_2}	$= 94{,}310$,
carbon-monoxide	q_{CO}	$= 26{,}100$,
steam	q_{H_2O}	$= 58{,}100$ (as gas),

we have

$$\begin{aligned} Q &= P - q \\ &= xq_{CO_2} + Cq_{CO} + wq_{H_2O} - q \\ &= xq_{CO_2} + (A_c - x)q_{CO} + (A_o - A_c - x)q_{H_2O} \\ &= x[q_{CO_2} - q_{CO} - q_{H_2O}] + A_c q_{CO} + (A_o - A_c)q_{H_2O} - q \\ &= 10{,}110x + A_c q_{CO} + (A_o - A_c)q_{H_2O} - q. \end{aligned} \tag{7, 5}$$

(3) *Relation of specific heats of gaseous products.*

Let

C_{CO_2} be specific heat of carbon-dioxide,
C_h " " hydrogen,
C_d " " diatomic gas except hydrogen,
C_w " " steam.

The combustion of the powder produces x mols. of carbon-dioxide, C mols. of carbon-monoxide, w mols. of water, h mols. of hydrogen and A_n mols. of nitrogen. Let the quantity of heat evolved in the reaction be Q; then the relation between Q and temperature is:

(7,6)

$$Q=\int(xC_{CO_2}+cC_d+wC_w+hC_h+nC_d)dT$$

$$=\int[xC_{CO_2}+(A_o-A_c-x)C_w+(A_c+A_n-x)C_d+(A_h-A_o+A_c+x)C_h]dT$$

$$=\int[x(C_{CO_2}-C_w-C_d+C_h)+(A_o-A_c)C_w+(A_c+A_n)C_d+(A_h-A_o+A_c)C_n]dT.$$

If the specific heat of a diatomic gas be taken as given by an equation of the first degree in T as shown in Section V (5, 6−5, 7) and that of a triatomic gas be also taken as given by an equation of the first degree* then approximately:

(7, 7)

$$C_{CO_2}=7.31+2.44\times10^{-3}T,$$
$$C_w=2.56+4.68\times10^{-3}T,$$
$$C_h=4.454+0.0009T,$$
$$C_d=4.654+0.0009T,$$

whence

$$C_{CO_2}-C_w-C_d+C_h=4.55-2.24\times10^{-3}T.$$

We have the relation

(7, 8)

$$Q=\{4.55T-1.12\times10^{-3}T^2-1220\}x$$
$$+(A_o-A_c)\ [2.34\times10^{-3}T^2+2.56T-931]$$
$$+(A_c-A_n)\ [0.45\times10^{-3}T^2+4.654T-1412]$$
$$+(A_h-A_o+A_c)\ [0.45\times10^{-3}T^2+4.700T-1355].$$

(4) *Relation of Q, x and T.*

The following three equations (7,4), (7,5) and (7,8),

I $$\frac{(A_c-x)(A_o-A_c-x)}{x(A_h-A_o+A_c+x)}=K(T),$$

II $$Q=10{,}110x+A_c q_{CO}+(A_o-A_c)q_{H_2O}-q,$$

III $$Q=(4.55T-1.12\times10^{-3}T^2-1220)x+a+bT+cT^2),$$

contain three unknown variables Q, T and x, solving for which we would

*Blom: Z. f. Schiess u. Spreng. (1916), S. 219.

theoretically be able to calculate the composition of the gaseous products, temperature of explosion and heat of decomposition not affected by cooling. But owing to the difficulties of solving these equations simultaneously, it is better to find the values graphically. By eliminating Q from equations II and III, x can be expressed by an equation of the second degree in T; the graphical representation can now be easily made (see Figure 6) by taking x and T as axes.

Next trace a curve (see Figure 7) for the equation (7, 3),

$$K=\frac{(A_c-x)(A_o-A_c-x)}{x(A_h-A_o+A_c+x)}$$

taking x and K as axes.

From these two curves plot a curve (see Figure 8) taking T and K as axes for the same corresponding values of x in each $(x:T)$ and $(x:K)$ figures. Next plot equation (9) in Section V on the new (T, K) graph, then the coordinates of the point of intersection of these two curves, are the values of K and x required.

The following are examples of the calculations for some modern typical smokeless powders: M.D. Cordite, M.K. 1. Cordite and nitrocellulose powder.

(5) *M.D. Cordite.*

English M.D. Cordite has the following composition:

Guncotton............	65.0 wt. per cent.,
Nitroglycerin..........	30.0 " ,
Mineral jelly..........	5.0 " .

The atomic constituents expressed in weight percentages are:

(7, 9)

Carbon	C............	25.40,
Oxygen	O............	57.50,
Hydrogen	H............	3.10,
Nitrogen	N...........	14.00,

or in atomic ratios:

$$C \quad \frac{25.40}{12}=2.12,$$

$$O \quad \frac{57.50}{10}=3.59,$$

$$H \quad \frac{3.10}{1.008}=3.10,$$

$$N \quad \frac{14.00}{14.01}=1.00,$$

9.81.

The numerical concentrations are:

$$A_c = C = \frac{2.12}{9.81}\times 100 = 21.60\%, \quad A_c = 21.60,$$

$$A_o = O = \frac{3.59}{9.81}\times 100 = 36.60\%, \quad A_o = 36.60,$$

(7, 10)

$$2A_h = H = \frac{3.10}{9.81}\times 100 = 31.60\%, \quad A_h = 15.80,$$

$$2A_n = N = \frac{1}{9.81}\times 100 = 10.20\%, \quad A_n = 5.10,$$

100.00.

Here

$$A_o - A_c = 15.00, \quad A_h - (A_o - A_c) = 0.8,$$

(7, 11) $$K = \frac{(21.6-x)(15.0-x)}{x(0.8+x)}.$$

Next the heat of formation of M.D. Cordite, q, is by experiment

$$q = 520 \text{ cal./gram.}$$

To calculate the total heat of formation q for this number of mols. of powder we then have

$$\begin{aligned} A_c\times 12.00 &= 21.60\times 12.00 = 259.20,\\ A_o\times 16.00 &= 36.60\times 16.00 = 585.60,\\ 2A_h\times 1.008 &= 31.60\times 1.008 = 31.80,\\ 2A_n\times 14.01 &= 10.20\times 14.01 = 142.90,\\ & \qquad 1019.50 \text{ g}, \end{aligned}$$

and therefore

$$q = 520\times 1{,}019.5 = 530{,}140;$$

accordingly

(7, 12) $$\begin{aligned} Q &= 10{,}110x + 26{,}100\times 21.6 + 58{,}100\times 15.0 - 520\times 1{,}019.5\\ &= 10{,}110x + 563{,}760 + 871{,}500 - 530{,}140\\ &= 10{,}110x + 905{,}120. \end{aligned}$$

Equations III,

(7, 13) $$\begin{aligned} Q &= \int_{288}^{T} x(4.55 - 2.24\times 10^{-3}T)\,dT\\ &+ \int_{288}^{T} \{15(2.56 + 4.68\times 10^{-3}T) + 26.7\,(4.654 + 0.9\times 10^{-3}T)\\ &\quad + 0.8\,(4.454 + 0.9\times 10^{-3}T)\}\,dT \end{aligned}$$

$$= \int_{288}^{T} x(4.55-2.24\times10^{-3}T)dT$$
$$+\ (38.4+124.26+3.56)\ (T-288)$$
$$+\ \tfrac{1}{2}\ (70.20+24.03+0.72)\times10^{-3}(T^2-\overline{288}^2)$$
$$=\quad 4.55xT-0.00114xT^2-1215.8x$$
$$+166.2T+0.04748T^2-36.657$$
$$=\quad (4.55x+166.2)T+(0.04748-0.00114x)T^2-1215.8x-36,657$$
$$Q=\quad 10,110x+905,120$$
$$=\quad (4.55x+166.2)T+(0.04748-0.00114x)T^2-1215.8x-36,657$$

or

(7, 14) $(0.04748-0.00114x)T^2+(166.2+4.55x)T-941,777-11,326x=0.$

Figure 6 represents the graph of these equations. Similarly Figure 7 shows the equation (7, 11); these two figures plot Figure 8 taking T and K as axes for the same corresponding value of x, as already stated. On this Figure 8, plot the equations ((9)) and ((8)); the intersections of these two curves give respectively,

	From ((9))	From ((8))
T	3115.0° abs.,	3111.0° abs.,
K	8.10,	9.0.

The temperature of explosion is nearly the same by either equation so that x can be easily calculated from Figure 6 or Figure 7 for corresponding values of T and K:

from ((8)) $x=4.20$, from ((9)) $x=5.50$.

Thus we have found three elements x, T and K which are necessary for the calculation of explosion data.

To find how the composition of the gaseous products of explosion changes due to cooling, it is necessary to find the concentration of carbon-dioxide and other gases for various temperatures by aid of equation (7, 3). Table X and Figure 9 show the change in per cent. of these gases with the lowering of temperature (below 1500°, however, the formation of methane must be taken into account).

TABLE X

		CO_2	CO	H_2O	H_2	N_2
T	K	%	%	%	%	%
3325	9.24	9.705	41.02	25.59	11.59	11.99
3200	8.55	10.02	40.74	25.23	11.89	11.99
3115	8.14	10.19	40.55	25.06	12.05	11.99
3000	7.64	10.40	40.36	24.89	12.27	11.99
2500	5.90	11.46	39.36	23.82	13.34	11.99
2325	5.34	11.99	38.82	23.23	13.87	11.99
2000	4.26	12.85	37.93	22.39	14.76	11.99
1500	2.43	15.52	35.24	19.77	17.38	11.99
1000	0.63	22.39	28.38	12.88	24.27	11.99

(6) *M.K.*1, *Cordite.*

The constituents of Cordite are,

Guncotton	37.0,
Nitroglycerine.............	58.0,
Mineral jelly..............	5.0,

and the atomic composition in percentage by weight is

Carbon	C..............	23.00%,
Oxygen	O	58.96%,
Hydrogen	H	3.04%,
Nitrogen	N	15.00%.

The numerical concentrations are therefore

$$A_c=\frac{23.00}{12}\times\frac{100}{9.67}=19.60,\qquad A_c=19.60,$$

$$A_o=\frac{55.96}{16}\times\frac{100}{9.67}=38.00,\qquad A_o=38.00,$$

$$2A_h=\frac{3.04}{1.008}\times\frac{100}{9.67}=31.00,\qquad A_h=15.50,$$

$$2A_n=\frac{15.00}{14.01}\times\frac{100}{9.67}=11.40,\qquad A_n=5.70,$$

and

$$K=\frac{(19.6-x)(18.4-x)}{x(x-2.9)}.$$

The heat of formation of M.K.1, Cordite is

$$q=569 \text{ cal.}$$

We have

$$\begin{aligned} A_c\ &\times 12=19.60\times 12 &&=235,\\ A_o\ &\times 16=38.00\times 16 &&=608,\\ 2A_h\ &\times 1.008=31.00\times 1.008 &&=\ 31,\\ 2A_n\ &\times 14.01=11.4\ \times 14.01 &&=160,\\ \hline & && 1.034, \end{aligned}$$

and therefore

$$q=569\times 1{,}034=588{,}350;$$

accordingly

$$\begin{aligned} Q&=10{,}110x+26{,}100\ A_c+58{,}100\ (A_o-A_c)-q\\ &=10{,}110x+511{,}560+1{,}069{,}040-588{,}350\\ &=10{,}110x+992{,}250. \end{aligned}$$

Equation III gives

$$Q=\int_{288}^{T} x(4.55-2.24\times10^{-3}T)dT$$

$$+\int_{288}^{T}\{18.4\ (2.56+4.68\times10^{-3}T)+25.3\ (4.654+0.9\times10^{-3}T)$$

$$-2.9\ (4.454+0.9\times10^{-3}T)\}dT$$

$$=4.55xT-1.12\times10^{-3}xT^2-1215.8x+151.9T+0.0531T^2-48{,}146.$$

Combining equations II and III,

$$10{,}110x+992{,}254$$

$$=4.55xT-1.12\times10^{-3}xT^2-1215.8x+151.9T+0.0531T^2-48{,}146,$$

$$(0.0531-0.00112x)T^2+(4.55x+151.9)T-(1{,}040{,}400+11{,}325.8x)=0.$$

Figure 6 shows the results of the combination. In the same manner as in the previous example two values each of K, T and x can be found from Figure 8 corresponding to the equations ((8)) and ((9)). These values are

	From ((8))	From ((9))
T	3320.0°,	3323.5°,
K	10.55,	9.2,
x	5.75,	5.95.

(7) *Nitrocellulose or French B Powder.*

Nitrocellulose	94 −95 %,
Diphenylamin	1.5− 2.0%,
Solvent	2.5− 4.5%,
Carbon	C =27.91% by weight,
Oxygen	O =58.96 " "
Hydrogen	H = 3.18 " "
Nitrogen	N =11.90 " " .

Numerical concentration

$$A_c=23.50,$$
$$A_o=36.04,$$
$$A_h=15.95,$$
$$A_n=4.28,$$

$$K=\frac{(23.50-x)(12.54-x)}{x(x+3.41)}.$$

Heat of formation of nitrocellulose powder

$$q = 695 \text{ cal./gram.}$$

and

$$\begin{aligned} A_c \times 12 &= 282.0, \\ A_o \times 16 &= 576.6, \\ 2A_h \times 1.008 &= 32.15, \\ 2A_n \times 14.01 &= 119.92, \\ \hline & 1010.7, \end{aligned}$$

$$q = 695 \times 1010.7 = 702440;$$

accordingly

$$\begin{aligned} Q &= 10{,}110x + 23.50 q_{CO} + 12.54 q_{H_2O} - q \\ &= 10{,}110x + 613{,}350 + 728{,}570 - 702{,}440 \\ &= 10{,}110x + 639{,}480. \end{aligned}$$

Equation III gives

$$Q = 4.55xT - 1.12 \times 10^{-3} xT^2 - 1215.8x + 176.6T + 0.0434T^2 - 54{,}460.$$

Combining equations II and III and eliminating Q we have

$$(0.0434 - 0.00112x)T^2 + (176.6 + 4.55x)T - 11{,}326x - 693{,}940 = 0.$$

Figure 6 represents this equation and as in the two previous examples, we can calculate the values T, K and x corresponding to the equations ((8)) and ((9)). They are found to be

	(8)	(9)
T	2316.0°,	2325.0°,
K	5.6,	5.3,
x	3.9,	4.0.

VIII

Conclusion and Method of Calculation

If the elementary composition and heat of formation of powder are known, the temperature of explosion T_0, the equilibrium constant K and the concentration of carbon-dioxide x can be found and the various relations, which are necessary to the theory of explosives, such as the force of explosives, the composition of gaseous products, etc., can be calculated as we have shown in Section VII.

(1) *Elementary analysis of smokeless powder.*

The methods for the elementary analysis of smokeless powder are not different from those of ordinary chemical analysis. Two methods are generally used.

(*a*) *Combustion analysis.*

This is the same as is used in the ordinary organic chemical laboratory. There are a few suggestions which might be added here; the copper gauze at the end of the combustion tube must be very long in order to completely decompose the oxy-nitrogen compounds, and it is better to mix the sample with some fine copper-dioxide powder to prevent rapid decomposition.

For the determination of nitrogen, the Dumas method is preferable to that of Lunge or Kjeldahl. Sometimes nitrogen* can be calculated from the gases produced by the explosion of powder in an atmosphere of CO_2 by a method similar to that of Dumas.

(*b*) *Analysis by the method of gas analysis.*

The explosive to be tested is fired in a Vieille closed vessel, and the resulting permanent gases (CO_2, CO, CH_4, and H_2) are calculated by Hempel's exact method of gas analysis†.

Water is estimated gravimetrically by absorbing it with calcium chloride or phosphorous pentoxide. The percentage composition of carbon, hydrogen, oxygen and nitrogen can be calculated from the results of analysis.

(2) *Heat of formation.*

The heat of formation can be found by calorimetric measurements with Berthelot's bomb using oxygen up to a pressure of 5 atmospheres in order to ensure complete oxidation.

In this case, some dissociations such as

$$CO_2 \rightleftharpoons CO + O,$$
$$H_2O \rightleftharpoons H_2 + O,$$

may occur, but these dissociations are negligible‡ under 2000° C. and the only correction is that due to the formation of nitric acid.

(3) *Calculation.*

By the previous experiments, if the elementary composition of the explosives and the heat of formation are known, we can find graphically the temperature of explosion T_0, the equilibrium constant K of the gaseous products, and the numerical concentration of carbon-dioxide. From these three fundamental quantities, we can calculate the other characteristic constants of explosives, such as the numerical concentrations of all gases at the moment of explosion, and consequently by the aid of Table I, the weight and the volume percentage of gases, the specific volume of the gases, the heat of explosion, the force of explosive, etc.

*Poppenberg u. Stephan: Z. f. Schiess u. Spreng. (1909), S. 350.
†W. Hempel: *Gasanalytische Methoden*, p. 55.
‡Nernst: *Thermodynamics and Chemistry* (1907), p. 81.

Examples M.D. Cordite

	Unit			
Heat of formation..........	g-cal/g.	q	520	Experiments.
Elementary composition.....	Wt.%	C O H N	25.40 57.50 3.10 14.00	Experiments.
Molecular concentration.....		A_c = A_o = A_h = A_n =	21.6 36.6 15.8 5.10	
Equilibrium constant		K	$\frac{(21.6-x)(15-x)}{x(x+0.8)}$	
Temperature of explosion....	abs.	T_o T_o	3111° 3115°	by ((8)) by ((9))
K......................			9.00 8.10	by ((8)) by ((9))
x......................			4.40 5.50	by ((8)) by ((9))

Percentage composition of gases by weight:

	((8))	((9))
CO_2	18.11	23.70
CO	47.75	44.20
H_2	0.99	1.25
H_2O	19.05	16.75
N_2	14.10	14.10

Percentage composition of gases by volume.

	((8))	((9))
CO_2	9.84	12.88
CO	40.95	37.90
H_2	11.78	14.84
H_2O	25.41	22.38
N_2	12.02	12.02

Total volume of gas evolved from 1 kg. of explosive:

((8)) 933.5 litres,

((9)) 933.3 " .

The heat of reaction can be calculated from equation (7, 5):

$$Q = 10{,}110x + 905{,}150.$$

From ((8))

$$Q = 930 \text{ cal./gram.}$$

From ((9))

$$Q = 943 \quad " \quad .$$

The heat of evaporation of water must be added to these figures,

((8)) $Q = 930 + 0.19 \times 600 = 1044,$

((9)) $Q = 943 + 0.168 \times 600 = 1044.$

The force of explosive f is a product of the specific volume of the powder and the temperature of explosion:

$$f = \frac{p_0 v_0}{273} T.$$

From ((8))

$$f = \frac{1.0332 \times 933.5}{273} \times 3111 = 10{,}993 \; kg./cm^2.$$

From ((9))

$$f = \frac{1.0332 \times 933.3}{273} \times 3115 = 11{,}003.$$

The values of f for M.D. Cordite calculated from Abel and Noble's formulae by the measurement of pressure in a closed vessel are from 10,000 to 11,100; the two results are in very good agreement.

From the calculation of the explosion temperature from the results of gas analysis after cooling we get $f = 10{,}300$ to $10{,}200$; this shows a very remarkable agreement.

To obtain accuracy, however, in the calculation of f, the specific volume v_0 must be corrected for the covolume.

In the equation of Abel and Noble,

$$p(v - \alpha) = RT,$$

the value of α depends, of course, on the units of p and v. Consider a gas of unit volume under standard conditions; then by analogy with the equation for ideal gases, we have

$$p_0 = \bar{v}_0 = 1,$$

$$p(\bar{v} - \beta) = (1 - \beta)\frac{T}{273},$$

where β is the covolume; therefore for unit weight of powder

$$p(v - \beta v_0) = \frac{(1-\beta)v_0}{273}T;$$

consequently f should be modified by introduction of the factor $(1-\beta)$,

$$f = \frac{p_0 v_0 (1-\beta)}{273}T.$$

According to the results of experiments in a closed vessel by Charbonnier and others and also in accord with theoretical calculations* βv_0 or α is nearly equal to 1, then

$$\beta = \frac{1}{1000}$$

and the difference in f is then less than 0.1 per cent.

Calculation of Explosion Pressure

By the aid of Noble and Abel's equation, the pressure of explosion P can be easily calculated for a loading density Δ by taking the covolume $\alpha = 1$.

There is still room for plenty of discussion on the question as to whether or not Noble and Abel's equation satisfies the relation for covolume, explosion pressure and density of loading.

(4) *Results.*

Calculations are given for three different explosives,

M.D. Cordite,

M.K. 1, Cordite,

Nitrocellulose powder.

*D. Berthelot, Comptes Rendus Acad. Sciences, t. 161, p. 209, 1915.

			M.D. Cordite	Cordite M.K.1.	Nitrocellulose powder
Heat of formation		cal./g.	520	569.0	695
Elementary composition	Wt. %	C	25.40	23.00	27.91
		O	57.50	58.96	58.96
		H	3.10	3.04	3.18
		N	14.00	15.00	11.90
	Mol.%	A_c	21.60	19.60	23.50
		A_o	36.60	38.00	36.04
		A_h	15.80	15.50	15.95
		A_n	5.10	5.70	4.28
Equilibrium constant		K	$\frac{(21.6-x)(15-x)}{x(x+0.8)}$	$\frac{(19.6-x)(18.4-x)}{d(x-2.9)}$	$\frac{(23.50-x)(12.54-x)}{x(x+3.41)}$
Temperature of explosion		Absolute	3115	3323.5	2325°
K			8.10	9.20	5.30
x			5.50	5.95	4.0
Component of explosion gases	Wt. %	CO_2	23.70	25.32	17.42
		CO	44.20	36.88	54.03
		H_2	1.25	0.60	1.48
		H_2O	16.75	21.78	15.22
		N_2	14.10	15.42	11.85
	Vol. %	CO_2	12.88	14.54	9.11
		CO	37.90	33.34	44.65
		H_2	14.84	7.47	16.95
		H_2O	22.38	30.65	19.55
		N_2	12.02	14.00	9.74
Total volume of gas		L/kg.	933.0	884.0	970.0
Heat of explosion	Water as gas	cal/gm.	943	1047.0	663.0
	Water as liquid	cal/gm.	1044	1177.0	754.0
Force of explosive f		kg./cm²	11003	11118	8526

(5) *Summary.*

(i) The gaseous products of explosion undergo changes in their composition according to the variation of the equilibrium constants due to cooling. Hence the constituents of a gas as given by analysis after cooling are not the same as at the moment of explosion.

(ii) Therefore, the previous calculations used in the theory of explosives are incorrect as they are founded on the composition of gases after cooling.

(iii) The chemical reaction at the instant of explosion is simply that of water-gas. The formation of methane does not occur at the moment of explosion but during the cooling stages of the gas.

(iv) A discussion was given on the thermodynamic relations of the actual gas mixture. The difference from the ideal gas mixture is very small at high temperatures as in the case of explosion gas, but the correction cannot be found by experiment.

(v) The characteristic equation of a gas at high temperature and under high pressure was examined.

(vi) The equilibrium constant of a water gas reaction near the explosion temperature was calculated by means of Pier and Bjerrum's equations for the specific heat applied at high temperature.

(vii) The reaction for methane formation at high temperature was similarly calculated, and it was proved that actual methane formation only occurs at temperatures less than 1500° C.

(viii) The three fundamental quantities, the temperature of explosion T_0, the concentration of carbondioxide x and the equilibrium constant K were graphically evaluated.

(ix) By the aid of the above quantities T, K and x, we were able to calculate the necessary data of explosives such as the heat of explosion, the composition of gases at the moment of explosion, and the force of explosive.

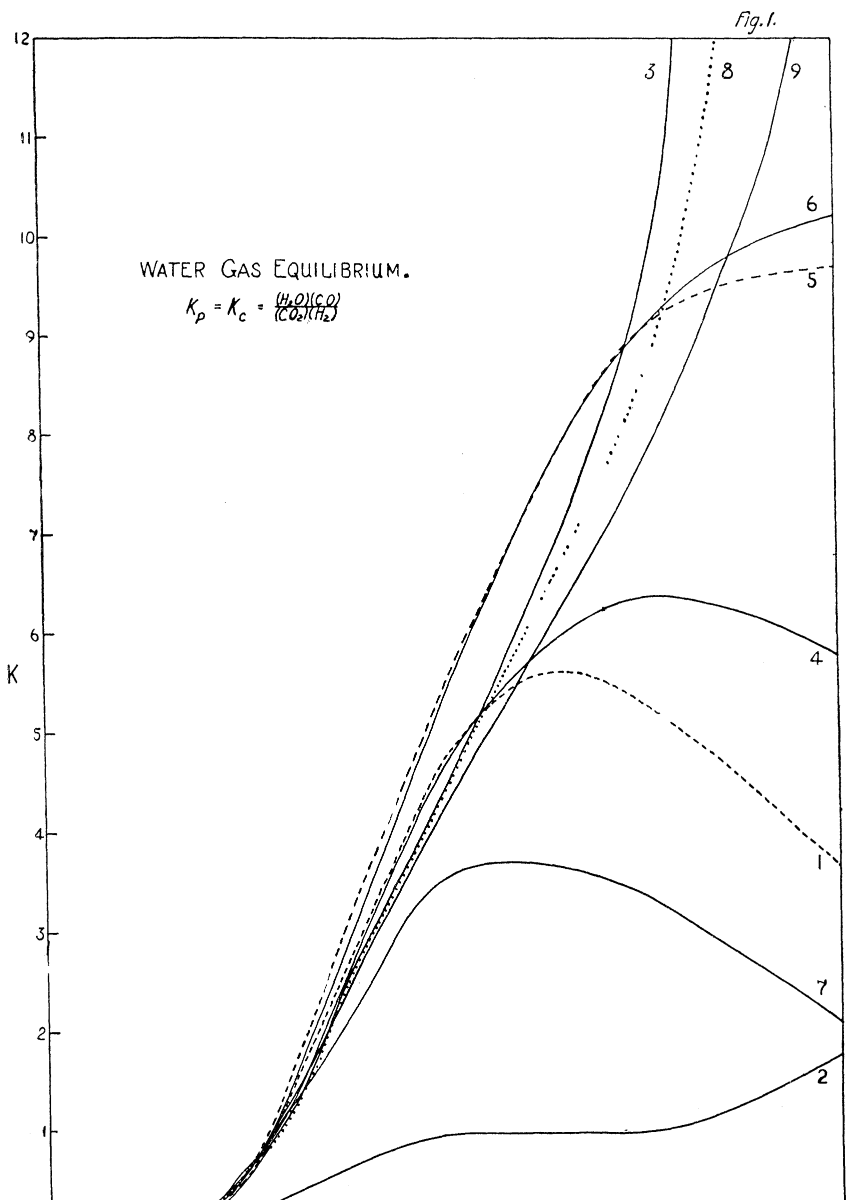


Fig. 1—See p. 628

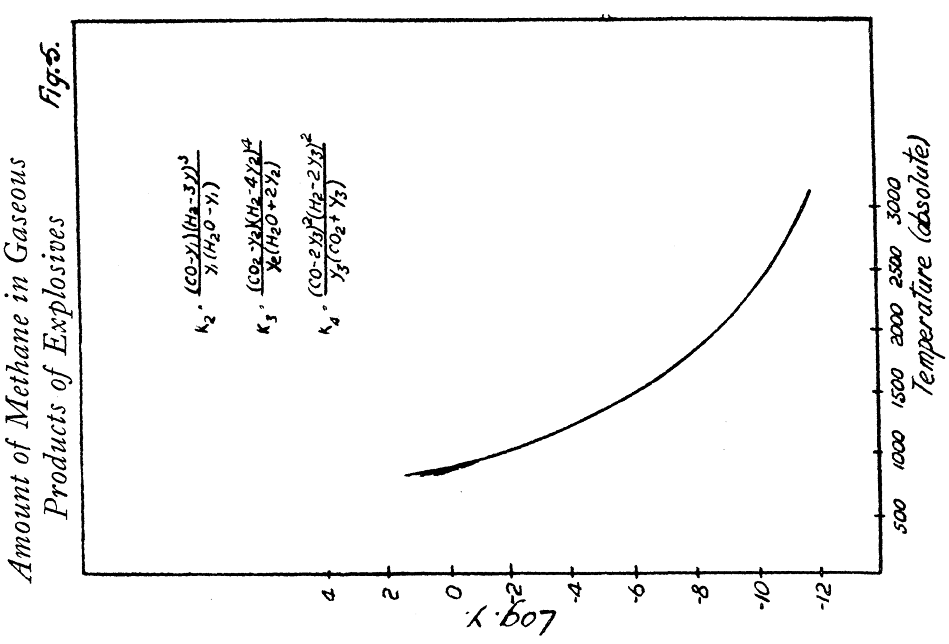


Fig. 5—See p. 643

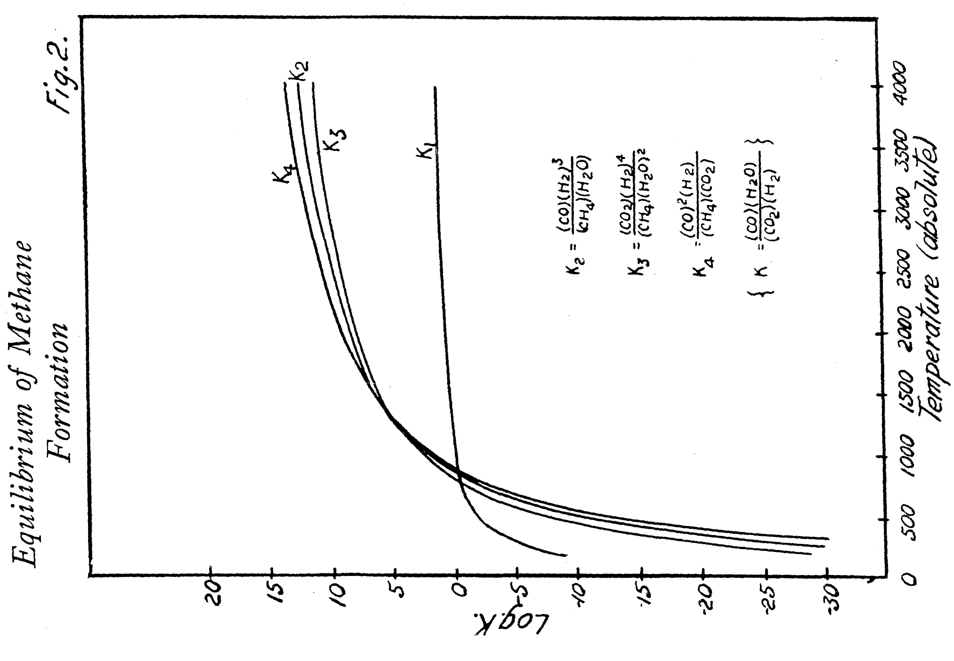
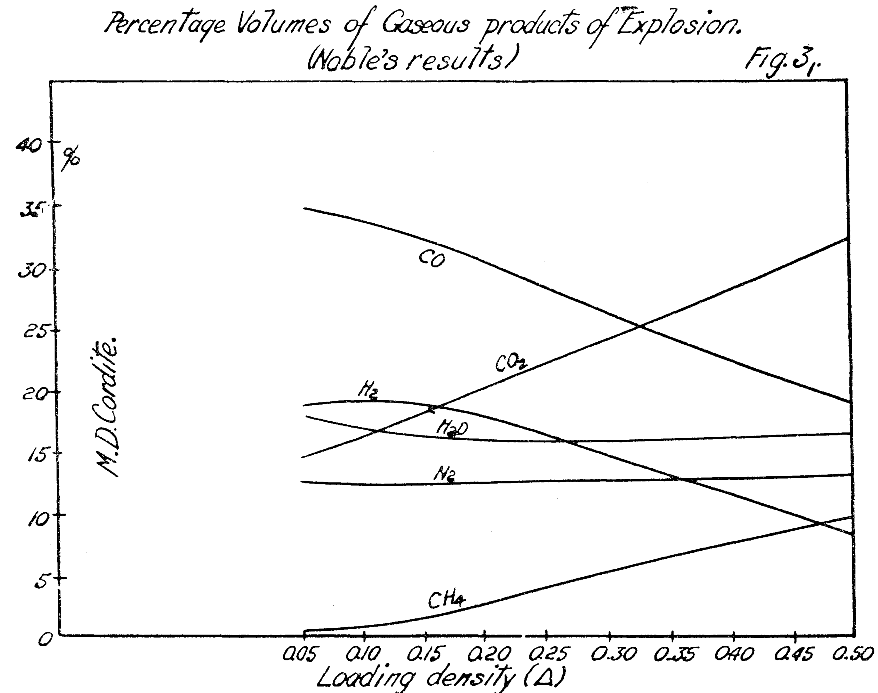


Fig. 2—See p. 635



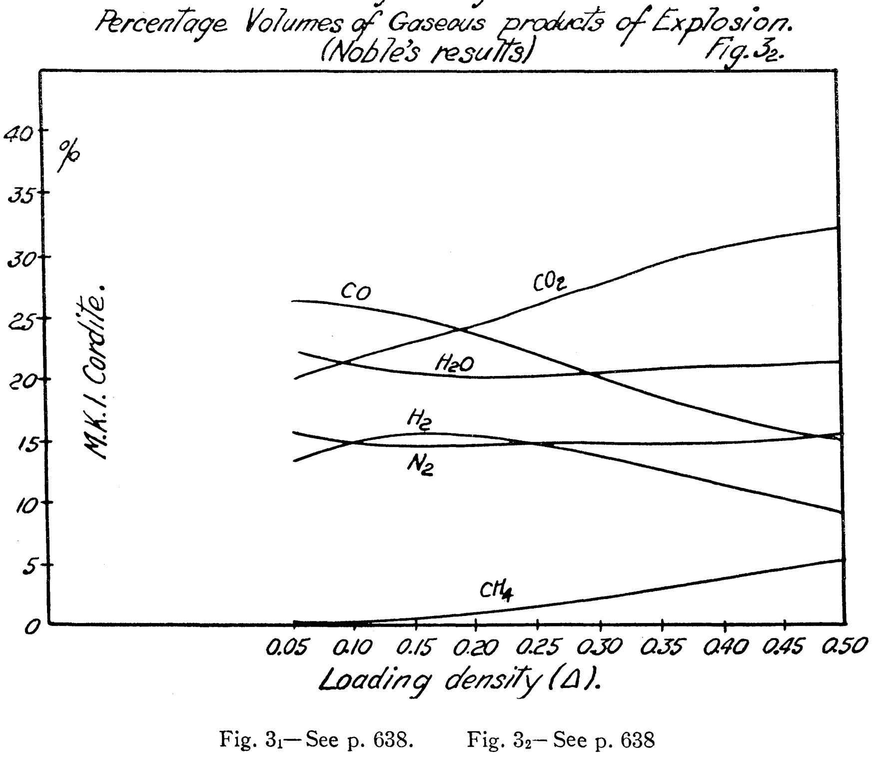
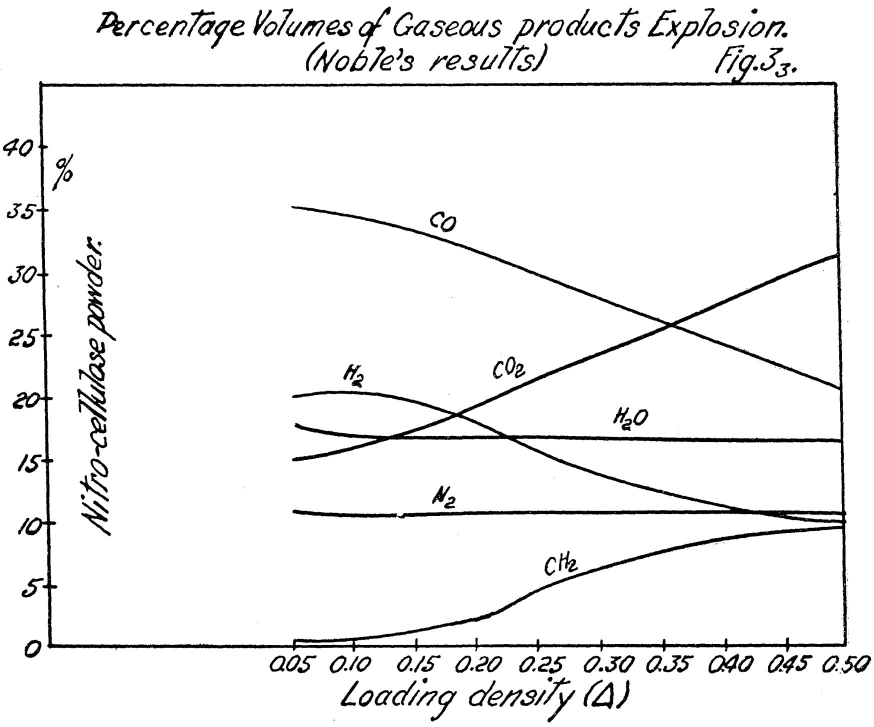


Fig. 3_1—See p. 638. Fig. 3_2—See p. 638



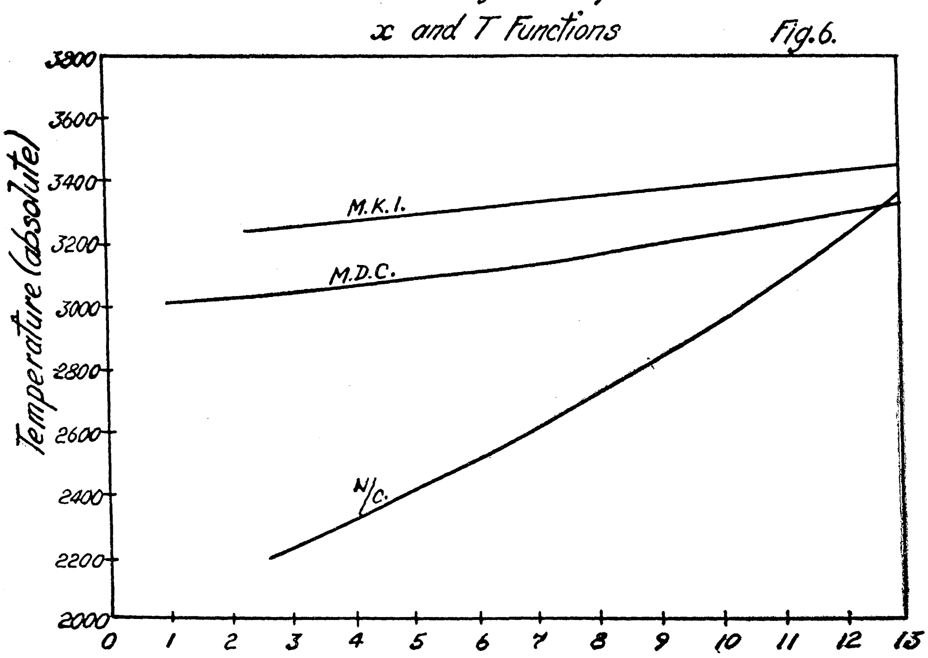
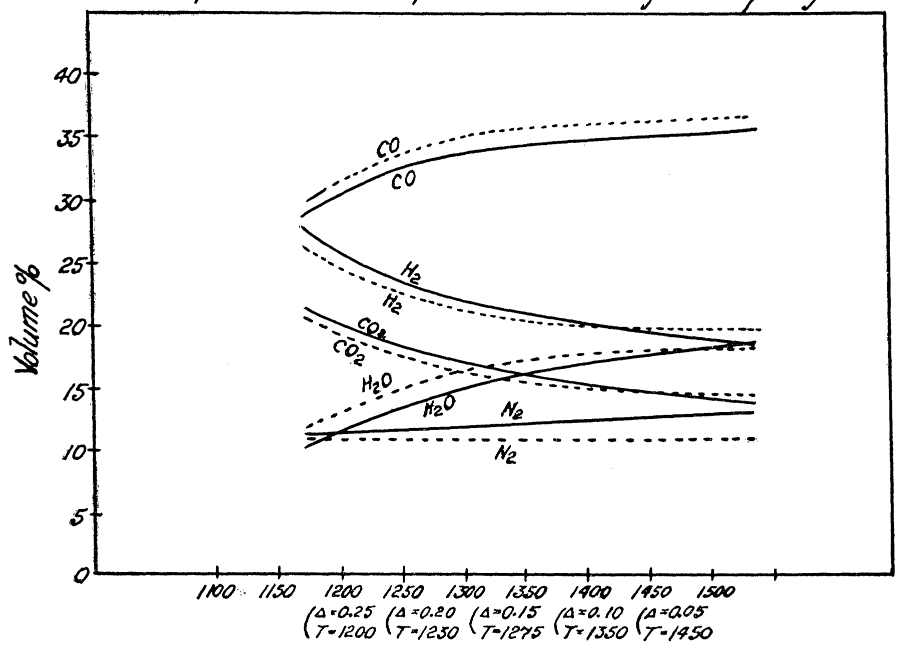


Fig. 3_3—See p. 638. Fig. 6—See p. 646

Temperature. Gas composition and Loading density. Fig. 4.



Equilibrium Composition Curve. Fig. 7.

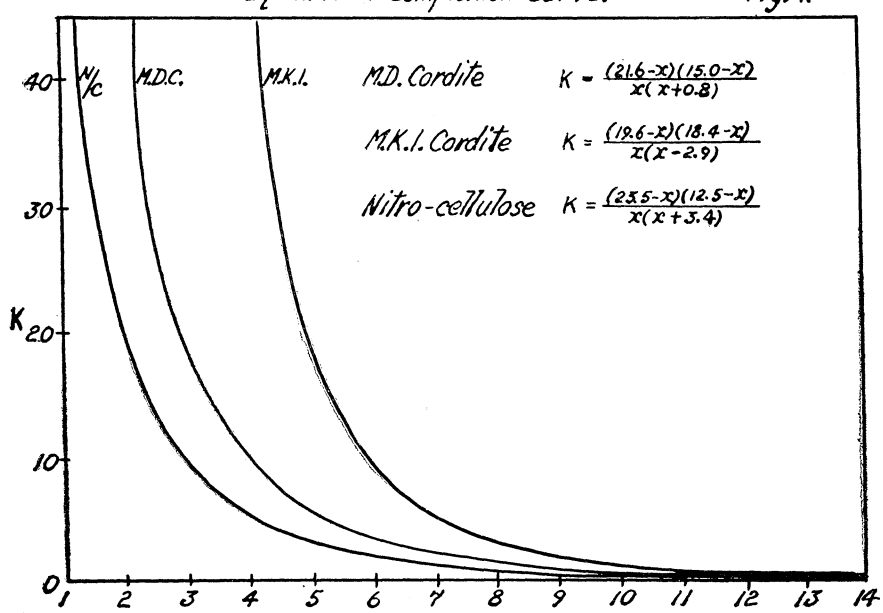
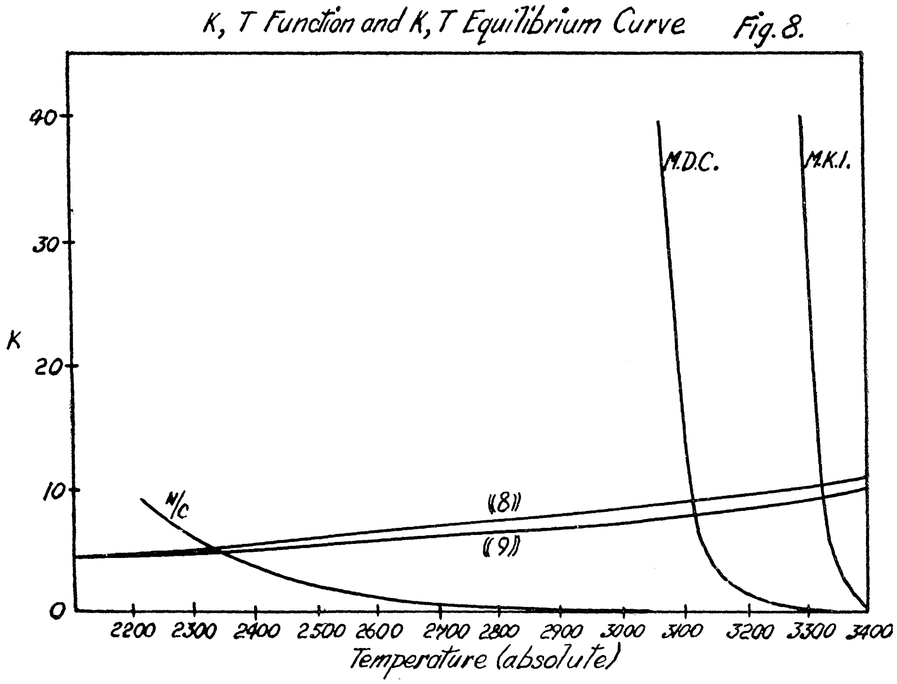


Fig. 4—See p. 640. Fig. 7—See p. 646



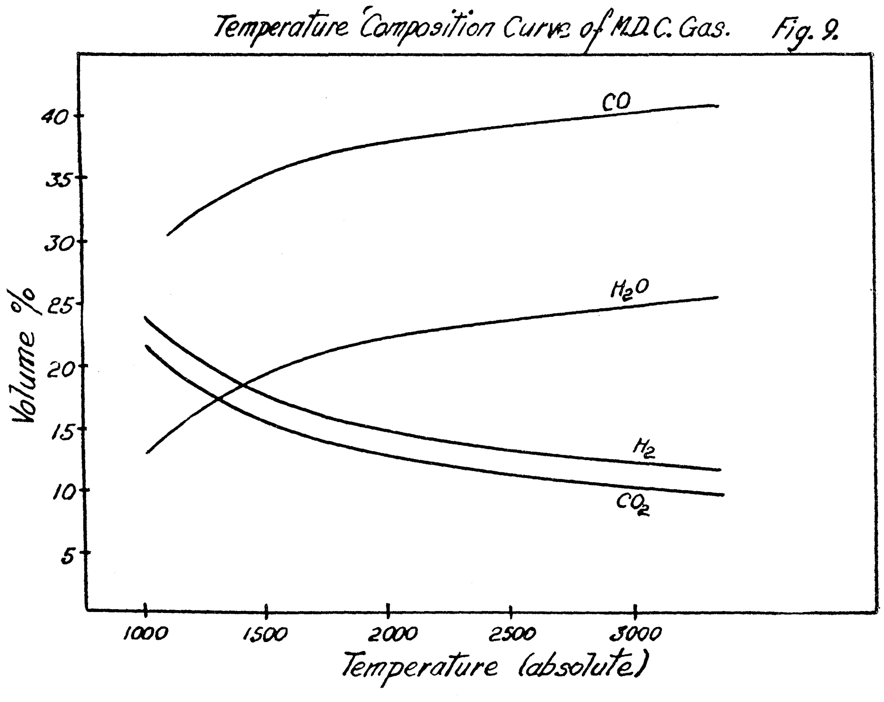


Fig. 8—See p. 646. Fig. 9—See p. 648

APERÇU HISTORIQUE SUR LES TRAVAUX AÉRODYNAMIQUES FAITS EN DANEMARK AVANT 1900

PAR M. ERIK SCHOU,
Professeur à l'École Polytechnique de Copenhague, Copenhague, Danemark.

Le développement de l'Aérodynamique pendant les vingt-cinq dernières années a été si rapide et les ressources qui sont à sa disposition au moment actuel sont si riches qu'il est difficile de se figurer comment les savants d'autrefois ont pu parvenir aux résultats sur lesquels est fondée l'Aérodynamique d'aujourd'hui. Il me semble bien justifié de conserver avec gratitude les noms de ces hommes et c'est pourquoi je me permets à cette occasion, de jeter un coup d'œil sur les travaux anciens de trois de mes compatriotes qui, par l'originalité de leur pensée et par leurs expériences faites par des méthodes reconnues bonnes aujourd'hui encore, méritent bien de ne pas être oubliés.

Je commence par M. Vogt dont le nom est peut-être encore dans la mémoire de quelques lecteurs de Engineering ou de The Steamship des années 1890 ou environ. M. Vogt a eu le bonheur d'étudier à l'Université de Glasgow vers 1880 avec Lord Kelvin, et il voue encore une profonde reconnaissance à cet homme éminent. A cette époque M. Vogt s'occupait déjà de recherches aérodynamiques et il avait la hardiesse de faire des spéculations sur la possibilité de construire des machines à voler. Il est assez caractéristique que Lord Kelvin le sermonna en lui démontrant qu'il n'était pas convenable à un homme de science de s'occuper d'une telle question. Ce qui n'empêcha pas, du reste que Lord Kelvin lui donna aimablement une introduction à l'Association Aéronautique Britannique.

M. Vogt appartenant à une vieille famille de marins s'intéressa de bonne heure de tout ce qui concerne la navigation et doué d'une forte imagination scientifique il étudia en théorie et en pratique la manière d'action des voiles, des rames et des hélices. De cette manière il parvint de très bonne heure à sa conception fondamentale, à savoir que, pour tous les propulseurs agissant sur des fluides, c'est la raréfaction du milieu sur la face postérieure qui joue le plus grand rôle, tandis que la compression sur la face antérieure qu'on avait considérée uniquement jusqu'ici ne compte relativement que pour peu de chose.

Il est bien probable que cette manière de voir était à cette époque, acceptée par d'autres, mais il est certain que pour M. Vogt l'idée était tout à fait originale et qu'elle rebutait presque tous ceux qui s'occupaient alors de la théorie de la navigation. En effet, ce principe ne fut généralement admis qu'une trentaine d'années plus tard.

Pour M. Vogt l'effet des propulseurs reposait sur le mouvement du fluide autour des surfaces en question et il considérait que la théorie des hélices comparant l'effet de ce propulseur avec le mouvement d'une vis dans son écrou n'a aucune relation naturelle avec ce qui se passe en réalité. Ainsi, pour M. Vogt, il n'y rien de surprenant dans le phénomène du *slip* négatif, et il écrivit la-dessus dans l'Engineering de 1891 à l'occasion d'une expérience donnant un *slip* très considérable: «Now, in accordance with a correct theory of propulsion, there is not the least singularity in this, because the centrifugal force which causes the rarefaction on the passive sides of the blades is one function of the revolutions, the slip, in its ordinary sense, is another function of the revolutions, resistance, pitch, etc., but these two functions are of entirely different character».

L'expérience en question, bien intéressante en elle-même, concernait une hélice d'un pas moyen de 1 pied qu'on anima d'une rotation de 70 révolutions par seconde, après quoi elle fut lâchée. L'hélice monta alors jusqu'à une hauteur de 200 pieds et l'on put constater que la montée de 30 à 130 pieds se fit en moins d'une seconde. Le nombre de revolutions étant certainement pendant cette montée moindre de 70 par seconde l'hélice a eu un *slip* négatif très considérable.

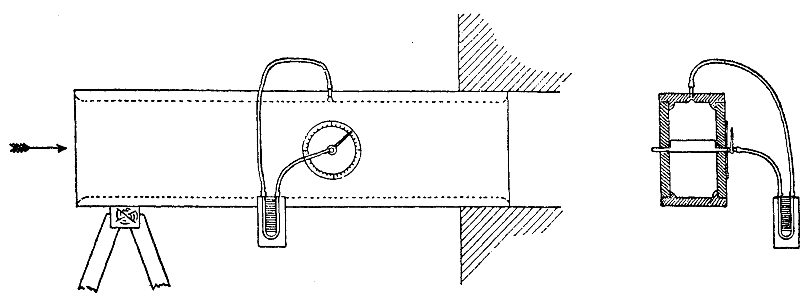


Fig. 1

Dans le cas des voiles M. Vogt établit la raréfaction sur les faces postérieures en calculant d'une part la pression sur les faces antérieures par la formule de Newton, et d'autre part la pression entière du vent. Il montra alors que la formule de Newton ne donnait qu'une fraction de la pression entière et il en tira la conclusion que la raréfaction devait jouer un rôle très considérable.

Outre les questions indiquées plus haut, M. Vogt s'est occupé de beaucoup d'autres considérations concernant le mouvement des fluides et la navigation, mais la place ne permet pas de les mentionner ici. M. Vogt n'a pas fait des expériences lui-même, mais ses travaux ont donné lieu à des expériences faites par d'autres, par lesquelles l'exactitude de ses idées a été confirmée. Parmi les savants qui ont été influencés par les idées de M. Vogt il faut nommer en premier lieu M. Irminger qui par une série d'expériences admirables faites en 1893-94 a demontré que c'est bien la raréfaction du milieu à l'arrière d'un corps exposé à un courant d'air qui à de petits angles d'incidence joue un rôle prépondérant vis à vis de la pression exercée sur le corps. Ces expériences, remarquables par leur méthode ingénieuse, mettent en lumière d'une manière directe le phénomène en question.

Dans ses recherches M. Irminger employa un des premiers la méthode d'un courant d'air artificiel, le corps se trouvant en repos, le mouvement de l'air étant provoqué par le tirage d'une haute cheminée. Le corps à étudier était placé dans une conduite rectangulaire de 11 cm. sur 23 cm., longueur 1 m., mise en communication avec la cheminée (voir Fig. 1). Pour étudier la répartition de la pression sur la surface du corps celui-ci était creux et la surface était percée d'un certain nombre de petits trous de sorte qu'en bouchant tous les trous à l'exception d'un seul on était à même de mesurer la pression sur la surface au point ou se trouvait ce trou. A cet effet le corps était monté sur un axe creux traversant la conduite, et en tournant cet axe on pouvait donner au corps

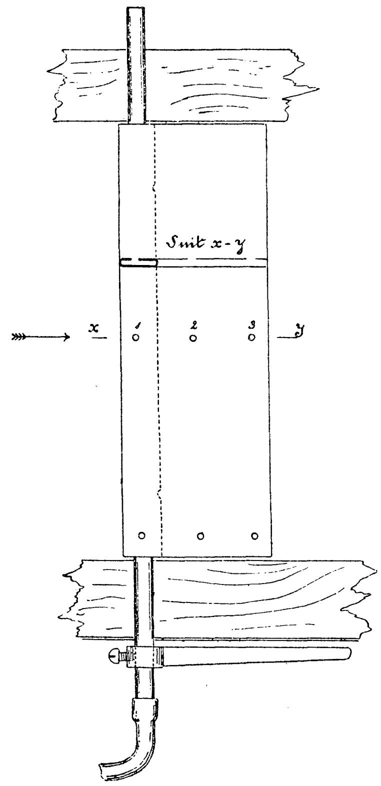


Fig. 2

toutes les inclinaisons par rapport au courant d'air. L'axe creux était mis en communication avec l'une des extrémités d'un tuyau de verre incliné et contenant un liquide approprié, l'autre bout du tuyau communiquant avec un petit trou percé dans la paroi de la conduite. La différence entre la pression au point considéré du corps et au trou de la paroi pouvait ainsi être mesurée par le mouvement du liquide. Dans une série d'expériences le corps avait la forme d'une plaque creuse de 2,2 mm. d'épaisseur (voir Fig. 2) la surface étant percée de six trous, trois dans la ligne du milieu et trois sur le bord, ces derniers donnaient les mêmes pressions que les trous se trouvant au milieu de la plaque et ils servaient ainsi de contrôle. Les résultats obtenus sont résumés dans le tableau suivant où l'on a indiqué le pourcentage de la pression entière due à la raréfaction de l'air sur la face postérieure de la plaque.

Angle d'incidence.	Effet de la raréfaction.
5 p.c.	100 p.c.
10 "	87 "
20 "	83 "
40 "	67 "
60 "	63 "
90 "	57 "

Les résultats sont encore representés sous forme graphique dans la Fig. 3 où les abscisses représentent les angles d'incidence, tandis que les ordonnées donnent les pressions observées, les dépressions étant portées vers le haut et les

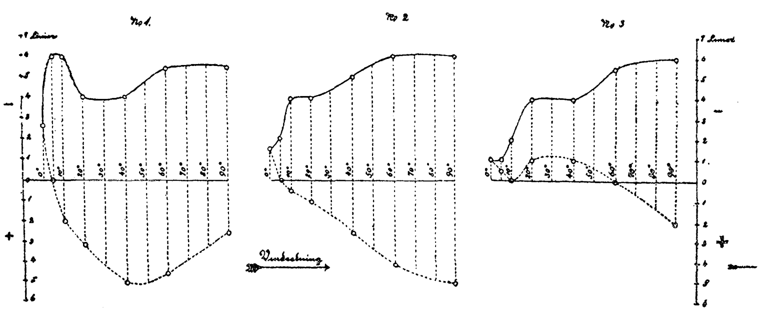


Fig. 3

pressions positives vers le bas. La courbe pleine se rapporte au côté à l'abri du vent, la courbe pointillée au côté opposé. L'on voit bien que ces résultats confirment d'une manière complète la théorie de M. Vogt. Dans d'autres séries d'expériences M. Irminger substitua à la plaque creuse d'autres corps. De

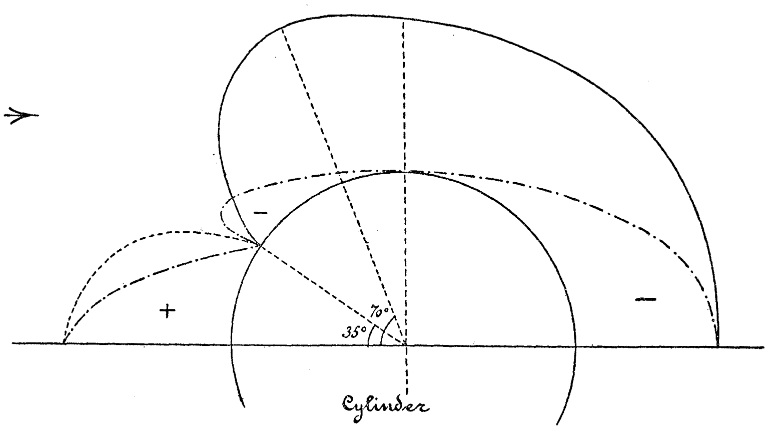


Fig. 4

cette façon il étudia la répartition de la pression sur la surface d'un cylindre (Fig. 4) et sur des corps formant des modèles de maisons (Fig. 5 et 7). Ces dernières expériences établissent la conclusion inattendue que la résultante de la pression est dirigée vers le haut, ce qui explique l'action destructive des ouragans sur des constructions. Les essais de M. Irminger ont été rapportés dans l'Engineering de 1895.

Ces expériences constituaient un progrès important, tant pour la connaissance des lois aérodynamiques, que pour la technique à appliquer dans les recherches de ce genre. A côté de ces deux savants qui heureusement continuent encore leurs recherches, je nommerai un troisième savant danois, La Cour, mort prématurément depuis longtemps.

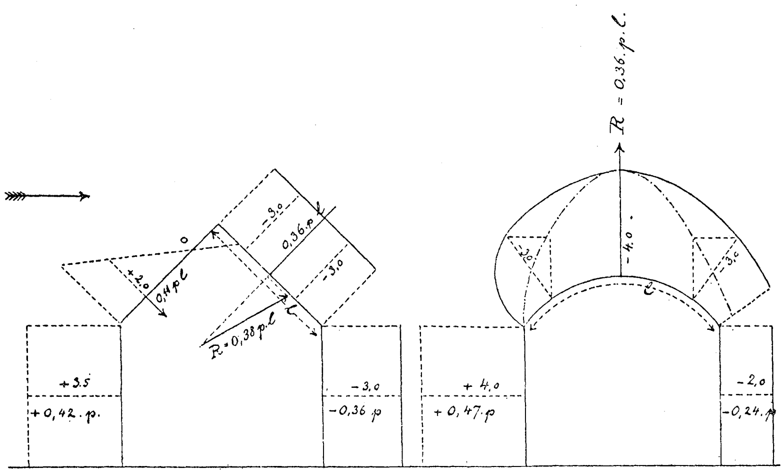


Fig. 5

La Cour fut amené à s'occuper de l'aérodynamique en étudiant la question du perfectionnement des moulins à vent. Pour le Danemark qui ne possède que des sources d'énergie restreintes, ces machines ont une certaine importance, et pour La Cour qui était un physicien éminent et qui demeurait à la campagne il était naturel de chercher une base solide pour leur construction.

La Cour eut le bonheur d'être aidé, dans ses recherches par le gouvernement danois qui à partir de 1891 vota des subventions considérables à une Institution, *Le Moulin d'Essais*, créée par lui. Comme base à des travaux ultérieurs d'ordre pratique La Cour commença par la détermination de la pression d'un courant d'air artificiel sur des plaques de formes différentes. Les expériences étaient donc exercées d'après le même principe qu'adopta M. Irminger quelques années plus tard. La Cour avait à sa disposition deux canaux de forme cylindrique de diamètre 0.5 et 1m. respectivement et ayant tous les deux la longueur de 2.25 m. (Fig. 6). Le ventilateur, placé à l'un des bouts du canal, soufflait un courant d'air par le canal sur le corps à étudier, placé à l'autre bout. A l'intérieur du canal on avait prévu des plaques de manière que tout mouvement de rotation du courant d'air fût évité.

L'on voit bien que les canaux étaient de dimensions très modestes en comparaison avec ceux qu'on emploie maintenant, mais au moment où La Cour faisait ses recherches cet appareil réalisait tout de même un progrès important, et il permettait à La Cour de faire des recherches qu'on n'avait pu réaliser auparavant.

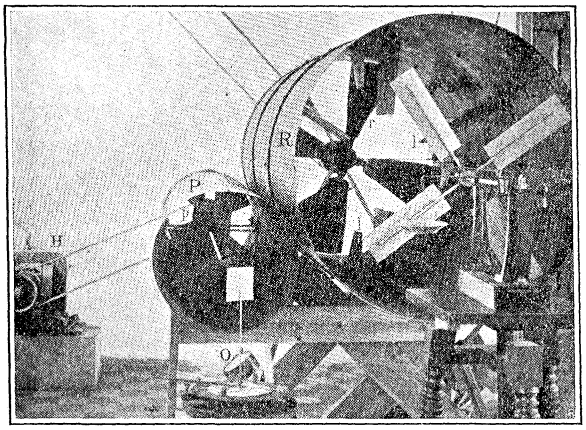


Fig. 6

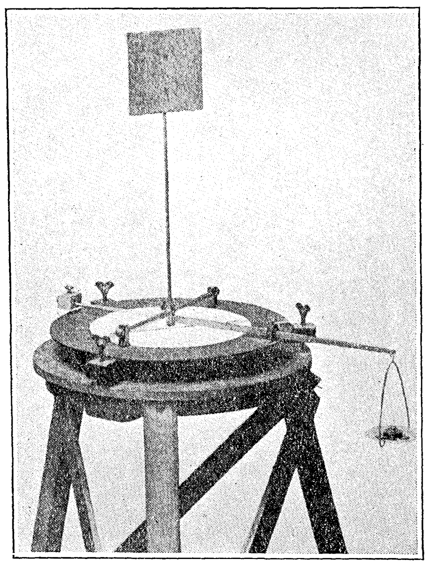


Fig. 7

La Cour commença par la détermination de la pression normale sur une plaque plane, et ayant remarqué que la résultante de la pression totale ne se confond pas avec la normale, ce qu'on avait admis jusqu'alors, La Cour détermina, pour des angles d'incidence variés, l'angle en question. L'appareil très simple dont se servit La Cour dans cette étude est représenté dans la Fig. 7.

La plaque qui, dans la plupart des expériences, avait la forme d'un carré d'un décimètre de côté, était fixée sur une barre en acier qui pouvait tourner dans un plan vertical autour d'un axe horizontal. A cet axe était fixé un levier de sorte qu'on pouvait équilibrer la pression sur la plaque au moyen de poids placés sur un plateau suspendu au levier. Tout le système pouvait être orienté autour d'un axe vertical de sorte qu'il fût possible de varier à volonté l'angle d'incidence du courant d'air sur la plaque.

En se servant de cet appareil, La Cour put déterminer assez exactement la composante normale de la pression, mais pour trouver la valeur de l'angle que faisait la direction de la pression totale avec la normale de la plaque, il fallut employer un autre procédé. Pour cela on fit tourner tout l'appareil autour de l'axe vertical jusqu'à ce que la direction de la pression totale devînt parallèle à l'axe horizontal: le système se trouvant alors en équilibre astatique la position voulue était facile à reconnaître (voir Fig. 8).

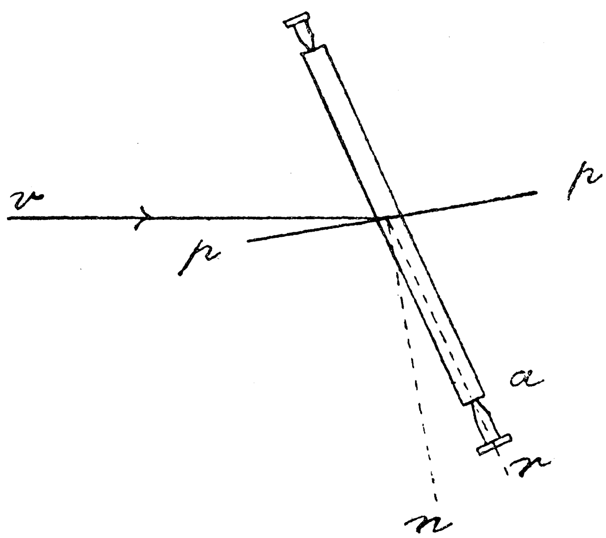


Fig. 8

Ces expériences furent exécutées non seulement avec des plaques planes, mais aussi avec des plaques de forme cylindrique et enfin avec des plaques composées de deux plans se rencontrant sous un angle obtus. Les résultats obtenus avec les dernières plaques ont une grande importance pour la construction des ailes du moulin à vent, parce que pour ces plaques la pression totale est dirigée d'une manière particulièrement avantageuse.

La Cour employa en outre ces canaux à vent pour étudier des modèles de moulins à vent de formes diverses, et parvenait ainsi à des résultats d'une grande utilité pratique.

Les travaux de La Cour ont eu une grande importance pour développer l'usage des moulins à vent en Danemark, en particulier pour leur adaptation à la production d'électricité. Bien entendu, il ne s'agit ici que de toutes petites installations, mais toutefois pendant la guerre mondiale où il était extrêmement difficile pour l'industrie danoise de se procurer les combustibles solides et liquides nécessaires, l'usage du moulin à vent a joué un rôle considérable pour nombre de petites installations électriques.

Les travaux exécutés par les trois savants MM. Vogt, Irminger et La Cour sont naturellement modestes en comparaison avec tout ce qui a été fait ensuite, mais je crois que ces trois hommes ont néanmoins par leurs idées originales contribué au premier début du développement de l'Aérodynamique et par suite j'ai considéré que c'était un devoir de mentionner leurs noms et leurs travaux à cette occasion.

ANALYSE DES EFFETS DES PULSATIONS DU VENT SUR LA RÉSULTANTE AÉRODYNAMIQUE MOYENNE D'UN PLANEUR

PAR M. LOUIS BRÉGUET,
Ingénieur-Constructeur, Paris.

I

L'ÉNERGIE INTERNE DU VENT

On appelle vent, sans autre qualificatif, le déplacement moyen d'ensemble des masses d'air de l'atmosphère, par rapport à un observateur immobile au sol.

Ce mouvement, généralement horizontal, peut cependant avoir une composante ascendante lorsqu'il est dévié par un obstacle naturel ou sous l'action de causes thermiques locales.

Mais si l'observateur, à bord d'un appareil volant, quitte le contact du sol, en cessant d'être lié à lui, *il n'existe plus pour lui qu'un vent relatif, égal à sa vitesse aérodynamique.*

C'est ainsi que seule cette vitesse aérodynamique intervient dans l'étude du vol dans l'atmosphère, et un vent de vitesse uniforme ne fait que modifier, par une simple composition de vitesses, la trajectoire par rapport au sol, *sans jamais exercer sur l'oiseau ou l'aéronef aucune influence aérodynamique.*

Ceci posé, avant d'aborder le sujet de cette étude qui est le vol à voile dynamique, disons un mot de ce qu'on appelle *le vol à voile statique* dont la technique est élémentaire et des plus simples. On démontre qu'il existe une incidence de vol, qui est celle du minimum de puissance consommée en vol horizontal, pour laquelle la vitesse de descente d'un planeur suivant une trajectoire rectiligne dans la masse d'air qui l'environne est minimum est égale à u_m.

Si cette descente s'effectue dans un vent dont la vitesse verticale est égale ou supérieure à u_m, un observateur verra du sol le planeur rester dans un même plan horizontal, ou même s'élever d'un mouvement uniforme, et l'on dira que le planeur effectue du vol à voile statique.

Rappelons que la vitesse u_m est d'ailleurs d'autant plus faible que la charge au mètre carré des ailes est elle-même plus petite, mais alors, sa vitesse horizontale diminuant dans le même rapport, le planeur peut se trouver entraîné par le vent loin de la zone favorable d'ascendance. Cette remarque explique que, par vent violent, les oiseaux trop peu chargés au mètre carré ne peuvent utiliser ce mode de stationnement.

Mais si les oiseaux qui volent à voile recherchent, pour s'y maintenir, les zones d'ascendance du vent, l'observation courante révèle qu'un grand nombre d'entre eux sont susceptibles de parcourir, sans perte d'altitude, de longues distances dans un vent uniquement horizontal, sans qu'apparaisse le moindre battement de leurs ailes.

Pour expliquer ce vol, dit *vol à voile dynamique*, il est nécessaire d'introduire et d'étudier une notion nouvelle, *celle de l'énergie interne du vent* dans laquelle l'oiseau puise la puissance nécessaire à son vol.

Cette énergie, dont l'existence n'a pas échappé à certains expérimentateurs tels que Langley qui en avait, dès 1893, entrevu l'utilisation, résulte de l'existence, quand il y a vent, d'accélérations alternatives des masses d'air déplacées. Toute masse d'air en mouvement possède donc, par rapport au sol, une vitesse moyenne V que nous supposerons horizontale, qui est la vitesse du vent, et une vitesse alternative v qui peut être définie par trois composantes v_x, v_y, v_z: v_x et v_y étant horizontales et v_z verticale.

Chacune de ces trois *pulsations élémentaires*, supposée périodique, est caractérisée par sa vitesse maximum ou intensité et par sa période.

Mais en examinant les diagrammes en fonction du temps des vitesses de ces pulsations relevées dans le vent, on constate qu'ils présentent généralement certaines irrégularités, dues à des pulsations secondaires à plus courte période, oscillant autour d'une ligne moyenne assez régulière représentant la pulsation principale.

Quelles que soient cependant la complexité et l'irrégularité des pulsations périodiques du vent, on sait qu'on peut représenter leur vitesse en fonction du temps par la somme algébrique d'une série de fonctions sinusoidales simples différant par leur amplitude, leur fréquence et leur phase.

Un tel développement en série de Fourier peut d'ailleurs se déduire de la connaissance du diagramme relevé dans le vent.

On peut donc représenter la pulsation v_x par un développement de la forme

$$v_x = \sum v_{x_n} \sin(n\omega t + \psi_n), \tag{1}$$

la période de la pulsation fondamentale étant $T = \frac{2\pi}{\omega}$, ψ_n représentant le décalage de phase des divers harmoniques et n recevant certaines valeurs entières à partir de l'unité.

La pulsation fondamentale est celle de plus longue période; toutefois nous ne considérerons pas ici les pulsations qui auraient des périodes de plus de 20 secondes.

Des développements analogues permettront d'obtenir les expressions analytiques de v_y et v_z.

Il convient de remarquer, ainsi que nous l'avions déjà indiqué, il y a deux ans, que l'existence d'une pulsation v_x suivant la vitesse du vent entraîne nécessairement, ainsi que le montre l'application de l'équation de continuité, l'existence de pulsations de même période dans un plan normal à cette vitesse.

Soient, en effet, σ une section droite du courant, σ_0 la valeur de σ lorsque $v_x = 0$. En négligeant les variations de densité de l'air, la continuité du débit permet d'écrire.

$$\sigma(V + v_x) = \sigma_0 V = \text{const.},$$

d'où

$$\frac{d\sigma}{dt} = -\frac{\sigma_0 V}{(V+v_x)^2}\frac{dv_x}{dt}. \tag{2}$$

Les pulsations normales au courant ayant cette origine auraient donc une vitesse proportionnelle à l'accélération de la pulsation horizontale v_x.

Ce résultat est confirmé par certains diagrammes, le vent instantané descendant quand sa vitesse augmente, montant quand sa vitesse diminue et étant horizontal lorsque son intensité est maximum ou minimum.

D'après M. Idrac, les pulsations latérales du vent seraient en général plus faibles que les pulsations verticales.

Notons aussi que des mouvements tourbillonnaires peuvent donner lieu à des pulsations de toute nature se superposant à celles dont nous venons d'envisager l'origine.

Ceci posé, si une masse d'air M, possédant une énergie cinétique de translation $\frac{1}{2}MV^2$, est animée de pulsations périodiques en tous sens, ainsi que nous venons de l'exposer, elle possède de ce fait une certaine énergie cinétique interne, également périodique, dont une fraction est utilisable par un oiseau ou un planeur.

Au point de vue du vol à voile, cette énergie est caractérisée par *sa valeur moyenne* $\mathfrak{J}$ *dans la période*. En supposant v_x, v_y, v_z représentés par des formules analogues à la formule (1), on trouve sans difficulté

$$\mathfrak{J} = \frac{M}{4}\Sigma(v_{x_n}^2 + v_{y_n}^2 + v_{z_n}^2). \tag{3}$$

L'énergie interne moyenne est donc la somme des énergies apportées par tous les harmoniques des pulsations, tant horizontales que verticales, chaque harmonique intervenant comme s'il était seul. Cette propriété de la superposition des effets des divers harmoniques se retrouve dans leur utilisation pour le vol à voile dont c'est une des particularités les plus dignes d'être signalées.

Notons enfin que si l'énergie de translation $\frac{1}{2}MV^2$ n'est pas perceptible par un oiseau ou un planeur, l'énergie interne $\mathfrak{J}$ dépend cependant de cette énergie à laquelle elle semble être proportionnelle.

Les intensités des pulsations relevées dans le vent croissent en effet avec l'intensité du vent et, pour certaines mesures, la proportionnalité paraît se vérifier.

Les pulsations horizontales dirigées suivant le vent sont généralement de beaucoup les plus intenses, et leur intensité augmente avec leur période.

L'extrémité du vecteur représentant la pulsation résultante décrit donc une surface ayant la forme d'un ellipsoïde dont le grand axe est dirigé suivant la vitesse du vent.

Bien que le vent ne soit pas directement perceptible pour un planeur, il apparaît cependant que si un observateur, à bord de ce planeur, pouvait déterminer le grand axe de cette surface, il connaîtrait ainsi la direction de la vitesse du vent.

Certains auteurs ont prétendu relever dans les vents violents des pulsations horizontales de 5 à 7 mètres, d'intensité, pour des périodes de 6 à 8 secondes.

Les valeur des intensités des pulsations verticales pour différentes fréquences et différents vents n'est pas encore bien déterminée. D'après M. Idrac, des intensités de plus de 1^m par seconde ne s'observeraient que rarement et dans des vents d'au moins 15^m, les valeurs le plus couramment observées étant de 0^m, 20 à 0^m, 50 pour des vents de 7 à 15^m.

Ces chiffres nous paraissent assez pessimistes et nous pensons qu'avec les appareils à fil chaud de MM. Huguenard, Magnan et Planiol, de meilleures valeurs pourront être mesurées.

Après ces quelques préliminaires, nous allons analyser le mécanisme par lequel l'oiseau peut utiliser ces différentes pulsations en étudiant séparément l'effet des pulsations verticales, puis celui des pulsations horizontales reçues latéralement et enfin celui des pulsations horizontales reçues de front.

Nous verrons aussi comment s'opère la superposition des effets des diverses pulsations grâce à laquelle il semble qu'aucune agitation interne du vent ne soit perdue pour l'oiseau qui vole à voile.

Nous donnerons également l'analyse du vol par battements en air calme car elle met en jeu les mêmes phénomènes et s'effectue par des calculs en tous points semblables.

II

L'Effet de Pulsations Aériennes Verticales

C'est depuis longtemps que nous avions indiqué que l'oiseau doit trouver une partie de la puissance nécessaire à son vol sans battement dans l'utilisation des accélérations verticales alternatives des masses d'air (conférences faites à Lille en 1909, à l'Association Française Aérienne en 1922 et à la Société de Navigation aérienne en 1923).

Des essais de laboratoire, commencés à Vienne en 1921 et poursuivis récemment au Laboratoire de Saint-Cyr, ont pleinement confirmé nos théories en montrant que si l'on envoyait sur une aile immobile un courant d'air ondulé verticalement à l'aile de volets oscillants, on voyait la résultante aérodynamique moyenne se redresser jusqu'à donner une composante propulsive.

Un calcul simple, effectué simultanément par M. Rateau et par nous, permet d'analyser ce phénomène d'une façon tout à fait satisfaisante et de chiffrer le décalage observé de la polaire.

Cependant, la période des pulsations qui n'intervient pas pour une aile immobile devient, pour un planeur libre, le facteur prépondérant de la discussion, par suite du mouvement vertical ondulé que prend son centre de gravité.

C'est ainsi que nous avons été conduit, tout récemment, dans deux Notes communiquées à l'Académie des Sciences le 21 janvier et le 10 mars 1924, à présenter l'analyse qui sert de base à cette étude.

Pour la facilité de la discussion, nous étudierons d'abord l'effet d'une pulsation simple de vitesse $v = v_M \sin \omega t$, en montrant par la suite que lorsqu'une pulsation comporte un certain nombre d'harmoniques, les effets de ces harmoniques se superposent simplement.

Soient:

i, c_x, c_z, l'incidence en radians de l'aile au temps t, comptée à partir de celle du minimum de traînée, et les coefficients de traînée et de poussée rapportés à la vitesse aérodynamique et à sa normale;

c_x', c_z', les mêmes coefficients rapportés à l'horizontale et à la verticale, c' étant positif dans le sens des résistances;

a, g, le poids spécifique de l'air et l'accélération de la pesanteur;

v, v_M, T, la vitesse de la pulsation verticale agissant sur le planeur au temps t, sa valeur maximum et sa période; en posant $\omega T = 2\pi$, nous admettrons que $v = v_M \sin \omega t$;

V, θ, la grandeur de la vitesse aérodynamique du planeur, dont nous négligerons les variations, et l'inclinaison au temps t de cette vitesse sur l'horizontale;

P, S, le poids et la surface du planeur.

Les angles i et θ seront considérés comme positifs dans le cas de la figure 1.

Pour rechercher quelle est, pour chaque valeur de θ, l'incidence la meilleure à donner à l'aile, nous admettrons la loi linéaire

(4) $$i = k\theta + i_0,$$

k et i_0 étant ultérieurement déterminés pour que les effets des pulsations soient les plus favorables un vol à voile.

Nous remarquerons que l'incidence de l'aile sur l'horizontale est

(5) $$i' = i - \theta = (k-1)\theta + i_0,$$

de sorte que i_0 n'est autre que l'incidence moyenne de vol et que l'immobilité relative de l'aile correspond à $k=1$.

En nous bornant alors aux incidences, ne dépassant pas 10 à 12°, en fonction

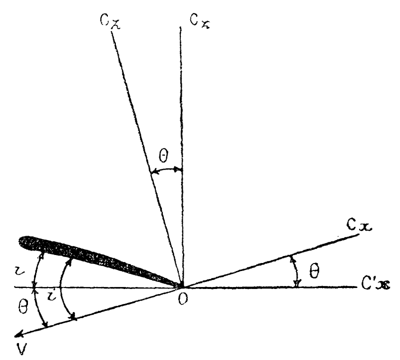


Fig. 1

desquelles on peut représenter c_x par un arc de parabole et c_z par une droite, nous écrirons, α, β, A, B étant des constantes:

(6) $$c_x = \alpha + \beta i^2,$$

(7) $$c_z = A + Bi,$$

d'où, en remplaçant i par sa valeur (4), c_{x_0} et c_{z_0} désignant les valeurs de c_x et c_z pour $i = i_0$:

(8) $$c_x = c_{x_0} + 2\beta k i_0 \theta + \beta k^2 \theta^2,$$

(9) $$c_z = c_{z_0} + Bk\theta.$$

En assimilant cos θ à l'unité, sin θ à θ et en remarquant que θc_x est négligeable devant c_z, il vient:

(10) $$c_x' = c_x - \theta c_z = c_{x_0} - (c_{z_0} - 2\beta k i_0)\theta - (Bk - \beta k^2)\theta^2,$$

(11) $$c_z' = c_z = c_{z_0} + Bk\theta.$$

Ces formules montrent dès maintenant que le redressement de la résultante aérodynamique moyenne s'opère *grâce au changement de sens de la composante horizontale θc_z de la poussée c_z*, composante tantôt propulsive, tantôt résistante, et dont la valeur moyenne dans la période n'est pas nulle, puisqu'elle contient un terme en θ^2.

Si maintenant $\frac{dz}{dt}$ désigne la vitesse verticale du centre de gravité du planeur, on a:

(12) $$\theta = \frac{1}{V}\left(v - \frac{dz}{dt}\right) = \frac{1}{V}\left(v_M \sin \omega t - \frac{dz}{dt}\right),$$

l'équation du mouvement vertical étant

$$\frac{P}{g}\frac{d^2z}{dt^2} = \frac{a}{2g} c_z' S V^2 - P = \frac{a}{2g} c_{z_0} S V^2 + \frac{a}{2g} BkSV^2\theta - P,$$

c'est-à-dire, en remplaçant θ par sa valeur (12).

$$\frac{P}{g}\frac{d^2z}{dt^2} = \frac{a}{2g} c_{z_0} S V^2 + \frac{a}{2g} BkSV\left(v_M \sin \omega t - \frac{dz}{dt}\right) - P.$$

La périodicité du mouvement exige que la somme des termes constants de cette équation soit nulle, condition qui fournit l'équation moyenne de sustentation:

(13) $$P = \frac{a}{2g} c_{z_0} S V^2,$$

avec

(14) $$\frac{2P}{aBkSV}\frac{d^2z}{dt^2} + \frac{dz}{dt} = v_M \sin \omega t.$$

En posant

(15) $$\text{tang}\, \Phi = \frac{2P\omega}{aBkSV},$$

on trouve facilement que la solution permanente de cette équation peut s'écrire

(16) $$\frac{dz}{dt} = v_M \cos \Phi \sin(\omega t - \Phi),$$

l'angle θ, connu par la relation (12), étant alors donné par l'expression

(17) $$\theta = \frac{v_M}{V} \sin \Phi \cos(\omega t - \Phi).$$

Nous voyons que le mouvement vertical du planeur présente, par rapport à la pulsation du vent, le décalage de phase Φ qui est d'autant plus grand que. la période est plus courte, en tendant vers $\frac{\pi}{2}$ pour une très grande fréquence C'est d'ailleurs grâce à ce décalage de phase que le planeur peut utiliser les accélérations périodiques du vent.

Ceci posé, la grandeur et la direction de la résultante aérodynamique moyenne se trouveront définies par les valeurs moyennes dans la période c'_{x_m} et c_z' de c'_{x_m} et c_z' dont les valeurs au temps t sont données par les équations (10) et (11).

La valeur moyenne de θ étant nulle dans la période, on voit que la valeur moyenne de c_z' est c_{z_0}, résultat déjà connu, tandis que c'_{xm} est donné par la formule

$$c'_{x_m} = \frac{1}{T}\int_0^T c_x' dt = c_{x_0} - \frac{Bk-\beta k^2}{2V^2} v_M^2 \sin^2\Phi. \tag{18}$$

La condition du vol à voile à trajectoire moyenne horizontale est que la valeur de c'_{x_m} soit nulle, la résultante aérodynamique moyenne étant alors verticale.

Il en résulte que le carré de l'intensité des pulsations qui permettraient le vol à voile doit être

$$v_M^2 = \frac{2V^2 c_{x_0}}{(Bk-\beta k^2)\sin^2\Phi}. \tag{19}$$

En remplaçant V par sa valeur tirée de l'équation (13) de sustentation, sin Φ par sa valeur, connue par la formule (15), et en désignant, suivant la notation habituelle, $\frac{c_{x_0}}{c_{z_0}}$ par tang ϕ_0, on trouve finalement:

$$v_M^2 = \frac{2g \text{ tang } \phi_0}{Bk-\beta k^2}\left[\frac{2P}{Sa} + \frac{gB^2k^2}{\omega^2 c_{z_0}}\right]. \tag{20}$$

Si nous annulons la dérivée de cette expression par rapport à k, il apparaît que la meilleure valeur à donner à k est la racine positive de l'équation du deuxième degré

$$\frac{k^2}{B-2\beta k} = \frac{2P}{Sag}\,\frac{c_{z_0}\omega^2}{B^3}. \tag{21}$$

La plus grande valeur de k est $\frac{B}{2\beta}$ correspondant aux pulsations de très faible période.

Il convient d'ailleurs de noter que, pour de telles périodes, l'aile d'un oiseau ne peut osciller en vue de satisfaire à cette condition qui est ainsi toute théorique.

Lorsque la période croît, k diminue en se rapprochant de zéro pour les très longues périodes.

Pour $c_{z_0}=0{,}5$, $B=5$, $\beta=1{,}3$, on trouve, en effet:

	$\frac{P}{S}$	Période (en secondes) 0,25	0,5	0,75	1	2	4	0,17
Albatros....	10kg	$k=1{,}68$	1,31	1,94	0,86	0,50	0,27	0,17
Pétrel......	5	$k=1{,}52$	1,07	0,82	0,66	0,37	0,19	0,13
Mouette....	2,5	$k=1{,}31$	0,85	0,63	0,49	0,26	0,14	0,094

Remarquons que si l'on admet que k ait toujours sa valeur la plus favorable, l'élimination de ω^2 entre les deux relations (20) et (21) permet d'écrire la formule donnant les valeurs exigées pour v_M sous la forme remarquable par sa simplicité

$$(22) \qquad v_M^2 = \frac{8gP}{Sa} \frac{\text{tang}\ \phi_0}{kB}.$$

Mais les pulsations du vent sont, comme nous l'avons à plusieurs reprises mentionné, constituées par la superposition d'une série de pulsations périodiques simples analogues à celle envisagée et de périodes différentes, et il ressort immédiatement de l'examen des formules (20) ou (22) que, parmi toutes les pulsations verticales du vent, ce sont celles à courte période qui sont le plus profitables à l'oiseau qui vole à voile.

L'adaptation de l'aile doit donc naturellement s'effectuer pour les courtes périodes, ne dépassant pas une seconde, c'est-à-dire que *pratiquement k doit être égal à l'unité.*

Ainsi que nous l'avons dit, il est par ailleurs impossible que, pour de courtes périodes, l'aile d'un oiseau puisse osciller avec une fréquence suffisante pour suivre le rythme des pulsations du vent et c'est en conservant son aile immobile, c'est-à-dire en adoptant $k=1$, que l'oiseau pourra pratiquement profiter au mieux des pulsations verticales du vent.

La formule (20) devient, dans ces conditions,

$$(23) \qquad v_M^2 = \frac{2g \tan \phi_0}{B-\beta}\left[\frac{2P}{Sa} + \frac{gB^2}{\omega^2 c_{z_0}}\right]$$

Il apparaît de cette équation que, pour les faibles périodes, le second terme de la parenthèse étant petit par rapport au premier, l'intensité réclamée v_M est minimum en même temps que tang ϕ_0 et que, par suite, la meilleure valeur de l'incidence moyenne i_0 est celle de la finesse du planeur.

C'est cette valeur que nous adopterons pour nos calculs, car elle donne pour v_M, même pour des périodes de 1 à 2 secondes, des chiffres très peu différents des valeurs minima.

Notons d'ailleurs que si, pour les grandes périodes, l'aile adaptait au mieux son incidence, la formule (21) montre que k^2 varierait alors proportionnellement à c_{z_0} et la formule (22) indique que la meilleure incidence serait celle du minimum de $\zeta = \frac{c_{x_0}}{c_{z_0}^{\frac{3}{2}}}$.

Discussion.—A l'examen de la formule (23), il apparaît immédiatement que l'intensité exigée v_M de la pulsation est d'autant plus faible que la période est plus courte, que le planeur est moins chargé au mètre carré et que l'altitude de vol est plus basse. Il semble ainsi qu'il y ait intérêt à ce que la charge au mètre carré soit d'autant plus faible que l'altitude du vol est plus élevée.

Le second terme de la parenthèse croissant comme le carré de la période, il en résulte que l'influence de la charge au mètre carré et de la densité de l'air est d'autant plus marquée que la période est plus courte, cette influence devenant presque négligeable pour les grandes valeurs de la période.

Enfin, en ce qui concerne les caractéristiques aérodynamiques du planeur, on constate que ses résistances nuisibles et surtout l'*envergure relative* de son aile jouent le rôle prépondérant.

Le coefficient β, voisin de 1,3, ne semble pas influencé par la forme du profil ou l'envergure relative de l'aile et peut pratiquement être considéré comme une constante quelles que soient les caractéristiques du planeur.

Quant à B, pour tout profil, il ne dépend que de l'envergure relative Λ en se rapprochant d'une certaine limite B_0 pour une envergure infiniment grande.

Un calcul assez simple, basé sur la théorie de Prandtl et dont le résultat se trouve vérifié par les essais de laboratoire, montre qu'on peut, en fonction de Λ. représenter les variations de B par un arc d'hyperbole d'équation

$$B = \frac{B_0}{1+\dfrac{B_0}{\pi\Lambda}}. \tag{24}$$

Remarquons maintenant que la finesse tang ϕ_0 est d'autant plus petite que les résistances nuisibles sont plus faibles et que l'allongement relatif Λ est plus grand. La valeur de tang ϕ_0 est en effet très grande pour de très faibles allongements, puis décroît lorsque l'allongement augmente en se rapprochant d'une certaine limite pour de très grands allongements.

Enfin la portance c_{z_0} atteinte à l'incidence de la finesse varie comme le coefficient B en augmentant avec l'allongement et en tendant vers une certaine limite pour les grands allongements.

Pour un planeur ou un oiseau à résistances nuisibles faibles, nous admettrons, en définitive, les caractéristiques aérodynamiques suivantes:

Λ	4	6	8	10	12
B	3,8	4,2	4,6	4,8	5
tang ϕ_0	0,06	0,05	0,047	0,046	0,045
c_{z_0}	0,4	0,45	0,47	0,49	0,5

C'est ainsi que, pour une envergure relative de 4, pour $\frac{P}{S}=5, a=1{,}225$, $\beta=1{,}3$ et une période de 1 seconde, l'intensité v_M réclamée aux pulsations serait de $2^m{,}75$, tandis que, dans les mêmes conditions, $\Lambda=12$ exigerait seulement $v_M=2^m{,}20$.

Il apparaît donc un avantage certain en faveur des ailes peu profondes et de grande envergure.

Cette conclusion nouvelle est digne de remarque, car elle explique l'étroitesse, étonnante a priori, *des ailes des oiseaux de mer bons voiliers, notamment des albatros dont la grande envergure relative est si frappante.*

A un point de vue différent, un autre avantage semble inhérent aux ailes peu profondes qui se trouvent toujours, de ce fait, affectées par des pulsations aériennes en synchronisme, la profondeur de l'aile devant être suffisamment petite devant la longueur de l'onde aérienne utilisée.

En admettant alors les coefficients d'une aile de grande envergure relative, c'est-à-dire $B=5$, tang $\phi_0=0{,}045$, $c_{z0}=0{,}5$ et en prenant $\beta=1{,}3$ et $a=1{,}225$, nous trouvons:

	$\frac{P}{S}$	Periode (en secondes) 0,25	0,50	0,75	1	2	4	6
Albatros	10kg	$v_M=2{,}02$	2,15	2,36	2,62	3,96	7,16	10,72
Pétrel . .	5	$v_M=1{,}45$	1,64	1,90	2,21	3,71	7,02	10,62
Mouette	2,5	$v_M=1{,}08$	1,31	1,62	1,98	3,58	6,95	10,58

Ainsi que nous l'avons dit, d'après certaines mesures de M. Idrac, des pulsations verticales de 1 m: s n'ont été observées par lui que rarement et dans des vents d'au moins 15^m, les intensités le plus généralement rencontrées étant de $0^m{,}20$ à $0^m{,}50$.

Il apparaît donc que ces pulsations verticales, considérées comme périodiques simples, ne pourraient, à elles seules, entretenir le vol à voile.

Mais, d'après l'équation générale (18), pour un planeur déterminé recevant une pulsation de période donnée, *la diminution de la résistance aérodynamique est proportionnelle au carré de l'intensité de la pulsation.* Il en résulte que si v_M est l'intensité qui serait réclamée pour un redressement total permettant le vol à voile, v'_M l'intensité effective, la résistance aérodynamique moyenne aura pour expression

$$c'_{x_m}=\left(1-\frac{v'^2_M}{v^2_M}\right)c_{z_0}. \tag{25}$$

La valeur angulaire du redressement peut se mesurer par l'inclinaison ϕ_m de la résultante aérodynamique moyenne sur la verticale, *la tangente de cet angle étant la résistance relative apparente.*

Le coefficient moyen de sustentation étant c_{z_0}, on tire de la formule (25) l'expression équivalente

$$\text{tang}\,\phi_m=\left(1-\frac{v'^2_M}{v^2_M}\right)\text{tang}\,\phi_0. \tag{26}$$

On voit de même que la puissance qui serait nécessaire au vol horizontal à l'incidence i_0 se trouve réduite, par l'effet de la pulsation, dans le même rapport $1-\frac{v'^2_M}{v^2_M}$.

C'est ainsi que, si la pulsation a une intensité moitié de celle qui serait nécessaire à l'entretien du vol à voile, la résistance relative et la puissance nécessaire au vol horizontal se trouvent, de ce fait, être encore les trois quarts de leurs valeurs en air calme.

Mais les mouvements internes du vent résultent de la superposition d'un grand nombre d'harmoniques de périodes différentes qui ajoutent leurs effets, ainsi que nous allons le démontrer.

Il est donc possible que l'addition des effets de la pulsation principale et des effets de plusieurs harmoniques secondaires à courtes périodes, communique aux ailes des oiseaux une énergie notablement plus importante.

La Superposition des Effets des Divers Harmoniques d'une Pulsation

Une pulsation du vent ne peut généralement pas se représenter par une seule fonction sinusoïdale simple, mais bien par la somme algébrique d'une série de fonctions sinusoïdales simples différant par leur amplitude, leur fréquence et leur phase. On est alors en droit de se demander si les effets des divers harmoniques de cette série s'ajoutent ou si, au contraire, des interférences réciproques peuvent atténuer l'effet cherché.

Les équations différentielles intervenant dans ce problème étant linéaires, il apparaît bien *a priori* que les effets individuels des harmoniques doivent s'additionner. L'importance pratique de cette constatation nécessite cependant une étude détaillée.

Pour effectuer cette analyse, nous représenterons, ainsi qu'il a été dit, l'intensité de la pulsation en fonction du temps par une série de Fourier

$$v = \sum v_n \sin(n\omega t + \psi_n), \tag{27}$$

la période de l'harmonique fondamental étant $T = \frac{2\pi}{\omega}$, ψ_n représentant le décalage de phase des divers harmoniques et n pouvant recevoir des valeurs entières à partir de l'unité.

L'équation (12) donnant l'inclinaison au temps t de la vitesse aérodynamique devient

$$\theta = \frac{1}{V}\left[\sum v_n \sin(n\omega t + \psi_n) - \frac{dz}{dt}\right], \tag{28}$$

l'équation (14) du mouvement vertical s'écrivant alors, en supposant $k=1$,

$$\frac{2P}{aBSV}\frac{d^2z}{dt^2} + \frac{dz}{dt} = \sum v_n \sin(n\omega t + \psi_n). \tag{29}$$

En posant

$$\operatorname{tang} \Phi_n = \frac{2Pn\omega}{aBSV}, \tag{30}$$

on trouve facilement que la solution permanente de cette dernière équation peut se mettre sous la forme

$$\frac{dz}{dt} = \sum v_n \cos \Phi_n \sin(n\omega t + \psi_n - \Phi_n), \tag{31}$$

l'angle θ étant de même connu par l'équation

$$\theta = \frac{1}{V} \sum v_n \sin \Phi_n \cos(n\omega t + \psi_n - \Phi_n). \tag{32}$$

Si l'on remarque maintenant que la valeur moyenne dans la période du carré d'une série trigonométrique est égale à la moitié de la somme des carrés des coefficients de ses termes, on voit finalement que la valeur moyenne de c'_x dans la période aura pour expression

$$c'_{x_m} = c_{x_0} - \frac{B-\beta}{2V^2} \sum v_n^2 \sin^2 \Phi_n, \tag{33}$$

la vitesse V étant toujours connue par l'équation de sustentation

$$P = \frac{a}{2g} c_{z_0} S V^2. \tag{34}$$

La conclusion remarquable à déduire de la formule (33) est que, ainsi que nous l'avions prévu, sans que l'aile bouge ou se déforme, l'incidence étant, comme il a été dit, celle de la finesse, les divers harmoniques de la pulsation additionnent simplement leurs effets, chacun opérant comme s'il était seul dans la diminution de c'_{x_m}.

Notons que les pulsations secondaires étant généralement d'un ordre n élevé, c'est-à-dire à courte période, l'angle Φ_n qui leur correspond est grand et l'importance de leur effet se trouve ainsi accentuée, $\sin \Phi_n$ augmentant avec n en se rapprochant de l'unité.

L'importance de l'action des pulsations secondaires s'est d'ailleurs manifestée dans les expériences de laboratoire effectuées jusqu'ici. En recevant sur une aile fixe un courant d'air ondulé verticalement par les oscillations d'une série de persiennes, on a constaté, pour les oscillations correspondantes du courant d'air, un effet meilleur que celui qui résulte du calcul effectué avec le seul harmonique fondamental. Ce fait tient, à notre avis, à ce que les grandes oscillations provoquent dans le courant une turbulence certaine se traduisant par l'existence d'harmoniques secondaires qui additionnent leurs effets à celui de l'harmonique fondamental seul envisagé dans le calcul. Il ne convient donc pas de voir dans cette discordance, ainsi que certains l'ont pensé, une insuffisance du calcul, mais bien une analyse incomplète du phénomène.

Considérons maintenant, dans une pulsation verticale du vent, un harmonique quelconque d'intensité v'_n et de période $\frac{n\omega}{2\pi}$. Si cet harmonique agissait seul il devrait, pour permettre le vol à voile, présenter une certaine intensité v_n.

D'après le théorème de la superposition des effets, la condition de vol à voile sous l'action simultanée d'un certain nombre d'harmoniques peut s'écrire sous la forme simple

$$\sum \frac{v'^2_n}{v_n^2} = 1. \tag{35}$$

C'est ainsi que si l'on envisage, à titre d'exemple, l'action de 4 harmoniques qui, pour la simplicité du calcul, seront supposés de même intensité et si leurs périodes respectives sont de 1, 0,75, 0,5 et 0,25 seconde, il leur suffirait de posséder une intensité commune de $1^{m},13$ pour permettre le vol à voile d'un planeur chargé à 10^{kg} au mètre carré et de caractéristiques telles que celles que nous avons indiquées. Cette intensité s'abaisserait à $0^{m},80$ pour une charge au mètre carré de 5^{kg}.

Si l'intensité des pulsations n'est pas suffisante pour annuler la résistance aérodynamique moyenne, on voit facilement que la valeur de cette résistance a pour expression

$$c'_{x_m} = \left(1 - \sum \frac{v_n'^2}{v_n^2}\right) c_{x_0}. \tag{36}$$

Le redressement partiel ainsi obtenu de la résultante aérodynamique peut s'évaluer, ainsi qu'il a été dit, par l'inclinaison ϕ_m de cette résultante en arrière de la verticale. La tangente de cet angle n'est autre que *la résistance relative apparente*, immédiatement donnée par la formule

$$\operatorname{tang} \phi_m = \left(1 - \sum \frac{v_n'^2}{v_n^2}\right) \operatorname{tang} \phi_0. \tag{37}$$

La puissance nécessaire au vol se trouve, elle aussi, multipliée par le même facteur.

C'est ainsi qu'avec les 4 harmoniques précédemment envisagés, une intensité commune de $0^{m},50$ réduirait de 40 pour 100 l'inclinaison de la résultante aérodynamique ainsi, par suite, que la puissance nécessaire au vol.

En résumé, il semble bien que si un oiseau ne peut normalement voler à voile en utilisant uniquement les pulsations verticales du vent, il peut néanmoins y puiser une très notable fraction de l'énergie totale qui lui est nécessaire.

III

Le Vol par Battements d'une Aile en Air Calme

Bien que cette étude paraisse s'écarter du sujet qui nous occupe ici, nous avons pensé qu'en raison de l'analogie des phénomènes mis en jeu, l'analyse de la propulsion par battements et nos conclusions sur son rendement ne pourraient manquer d'intéresser le lecteur.

Les battements de l'aile d'un oiseau rameur en vol horizontal n'ont pas pour effet, ainsi que certains peuvent le penser, d'assurer directement sa sustentation, mais bien de propulser l'oiseau qui se soutient alors grâce à sa vitesse entretenue par les battements.

L'explication et l'analyse de la *propulsion* horizontale en air calme d'un oiseau par battements de ses ailes sont d'ailleurs en tous points semblables à celles qui montrent comment les pulsations verticales du vent peuvent redresser la résultante aérodynamique moyenne sur un oiseau qui les subit passivement.

Une réciprocité parfaite existe, en effet, entre les deux phénomènes, l'aile qui bat communiquant à la vitesse aérodynamique une oscillation périodique dans un plan vertical, tout comme le ferait une pulsation verticale du vent de même loi que le battement.

Certaines particularités du vol par battements, et notamment l'étude de son rendement, nous paraissent encore mal connues et nécessitent, de ce fait, un supplément d'analyse.

Dès 1908, nous avions indiqué, dans un article adressé à la Revue générale des Sciences, et par des considérations simples, que le rendement de ce mode de propulsion pouvait être très voisin de l'unité.

Cette étude ne fut pas publiée, le rédacteur de la Revue générale des Sciences ayant estimé le sujet trop nouveau et nos conclusions de nature à susciter des controverses auxquelles il n'a pas jugé opportun d'être mêlé.

Les procédés plus rigoureux d'analyse que nous employons actuellement vont nous permettre, en chiffrant très exactement le rendement, de justifier nos assertions.

Un point de l'aile battante d'un oiseau décrit, par rapport au corps, une courbe fermée de forme elliptique aplatie, mais nous ne retiendrons de ce mouvement complexe qu'un mouvement vertical joint à un mouvement d'oscillation de l'aile autour d'un axe parallèle à l'envergure, mouvement dont nous verrons l'importance à propos de la meilleure loi de variation de l'incidence.

Par ailleurs, le mouvement vertical de l'aile n'est pas exactement un mouvement de translation, surtout pour les points les plus rapprochés du corps.

Nous admettrons néanmoins qu'on peut, avec une précision qui suffit à notre analyse, assimiler le mouvement moyen de la partie active de l'aile à un mouvement de translation, auquel se superpose simplement le mouvement d'oscillation qui régit la loi des incidences.

Pour conserver à cette étude toute sa généralité, nous écrirons que la vitesse du battement par rapport au corps est définie en fonction du temps par la somme des termes d'une série de Fourier

$$v = \sum v_n \sin(n\omega t + \psi_n), \tag{38}$$

la période du battement étant $T = \dfrac{2\pi}{\omega}$.

Nous nous fixerons, comme précédemment, les variations de l'incidence en fonction de l'inclinaison θ de la vitesse aérodynamique par une loi linéaire

$$i = k\theta + i_0. \tag{39}$$

Nous admettrons dans nos calculs, sans que la justesse de nos conclusions s'en trouve pratiquement affectée, que l'amplitude de θ reste suffisamment petite pour que $\cos\theta$ soit assimilable à l'unité et $\sin\theta$ à θ. Des développements en série plus complets pour $\sin\theta$ et $\cos\theta$ ne conduiraient d'ailleurs qu'à une valeur du rendement *plus grande*, l'écart étant inférieur à 1 pour 100, même pour un maximum de θ de 20°.

La grandeur de la vitesse aérodynamique V sera d'autre part supposée constante, hypothèse que justifie la grande fréquence des battements de l'aile d'un oiseau.

Le mouvement de l'aile par rapport à l'air résulte, à chaque instant, de la composition de la vitesse horizontale V avec la vitesse v et la vitesse verticale périodique $\frac{dz}{dt}$ que possède le corps de l'oiseau, de sorte que l'on a, à chaque instant:

$$\theta = -\frac{1}{V}\left(v + \frac{dz}{dt}\right) = -\frac{1}{V}\left[\Sigma v_n \sin(n\omega t + \psi_n) + \frac{dz}{dt}\right]. \tag{40}$$

En suivant pas à pas la théorie de l'utilisation des pulsations verticales du vent, on voit immédiatement que l'équation du mouvement vertical du centre de gravité s'écrit

$$\frac{2P}{aBkSV}\frac{d^2z}{dt^2} + \frac{dz}{dt} + \Sigma v_n \sin(n\omega t + \psi_n) = 0. \tag{41}$$

En posant

$$\operatorname{tang} \Phi_n = \frac{2Pn\omega}{aBkSV}, \tag{42}$$

la solution permanente de cette équation peut s'écrire

$$\frac{dz}{dt} = -\Sigma v_n \cos \Phi_n \sin(n\omega t + \psi_n - \Phi_n), \tag{43}$$

d'où l'on déduit facilement

$$\theta = -\frac{1}{V}\Sigma v_n \sin\Phi_n \cos(n\omega t + \psi_n - \Phi_n). \tag{44}$$

D'après l'équation (10), on voit alors immédiatement que la condition de propulsion en vol horizontal s'écrit

$$c_{x_0} = (Bk - \beta k^2)\frac{1}{T}\int_0^T \theta^2 dt = \frac{Bk - \beta k^2}{2V^2}\Sigma v_n^2 \sin^2\Phi_n. \tag{45}$$

Ceci posé, le travail élémentaire dépensé par l'oiseau a pour valeur

$$d\mathfrak{T} = -\frac{a}{2g}c'_z S V^2 v dt. \tag{46}$$

La partie variable de c'_z étant $Bk\theta$, on voit que le travail dépensé pendant la période est connu par la formule

$$\mathfrak{T} = -\frac{a}{2g}SV^2 Bk\int_0^T \theta v\, dt = \frac{a}{4g}SVBkT\Sigma v_n^2 \sin^2\Phi_n. \tag{47}$$

Il est bien évident que le travail dû au poids de l'aile déplacée et aux forces d'inertie développées dans son mouvement n'a pas à être introduit dans le calcul, car ce travail est nul pour la période.

Or, en vol horizontal à la même incidence i_0, le travail utilisé pendant le temps T de la période est donné par l'expression bien connue

$$\mathfrak{T}_u = \frac{a}{2g} c_{x_0} S V^3 T, \tag{48}$$

d'où la valeur du rendement de la propulsion par battements:

$$\rho = \frac{\mathfrak{T}_u}{\mathfrak{T}} = \frac{2c_{x_0} V^2}{Bk \Sigma v_n^2 \sin^2 \Phi_n}. \tag{49}$$

Si nous tenons compte de l'équation de propulsion (45), nous obtenons finalement l'expression remarquable

$$\rho = 1 - \frac{\beta k}{B}. \tag{50}$$

Il apparaît tout d'abord que la loi suivant laquelle s'effectue le battement périodique de l'aile et l'incidence de vol sont sans influence directe sur le rendement, contrairement à ce qu'en pensent certains auteurs qui ont cru discerner une facteur essentiel dans le fait que la durée de la remontée de l'aile paraît plus faible que celle de la descente.

Cette inégalité des durées, relatée notamment par Marey, n'a d'ailleurs rien qui doive nous surprendre, si nous notons que la montée est la course de récupération pendant laquelle l'aile n'a qu'à se laisser entraîner par la réaction verticale de l'air. La descente, au contraire, est la course de travail pendant laquelle l'aile doit vaincre cette résistance, donc en se déplaçant d'un mouvement plus lent.

Ainsi que nous l'avons dit, β est à peu près constant pour toutes les ailes, tandis que B ne dépend que de l'envergure relative avec laquelle il croît jusqu'à une limite de 5,5.

L'aile, au point de vue de ses qualités aérodynamiques, n'intervient donc que par son envergure relative, le rendement étant d'autant meilleur que cette envergure est plus grande.

Si l'aile n'oscillait pas pour adapter son incidence, il conviendrait de prendre $k=1$. Avec les chiffres $\beta=1{,}3$ et $B=5$ que nous avons admis, le rendement serait de 0,74.

En réalité, il apparaît que le rendement est d'autant plus grand que k est plus petit, c'est-à-dire que les variations de l'incidence aérodynamique i sont moins prononcées; k étant alors plus petit que l'unité, l'oiseau doit suivre avec son aile les oscillations de la vitesse aérodynamique.

Nous verrons que k peut couramment atteindre des valeurs de 0,10 à 0,20, le rendement correspondant étant de 0,97 à 0,95.

Ainsi que nous l'avions depuis longtemps annoncé, il n'est donc en rien exagéré d'assigner au vol par battements un rendement généralement supérieur à 0,95.

Pour approfondir cette discussion, nous admettrons maintenant que la loi du battement soit sinusoïdale simple

$$v = v_{\mathrm{M}} \sin \omega t. \tag{51}$$

La vitesse V étant éliminée par l'équation de sustentation

$$P=\frac{a}{2g}c_{z_0}SV^2, \tag{52}$$

l'équation de propulsion (45) nous donnera à nouveau la condition (20) déjà obtenue à propos du vol à voile par les pulsations verticales du vent:

$$v_M^2=\frac{2g\tan\phi_0}{Bk-\beta k^2}\left[\frac{2P}{Sa}+\frac{gB^2k^2}{\omega^2 c_{z_0}}\right]. \tag{53}$$

Notons enfin qu'une condition supplémentaire, fixant l'amplitude des battements, lie v_M à ω. L'intégration de l'équation (51) donne, en effet, la loi des espaces, l'amplitude totale du battement étant $\frac{2v_M}{\omega}$.

En admettant que cette amplitude soit proportionnelle à l'envergure L de l'oiseau, on peut écrire:

$$\frac{2v_M}{\omega}=\lambda L, \qquad v_M=\frac{\lambda\omega L}{2}. \tag{54}$$

Les meilleurs rendements seront atteints, ainsi que nous l'avons vu, pour de faibles valeurs de k, donc pour de grandes valeurs de v_M et de ω, d'après les équations (53) et (54); k étant faible et ω grand, on peut, avec une précision suffisante, négliger le second terme de la parenthèse de l'équation (53). Dans ces conditions, l'angle Φ est très voisin de 90° et le maximum θ de θ, évalué en radians, est pratiquement égal à $\frac{v_M}{V}$.

En prenant, pour un oiseau de 1^{m} d'envergure,

$\Lambda=0{,}25, \quad \tan\phi_0=0{,}045, \quad B=5, \quad \beta=1{,}3, \quad c_{z_0}=0{,}5,$

nous obtenons ainsi:

		Valeurs de k.						
		0,10	0,15	0,20	0,25	0,30	0,35	0,40
ρ		0,974	0,961	0,948	0,935	0,922	0,909	0,896
θ_M^0		17°,4	14°,7	12°,5	11°,2	10°,3	9°,6	9°
$\frac{P}{S}=5^{kg}$	v_M m:s	3,85	3,26	2,76	2,48	2,28	2,12	2
	T sec	0,20	0,24	0,28	0,32	0,34	0,37	0,39
$\frac{P}{S}=2^{kg},5$	v_M m:s	2,72	2,30	1,95	1,75	1,61	1,50	1,41
	T sec	0,29	0,34	0,40	0,45	0,49	0,52	0,56

Il apparaît donc bien que des battements à grande fréquence peuvent s'effectuer avec un rendement compris entre 0,9 et 1, rendement nettement supérieur à celui des meilleures hélices.

Malgré cet avantage, l'application à un avion du principe des ailes battantes ne pourrait s'effectuer sans certaines difficultés pouvant compenser les avantages d'un excellent rendement. Il conviendrait, en effet, de tenir compte du rendement des transmissions, du poids du mécanisme de commande et surtout des difficultés d'équilibrage des forces d'inertie alternatives.

IV

L'Effet de Pulsations Aériennes Horizontales Reçues Latéralement par un Planeur a Ailes en M Aplati

M. Alexandre Sée, dès 1908, dans sa théorie *du vent louvoyant*, avait indiqué que des variations de direction alternatives du vent relatif horizontal agissant sur les ailes d'un planeur animées d'un balancement latéral, entraînaient une diminution moyenne de la résistance générale. Il avait ainsi expliqué par la géométrie comment la résultante aérodynamique sur une aile inclinée latéralement se redresse lorsque l'air vient la frapper obliquement par rapport à la direction de sa marche.

M. Alexandre Sée en avait conclu qu'il est nécessaire que l'oiseau change à chaque instant l'orientation de son plan alaire en synchronisme avec chaque changement de sens de la variation du vent. Il ajoutait: «C'est en cela que consiste le balancement latéral qui, malgré son apparence passive, est la clé du vol à voile».

En traitant par l'analyse mathématique les explications de M. Alexandre Sée, il apparaît qu'il n'est pas possible, comme il le prétend, que l'oiseau puisse, à tout instant, orienter sa voilure, soit d'une seule pièce, avec son corps, soit seulement ses ailes, et cela synchroniquement avec les pulsations latérales du vent.

La vitesse angulaire du balancement aurait d'ailleurs pour effet de dérober l'aile à l'action des pulsations qui cesseraient ainsi de pouvoir être utilisables.

C'est pour ces raisons que la théorie du vent louvoyant de M. Alexandre Sée n'a pas été prise en considération d'une façon sérieuse, malgré un point de départ très judicieux, auquel nous tenons à rendre hommage d'une façon très particulière, et qu'elle n'a pas été reprise ces temps derniers où cependant le problème du vol à voile a occupé nombre de techniciens.

Frappé de la forme si caractéristique de tous les oiseaux de mer qui volent à voile dynamiquement, dont les ailes se dessinent en M aplati, les branches de cet M faisant entre elles un angle voisin de 140°, il nous est apparu que cette forme devait jouer un rôle important. Cette forme présente en effet l'avantage d'offrir à tout instant aux vents relatifs obliques des surfaces d'ailes inclinées favorables au redressement de la résultante aérodynamique, avec une compensation des couples de roulis permettant à l'oiseau de subir et d'utiliser les rafales latérales en conservant son plan de vol.

S'il est exact que certains balancements sont observés, ils sont de peu d'amplitude, et il n'est pas prouvé qu'ils s'effectuent dans le sens envisagé par M. Sée.

L'analyse mathématique à laquelle nous avons soumis le problème ainsi posé a confirmé pleinement notre façon de voir et nous a révélé que, contrairement à ce que pensait M. Sée, ce n'est pas le balancement de l'oiseau qui est la clé du vol à voile par pulsations latérales, mais bien la forme si caractéristique en M aplati présentée par ses ailes.

En effectuant cette analyse, nous supposerons d'abord, pour la facilité de la discussion, que la pulsation du vent est représentée par une loi sinusoïdale simple, l'action simultanée d'une série d'harmoniques étant envisagée par la suite.

Soient:

i, c_x, c_z, l'incidence en radians au temps t d'une aile, comptée à partir de celle du minimum de traînée, et les coefficients de traînée et de poussée rapportés à la vitesse aérodynamique et à sa normale dans un plan perpendiculaire à

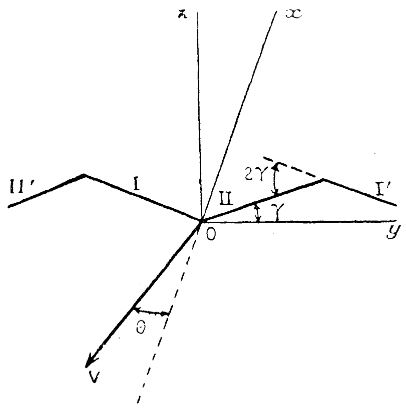


Fig. 2

l'aile contenant la vitesse aérodynamique, et affectés de l'indice 1 pour les parties d'ailes parallèles I et I' et de l'indice 2 pour II et II' (Fig. 2);

c'_x, c'_y, c'_z, les coefficients unitaires rapportés aux trois axes rectangulaires Ox, Oy, Oz, Oz étant vertical et Ox suivant la vitesse moyenne, positivement vers l'arrière;

$v = v_M \sin \omega t$, la vitesse au temps t de la pulsation de maximum v_M et de période telle que $\omega T = 2\pi$;

V, θ, la vitesse aérodynamique dont les variations de grandeur seront négligées et son inclinaison en radians au temps t sur Ox, dans le plan horizontal xOy;

P, S, γ, le poids du planeur, sa surface et l'inclinaison de ses ailes sur l'horizontale.

Notons d'abord que la constance à tout instant de la sustentation exige que l'aile conserve par rapport à la vitesse moyenne une incidence constante i_0; c'est ce que nous supposerons dans nos calculs.

En assimilant $\sin \theta$ à θ, $\sin i_0$ à i_0, $\cos \theta$ et $\cos i_0$ à l'unité, on trouve facilement que les incidences respectives des ailes au temps t seront

(55) $$i_1 = i_0 + \theta \sin \gamma$$

et

(56) $$i_2 = i_0 - \theta \sin \gamma.$$

En écrivant, comme nous l'avons précédemment admis,

$$c_x=\alpha+\beta i^2, \tag{57}$$

$$c_z=A+Bi, \tag{58}$$

et en désignant par c_x et c_z les valeurs de c_x et c_z pour l'incidence i_0, il vient:

$$c_{x_1}=c_{x_0}+2\beta i_0\,\theta\sin\gamma+\beta\theta^2\sin^2\gamma, \tag{59}$$

$$c_{x_2}=c_{x_0}-2\beta i_0\,\theta\sin\gamma+\beta\theta^2\sin^2\gamma, \tag{60}$$

$$c_{z_1}=c_{z_0}+B\,\theta\sin\gamma, \tag{61}$$

$$c_{z_2}=c_{z_0}-B\,\theta\sin\gamma. \tag{62}$$

Par projection sur les trois axes Ox, Oy, Oz, nous trouvons ensuite, avec le même degré d'approximation qui nous suffit dans ces calculs:

$$c'_x=\tfrac{1}{2}\,[c_{x_1}+c_{x_2}-\theta(c_{z_1}-c_{z_2})\sin\gamma], \tag{63}$$

$$c'_y=\tfrac{1}{2}\,[\theta(c_{x_1}+_{x_2})+(c_{z_1}-c_{z_2})\sin\gamma], \tag{64}$$

$$c'_z=\tfrac{1}{2}(c_{z_1}+c_{z_2})\cos\gamma. \tag{65}$$

La simple inspection de la première de ces trois équations montre dès maintenant que le redressement vers l'avant de la résultante aérodynamique s'opère, tout comme dans le cas des pulsations verticales, grâce à une composante horizontale de la poussée c_z.

En remplaçant c_{x_1}, c_{x_2}, c_{z_1}, c_{z_2} par leurs valeurs précédemment calculées et en négligeant dans l'expression de c'_y le terme en θ^3, nous trouvons sans difficulté les expressions suivantes:

$$c'_x=c_{x_0}-(B-\beta)\theta^2\sin^2\gamma, \tag{66}$$

$$c'_y=(c_{x_0}+B\sin^2\gamma)\,\theta, \tag{67}$$

$$c'_z=c_{z_0}\cos\gamma. \tag{68}$$

La sustentation étant constante à chaque instant, la trajectoire n'a pas d'ondulations verticales, l'équation de sustentation s'écrivant

$$P=\frac{a}{2g}\,c_{z_0}SV^2\cos\gamma. \tag{69}$$

D'autre part, grâce à la forme en M, les couples de roulis peuvent être annulés, les parties I et II des ailes engendrant un couple opposé à celui des parties I′ et II′. Nous supposerons cette compensation réalisée.

Le seul mouvement du planeur par rapport à sa route moyenne sera, en définitive, un mouvement de lacet, dans le plan horizontal xOy, défini par l'équation

$$\frac{P}{g}\frac{d^2y}{dt^2}=\frac{a}{2g}\,c'_ySV^2=\frac{aS}{2g}\,(c_{x_0}+B\sin^2\gamma)\theta V^2, \tag{70}$$

l'angle θ étant connu par la relation

$$\theta=\frac{1}{V}\left(v-\frac{dy}{dt}\right)=\frac{1}{V}\left(v_{\mathrm{M}}\sin\omega t-\frac{dy}{dt}\right), \tag{71}$$

d'où

(72) $$\frac{2P}{aSV(c_{x_0}+B\sin^2\gamma)}\frac{d^2y}{dt^2}+\frac{dy}{dt}=v_{\rm M}\sin\omega t.$$

En posant

(73) $$\tan\Phi=\frac{2P\omega}{aSV(c_{x_0}+B\sin^2\gamma)},$$

on trouve facilement que la solution permanente de cette équation peut s'écrire

(74) $$\frac{dy}{dt}=v_{\rm M}\cos\Phi\,\sin(\omega t-\Phi),$$

l'angle θ étant, par suite, donné par l'expression

(75) $$\theta=\frac{v_{\rm M}}{V}\sin\Phi\cos(\omega t-\Phi).$$

L'angle Φ, introduit dans les calculs, n'est autre que le décalage de phase du mouvement de lacet horizontal par rapport aux pulsations du vent; de même que dans le cas des pulsations verticales, c'est le facteur essentiel du redressement de la résultante aérodynamique. Ce décalage de phase croît quand la fréquence augmente, en prenant la valeur $\frac{\pi}{2}$ pour une fréquence infiniment grande.

Ceci posé, la valeur moyenne de c'_x dans la période sera, d'après l'équation (66),

$$c'_{x_m}=c_{x_0}-\frac{(B-\beta)\sin^2\gamma}{T}\int_0^T\theta^2\,dt,$$

c'est-à-dire

(76) $$c'_{x_m}=c_{x_0}-\frac{B-\beta}{2V^2}v_{\rm M}^2\sin^2\gamma\sin^2\Phi.$$

La condition de vol à voile horizontal est que c'_{x_m} soit nul, ce qui fait connaître la valeur de l'intensité $v_{\rm M}$ réclamée aux pulsations par la formule

(77) $$v_{\rm M}^2=\frac{2V^2c_{x_0}}{(B-\beta)\sin^2\gamma\sin^2\Phi}.$$

En remarquant que

(78) $$\frac{1}{\sin^2\Phi}=1+\frac{1}{\tan\!g^2\Phi}=1+\frac{a^2S^2V^2}{4P^2\omega^2}(c_{x_0}+B\sin^2\gamma)^2,$$

puis en remplaçant V^2 par sa valeur tirée de l'équation (69) de sustentation et en désignant $\frac{c_{x_0}}{c_{z_0}}$ par tang ϕ_0, nous obtenons finalement la formule cherchée:

(79) $$v_{\rm M}^2=\frac{2g\,\mathrm{tang}\,\phi_0}{(B-\beta)\cos\gamma\sin^2\gamma}\left[\frac{2P}{Sa}+\frac{g(c_{x_0}+B\sin^2\gamma)^2}{\omega^2c_{z_0}\cos\gamma}\right].$$

Le terme de la parenthèse contenant c_{z_0} est faible par rapport au premier, même pour de grandes valeurs de la période.

Il apparaît donc qu'en donnant à l'incidence i_0 la valeur qui correspond au minimum de tan ϕ_0, on obtiendra sensiblement les plus petites valeurs de v_M.

Par ailleurs, en annulant la dérivée de l'expression (79) par rapport à γ, on trouve, sans grande difficulté, que la valeur la plus favorable à donner à γ se tire de la formule

$$\frac{\sin^4\gamma}{(2-3\sin^2\gamma)\cos\gamma}=\frac{P\omega^2 c_{z_0}}{Sga(B^2+2c_{x_0}B)}\,. \tag{80}$$

On voit que l'angle γ croît avec la fréquence jusqu'à une valeur limite de 54°40′ correspondant à une fréquence infiniment grande. Dans les limites usuelles, de 10 à 30°, γ est pratiquement proportionnel à la racine carrée de la fréquence et à la racine quatrième de la charge au mètre carré.

Avec $c_{z_0}=0{,}5$, $c_{x_0}=0{,}02$, $B=5$, $a=1{,}225$, l'application de cette formule nous donne pour γ les chiffres indiqués dans le tableau ci-après:

$\frac{P}{S}$	T (en secondes)							
	0,75	1	3	6	8	12	15	20
kg 20....	$\gamma=49°$	47°	36°	27°	24°	20°	17°	15°
10....	$\gamma=47$	44	32	24	21	17	15	13
5....	$\gamma=44$	40	28	20	18	15	13	11
2,5...	$\gamma=40$	36	24	17	15	12	11	9

Remarquons par ailleurs que la valeur la plus favorable de γ est d'autant plus faible que l'altitude de vol est plus basse et que l'envergure relative de l'aile est plus grande, B et c croissant tous deux d'une façon analogue en fonction de l'envergure.

Discussion.—Il apparaît immédiatement que la formule (79) présente une grande analogie avec la formule (23) relative à l'utilisation des pulsations verticales.

Pour les très courtes périodes, pour lesquelles le terme contenant la fréquence devient négligeable, l'intensité réclamée aux pulsations horizontales n'est autre que celle qui serait nécessaire aux pulsations verticales divisée par $\sqrt{\cos\gamma}\,\sin\gamma$. Or le maximum de cette expression est 0,62 correspondant, ainsi qu'il a été dit, à $\gamma=54°\ 40'$. Il en résulte que, dans ce cas le plus favorable, l'intensité des pulsations doit être plus grande de 60 pour 100 que celle des pulsations verticales, cette majoration passant à 80 pour 100 pour $\gamma=40°$.

Par contre, pour des périodes croissantes, le terme de la formule (79) qui, dans la parenthèse, contient ω^2, reste toujours beaucoup plus petit que dans le cas des pulsations verticales.

Il en résulte que les pulsations à longue période, très peu profitables quand elles sont verticales, peuvent être utilisées quand elles sont horizontales.

Cette utilisation se fait d'autant mieux que l'observation du vent révèle, ainsi que nous l'avons dit, que la vitesse des pulsations horizontales croît avec leur période, leur accélération se trouvant de ce fait presque indépendante de la fréquence.

Remarquons à ce propos que si l'on admet que les pulsations verticales du vent sont surtout dues à la striction et à la contraction du courant, leur intensité est proportionnelle à chaque instant à l'accélération des pulsations horizontales, donc dépendrait peu de la période.

La conclusion remarquable à en tirer est donc bien que les pulsations verticales ne sont profitables que pour une très courte période puisqu'elles ne sauraient avoir pour les grandes périodes les intensités élevées réclamées par la formule, tandis que, pour un planeur à ailes en M aplati, toute la gamme des pulsations horizontales est utilisable, sans distinction des périodes, jusqu'à environ 10 secondes.

En admettant tang $\phi_0 = 0{,}045$, $c_{z_0} = 0{,}5$, $c_{x_0} = 0{,}02$, $a = 1{,}225$, $B = 5$, $\beta = 1{,}3$ et en admettant que, pour chaque valeur de la période, l'angle γ a sa valeur la plus favorable, l'application de la formule (79) nous donne les chiffres ci-après:

	$\frac{P}{S}$	Période (en secondes) 0,75	1	3	6	8	12
Albatros	10kg	$v_M = 3{,}54$	3,74	5,15	6,87	7,75	9,32
Pétrel	5	$v_M = 2{,}68$	2,86	4,20	5,62	6,45	7,78
Mouette	2,5	$v_M = 2{,}06$	2,24	3,44	4,65	5,35	6,50

Ce tableau montre que les intensités demandées aux pulsations horizontales sont de l'ordre de celles que certains auteurs ont prétendu relever dans les vents violents, car des vitesses de 5 à 7^m à la seconde, pour des périodes de 3 à 8 secondes, auraient été observées dans les fortes brises de mer.

De telles pulsations permettraient donc à elles seules le vol à voile.

L'intensité v_M réclamée croît avec la période, mais moins rapidement, semble-t-il, que l'intensité des pulsations observées dans le vent, jusqu'aux périodes de 8 à 10 secondes. Ce sont donc bien ces pulsations à longue période qui seront les plus efficaces.

Une confirmation remarquable de ces déductions se trouve dans la grandeur de l'angle γ observée le plus fréquemment sur les oiseaux bons voiliers et qui paraît être de 20° à 25°, indiquant ainsi que les ailes s'adaptent pour profiter au mieux des pulsations de 6 à 8 secondes.

Notons que, dans le cas du redressement partiel de la résultante aérodynamique, les formules (25) et (26) sont encore applicables.

Remarquons enfin qu'en ce qui concerne l'envergure relative de l'aile, des considérations toutes semblables à celles qui ont été développées à propos des pulsations verticales peuvent trouver place ici. L'avantage si marqué des grandes envergures relatives se trouve d'ailleurs confirmé si nous rappelons que des allongements de 10 sont courants chez les voiliers et que l'albatros, d'après Mouillard, pourrait avoir 5^m d'envergure pour une largeur d'aile d'au plus 25^{cm}, donc présenterait un allongement de 20.

LA SUPERPOSITION DES EFFETS DES DIVERS HARMONIQUES D'UNE PULSATION

Un raisonnement identique à celui qui a été donné à propos des pulsations verticales montre que si la pulsation envisagée est la somme d'une série d'harmoniques simples, tous ces harmoniques superposent leurs effets, chacun agissant, dans cette superposition, comme s'il était seul.

Supposons, en effet, que l'intensité de la pulsation soit la somme de la série de Fourier

$$v=\sum v_n \sin(n\omega t+\psi_n). \tag{81}$$

En posant

$$\operatorname{tang} \Phi_n=\frac{2Pn\omega}{aSV(c_{x_0}+B\sin^2\gamma)}, \tag{82}$$

on trouve que la résistance aérodynamique moyenne pendant la période sera

$$c'_{x_m}=c_{x_0}-\frac{(B-\beta)\sin^2\gamma}{2V^2}\sum v_n^2 \sin^2\Phi_n, \tag{83}$$

la vitesse V étant connue par l'équation de sustentation

$$P=\frac{a}{2g}c_{z_0}\, SV^2\cos\gamma. \tag{84}$$

La formule (83) montre bien que les différents harmoniques de la pulsation superposent leurs effets, ainsi qu'il a été annoncé.

Les formules (36) et (37) du redressement partiel de la résultante aérodynamique sont encore applicables.

L'ACTION SIMULTANÉE SUR UN PLANEUR A AILES EN M APLATI DE PULSATIONS HORIZONTALES REÇUES LATÉRALEMENT ET DE PULSATIONS VERTICALES

Ainsi qu'il a été dit à plusieurs reprises, il ne peut exister, dans un vent agité, de pulsations purement horizontales, toute pulsation horizontale engendrant nécessairement, par contraction ou dilatation de l'air en mouvement, une pulsation verticale qui peut d'ailleurs être accompagnée de pulsations d'autre origine.

Il convient donc de se demander si, pour un planeur à ailes en M aplati soumis latéralement à une pulsation horizontale, la présence simultanée d'une pulsation verticale peut, par interférence réciproque, atténuer l'effet de la pulsation horizontale, ou si, au contraire, les effets des deux pulsations s'ajoutent, chacune agissant pour son propre compte comme si elle était seule.

Pour répondre à cette question, nous reprendrons l'analyse de la page 691 en conservant la même notation et en nous reportant à la Figure 2.

Nous supposerons que, sous l'action d'une pulsation verticale agissant en même temps que la pulsation horizontale déjà considérée, la vitesse aérodynamique V se trouve, au temps t, inclinée de l'angle θ' sur le plan horizontal xOy, l'angle θ' étant considéré comme positif lorsqu'il augmente l'incidence.

Nous admettrons encore que l'aile conserve, par rapport à sa vitesse moyenne, une incidence constante i_0.

Les équations (55) et (56) donnant les incidences au temps t des parties d'ailes I et II deviennent alors, en assimilant cos θ' à l'unité et sin θ' à θ':

(85) $$i_1 = i_0 + \theta \sin\gamma + \theta' \cos\gamma,$$

(86) $$i_2 = i_0 - \theta \sin\gamma + \theta' \cos\gamma.$$

Par projection sur les trois axes Ox, Oy, Oz, nous obtenons les nouvelles valeurs de c'_x, c'_y et c'_z par les formules

(87) $$c'_x = \tfrac{1}{2}[c_{x_1} + c_{x_2} - \theta(c_{z_1} - c_{z_2})\sin\gamma - \theta'(c_{z_1} + c_{z_2})\cos\gamma],$$

(88) $$c'_y = \tfrac{1}{2}[\theta(c_{x_1} + c_{x_2}) + (c_{z_1} - c_{z_2})\sin\gamma],$$

(89) $$c'_z = \tfrac{1}{2}(c_{z_1} + c_{z_2})\cos\gamma.$$

Un calcul très simple nous donne alors:

$$\tfrac{1}{2}(c_{x_1} + c_{x_2}) = c_{x_0} + \beta(\theta^2\sin^2\gamma + \theta'^2\cos^2\gamma + 2\theta' i_0 \cos\gamma),$$

$$\tfrac{1}{2}(c_{z_1} + c_{z_2}) = c_{z_0} + B\theta'\cos\gamma,$$

$$\tfrac{1}{2}(c_{z_1} - c_{z_2}) = B\theta\sin\gamma.$$

En négligeant dans l'expression de c'_y les termes en θ^3, $\theta\theta'^2$ et $\theta\theta' i_0$, les expressions de c'_x, c'_y, et c'_z deviennent finalement:

(90) $$c'_x = c_{x_0} - (c_{z_0} - 2\beta i_0)\theta'\cos\gamma - (B - \beta)[\theta^2\sin^2\gamma + \theta'^2\cos^2\gamma],$$

(91) $$c'_y = (c_{x_0} + B\sin^2\gamma)\theta,$$

(92) $$c'_z = c_{z_0}\cos\gamma + B\theta'\cos^2\gamma.$$

Il apparaît dès maintenant que *les ondulations horizontales de la trajectoire ne dépendent que de la pulsation horizontale et les ondulations verticales que de la pulsation verticale*; il n'y a donc aucune interférence réciproque entre les deux mouvements.

Pour garder à cette analyse toute sa généralité, supposons, ce qui ne complique pas les calculs, que la pulsation horizontale et la pulsation verticale perçues par le planeur soient connues en fonction du temps par deux séries de Fourier:

(93) $$v = \sum v_n \sin(n\omega t + \psi_n),$$

(94) $$v' = \sum v'_n \sin(n\omega' t + \psi'_n).$$

Les angles θ et θ' sont alors respectivement donnés par les formules

(95) $$\theta = \frac{1}{V}\left(v - \frac{dy}{dt}\right) = \frac{1}{V}\left[\sum v_n \sin(n\omega t + \psi_n) - \frac{dy}{dt}\right],$$

(96) $$\theta' = \frac{1}{V}\left(v' - \frac{dz}{dt}\right) = \frac{1}{V}\left[\sum v'_n \sin(n\omega' t + \psi'_n) - \frac{dz}{dt}\right].$$

L'étude du mouvement horizontal a déjà été effectuée pages 692 et 696, ce mouvement n'étant pas influencé par la pulsation verticale.

En posant

(97) $$\operatorname{tang}\Phi_n = \frac{2Pn\omega}{aSV(c_{x_0} + B\sin^2\gamma)},$$

on trouve:

$$\theta = \frac{1}{V} \sum v_n \sin\Phi_n \cos(n\omega t + \psi_n - \Phi_n). \tag{98}$$

L'équation du mouvement vertical montre que l'équation moyenne de sustentation est encore

$$P = \frac{a}{2g} c_{z_0} S V^2 \cos\gamma, \tag{99}$$

les ondulations de la trajectoire étant déterminées par l'équation différentielle

$$\frac{2P}{aBSV\cos^2\gamma} \frac{d^2z}{dt^2} + \frac{dz}{dt} = \sum v'_n \sin(n\omega' t + \psi'_n). \tag{100}$$

Cette équation est tout à fait analogue à celle que nous avons intégrée page 683.

En posant

$$\operatorname{tang} \Phi'_n = \frac{2Pn\omega'}{aBSV\cos^2\gamma}, \tag{101}$$

on trouve finalement:

$$\theta' = \frac{1}{V} \sum v'_n \sin\Phi'_n \cos(n\omega' t + \psi'_n - \Phi'_n). \tag{102}$$

En évaluant par les formules (98) et (102) les valeurs moyennes de θ^2 et θ'^2 dans la période de chacune des pulsations, on voit que la résistance aérodynamique moyenne sera

$$c'_{x_m} = c_{x_0} - \frac{B-\beta}{2V^2} [\sin^2\gamma \sum v_n^2 \sin^2\Phi_n + \cos^2\gamma \sum v'^2_n \sin^2\Phi_n]. \tag{103}$$

Cette formule générale montre bien que non seulement la pulsation verticale superpose son effet à celui de la pulsation horizontale, mais que tous les harmoniques des deux pulsations additionnent également leurs effets, chacun agissant pour sa part comme s'il était seul.

Cette propriété de la superposition des effets permet de dire qu'aucune agitation intérieure du vent n'est perdue pour l'oiseau qui vole à voile.

Elle explique la facilité déconcertante avec laquelle les oiseaux franchissent, sans un battement d'aile, d'énormes distances, alors que l'intensité maximum des pulsations que nous pourrions percevoir nous paraîtrait insuffisante pour entretenir le vol à voile.

Pour poursuivre cette discussion, supposons les deux pulsations représentées chacune par un harmonique:

$$v = v_{\mathrm{M}} \sin \omega t, \qquad v' = v'_{\mathrm{M}} \sin(\omega' t + \psi').$$

Écrivons la condition de vol à voile $c'_{x_m} = 0$ en remplaçant $\sin\Phi_m$, $\sin\Phi'_m$ et V par leurs valeurs; il vient:

$$\frac{2g \operatorname{tang}\phi_0}{B-\beta} = \frac{v_{\mathrm{M}}^2 \cos\gamma \sin^2\gamma}{\frac{2P}{Sa} + \frac{g(c_{x_0} + B\sin^2\gamma)^2}{\omega^2 c_{z_0} \cos\gamma}} + \frac{v'^2_{\mathrm{M}} \cos^3\gamma}{\frac{2P}{Sa} + \frac{gB^2\cos\gamma}{\omega'^2 c}}. \tag{104}$$

Il apparaît que l'effet de la pulsation verticale se trouve atténué par l'inclinaison des ailes et d'autant plus que l'angle γ est plus grand.

Il semble donc qu'il n'y ait pas intéret à donner à un planeur devant profiter simultanément des pulsations horizontales et verticales des ailes présentant un angle γ trop accentué.

Supposons $\tan\phi_0 = 0{,}045$, $c_{z_0} = 0{,}5$, $c_{x_0} = 0{,}02$, $B = 5$, $\beta = 1{,}3$, $\gamma = 20°$, $\frac{P}{S} = 5^{\text{kg}}$, $a = 1{,}225$, $\omega = 1$, $\omega' = 6$, nous trouvons:

$$\left(\frac{v_M}{5{,}85}\right)^2 + \left(\frac{v'_M}{2{,}46}\right)^2 = 1,$$

c'est-à-dire:

v_M	1,5	3	3,5	4	4,5	5	5,5
v'_M	2,38	2,11	1,97	1,79	1,56	1,26	0,83.

Une pulsation verticale d'environ 1^{m} par seconde réduit l'intensité réclamée à la pulsation horizontale de $5^{\text{m}},85$ à $5^{\text{m}},30$ par seconde.

Il semble donc, ainsi que nous l'avions indiqué, que la plus grande partie de l'énergie nécessaire au vol à voile soit tirée des pulsations horizontales.

Cependant l'utilisation simultanée de tous les harmoniques à faible période existant dans les pulsations verticales doit être en même temps envisagée dans une explication rationnelle et complète du vol à voile.

V

L'Effet de Pulsations Aériennes Horizontales Reçues de Front par un Planeur

Dans une théorie dite *des montagnes russes*, certains auteurs ont cru voir une explication du vol à voile dans l'utilisation des variations horizontales d'intensité du vent suivant l'axe du planeur.

De simples considérations générales, telles que celles qui ont été données à ce sujet, sont impuissantes à établir l'efficacité de ce mode d'utilisation des pulsations horizontales et l'analyse mathématique du problème, quoique difficile, nous paraît indispensable pour fixer l'opinion à cet égard.

Il paraît tout d'abord que l'effet favorable ne peut se produire que grâce au mouvement vertical du centre de gravité du planeur et qu'ainsi les pulsations de courte période ne peuvent avoir d'efficacité.

Par ailleurs, dans une analyse un peu serrée de la question, il convient de tenir compte du mouvement horizontal périodique du centre de gravité, d'autant plus accentué que la période est plus longue. Il en résulte que les pulsations à trop longue période sont également inefficaces et que les seules pulsations utilisables paraissent être celles de période moyenne, de 4 à 10 secondes.

Le problème ainsi posé ne peut, avec une rigueur suffisante, se traiter directement par des méthodes de calcul analogues à celles que nous avons déjà utilisées. Nous écrirons d'abord les équations générales du mouvement et nous verrons ensuite comment en tirer les conclusions qui nous sont nécessaires.

Les vitesses étant positives dans le sens de la marche du planeur, soient:

V_m, V, θ, la vitesse aérodynamique moyenne et les valeurs au temps t de la vitesse aérodynamique et de son inclinaison sur l'horizontale, l'angle θ étant positif quand il augmente l'incidence de vol;

v, la vitesse au temps t de la pulsation horizontale reçue par le planeur;

$\frac{dz}{dt}$, $\frac{dx}{dt}$, les valeurs au temps t de la vitesse verticale du centre de gravité et de sa vitesse horizontale périodique.

Nous supposerons, comme précédemment, l'aile uniquement passive, l'incidence de vol étant à chaque instant $i=\theta+i_0$, et nous assimilerons $\cos\theta$ à l'unité et $\sin\theta$ à θ.

V et θ sont alors connus par les formules

$$V=V_m-v+\frac{dx}{dt}. \tag{105}$$

$$\theta=-\frac{1}{V}\frac{dz}{dt}. \tag{106}$$

Les coefficients unitaires c'_x et c'_z suivant l'horizontale et la verticale sont donnés par les formules (10) et (11), précédemment établies à propos des pulsations verticales, et dans lesquelles $k=1$, ce qui donne:

$$c'_x=c_{x_0}-(c_{z_0}-2\beta i_0)\theta-(B-\beta)\theta^2, \tag{107}$$

$$c'_z=c_{z_0}+B\theta. \tag{108}$$

Nous obtenons ensuite, en tenant compte de l'équation (106),

$$X=c'_x V^2=c_{x_0}V^2+(c_{z_0}-2\beta i_0)V\frac{dz}{dt}-(B-\beta)\left(\frac{dz}{dt}\right)^2, \tag{109}$$

$$Z=c'_z V^2=c_{z_0}V^2-BV\frac{dz}{dt}. \tag{110}$$

Une relation se prêtant mieux à la discussion se tire des deux précédentes par élimination de V^2, ce qui permet d'écrire:

$$X=Z\ \mathrm{tang}\,\phi_0+(B\ \mathrm{tang}\,\phi_0+c_{z_0}-2\beta i_0)V\frac{dz}{dt}-(B-\beta)\left(\frac{dz}{dt}\right)^2. \tag{111}$$

Remarquons maintenant que les équations du mouvement horizontal et du mouvement vertical sont respectivement:

$$\frac{2P}{Sa}\frac{d^2x}{dt^2}=-X, \tag{112}$$

$$\frac{2P}{Sa}\frac{d^2z}{dt^2}=Z-\frac{2Pg}{Sa}. \tag{113}$$

Il apparaît alors que la condition du vol à voile horizontal est que la valeur moyenne de la résistance X soit nulle dans la période, la valeur moyenne de la portance Z étant $\frac{2Pg}{Sa}$.

Il en résulte que, d'après l'équation (111), cette condition s'écrit, en toute rigueur:

$$\frac{2Pg}{Sa}\operatorname{tang}\phi_0+(B\operatorname{tang}\phi_0+c_{z_0}-2\beta i_0)\frac{1}{T}\int_0^T V\frac{dz}{dt}dt-(B-\beta)\frac{1}{T}\int_0^T\left(\frac{dz}{dt}\right)^2 dt=0. \tag{114}$$

Ceci posé, pour obtenir la relation générale liant le mouvement horizontal au mouvement vertical, il suffit de remplacer, dans la formule (111), X et Z par leurs valeurs tirées des équations (112) et (113), ce qui donne:

$$-\frac{2P}{Sa}\frac{d^2x}{dt^2}=\frac{2P}{Sa}\left(g+\frac{d^2z}{dt^2}\right)\operatorname{tang}\phi_0+(B\operatorname{tang}\phi_0+c_{z_0}-2\beta i_0)V\frac{dz}{dt}-(B-\beta)\left(\frac{dz}{dt}\right)^2. \tag{115}$$

L'intégration simultanée des équations régissant le mouvement horizontal et le mouvement vertical ne peut pratiquement s'effectuer.

Nous tournerons cette difficulté en partant, non pas de la pulsation du vent supposée connue, mais d'une loi périodique simple du mouvement vertical du centre de gravité et en recherchant quelle est la pulsation susceptible d'entretenir ce mouvement.

Le mouvement vertical est évidemment de même période que la pulsation du vent.

Nous supposerons $\omega=1$ correspondant précisément à une période de 6,28 secondes, ordre de grandeur des périodes les plus favorables, et nous écrirons:

$$\frac{dz}{dt}=b\sin t. \tag{116}$$

Pour évaluer la vitesse horizontale périodique, nous pouvons, sans erreur appréciable, puisque cette vitesse ne constitue qu'un terme correctif, remplacer dans l'équation (115) $V\frac{dz}{dt}$ par $V_m\frac{dz}{dt}$,

Le calcul rigoureux pourrait d'ailleurs, sans difficulté, se faire par intégration graphique. Remplaçons alors dans l'équation (115) $\frac{dz}{dt}$ et $\frac{d^2z}{dt^2}$ par leurs valeurs et intégrons en ne considérant que les termes périodiques, puisque les termes constants de $\frac{d^2x}{dt^2}$ doivent s'annuler lorsque la condition de vol à voile se trouve remplie.

Nous obtenons sans difficulté:

$$-\frac{2P}{Sa}\frac{dx}{dt}=\frac{2Pb}{Sa}\operatorname{tang}\phi_0\sin t-(B\operatorname{tang}\phi_0+c_{z_0}-2\beta i_0)b\,V_m\cos t+(B-\beta)\frac{b^2}{4}\sin 2t. \tag{117}$$

Ceci posé, l'équation (113) du mouvement vertical nous permettra d'évaluer quelle est, à chaque instant, la valeur de la vitesse aérodynamique susceptible d'entretenir le mouvement vertical arbitrariement fixé.

Cette équation s'écrit, en effet,

$$c_{z_0}V^2 - BV\frac{dz}{dt} - \frac{2P}{Sa}\left(g+\frac{d^2z}{dt^2}\right)=0, \tag{118}$$

d'où

$$V=\frac{B}{2c_{z_0}}\frac{dz}{dt}+\frac{1}{2c_{z_0}}\sqrt{B^2\left(\frac{dz}{dt}\right)^2+\frac{8Pc_{z_0}}{Sa}\left(g+\frac{d^2z}{dt^2}\right)}, \tag{119}$$

avec

$$\frac{dz}{dt}=b\sin t, \tag{120}$$

$$\frac{d^2z}{dt^2}=b\cos t. \tag{121}$$

Nous supposerons, comme précédemment, que l'incidence i_0 est celle de la finesse du planeur, cette incidence étant toujours, d'après l'équation (111), très voisine de l'incidence optimum. Nous prendrons donc $\tan\phi_0=0{,}045$, $c=0{,}5$, l'angle i_0 étant supposé de 8° et nous effectuerons les calculs pour $\frac{P}{S}=10$, $a=1{,}225$.

Pour estimer, en première approximation, la valeur à assigner à b, appliquons la condition (114) du vol à voile en négligeant le terme contenant en facteur la valeur moyenne de $V\frac{dz}{dt}$.

La valeur moyenne de $\left(\frac{dz}{dt}\right)^2$ étant $\frac{b^2}{2}$, nous voyons qu'en adoptant toujours $B=5$ et $\beta=1{,}3$, il conviendrait, avec ce degré d'approximation, de prendre $b=2$.

En réalité, le terme contenant la valeur moyenne de $V\frac{dz}{dt}$ étant positif, la valeur exacte de b est supérieure à 2 et nous verrons que $b=2{,}6$ vérifie la condition complète de vol à voile. On peut alors, avec cette valeur de b, tracer par points le diagramme de la vitesse aérodynamique V en fonction du temps dans l'intervalle de la période.

On trouve (Fig. 3):

t sec......	0	1	2	3	4	5	6	6,28
V m: s....	20,15	32,90	32,40	17,27	9,23	9,95	16,80	20,15

La courbe ainsi obtenue a la forme d'une sinusoïde à deux arches inégales et qui se trouve légèrement décalée par rapport à $\frac{dz}{dt}$. La valeur moyenne V_m de V est 20 m: s, sa valeur maximum est 33,80 m: s, atteinte pour $t=1{,}5$ sec, et sa valeur minimum est 8,50 m: s, atteinte pour $t=4{,}35$ sec. L'amplitude des variations de V est donc de 25,30 m: s.

Le décalage de phase moyen de la courbe de V par rapport à celle de $\frac{dz}{dt}$ est sensiblement $\psi = 0{,}3$ sec.

Pour éviter la détermination graphique, cependant facile, de la valeur moyenne de $V\frac{dz}{dt}$, assimilons, pour le calcul de cette valeur, la courbe des variations de V à une sinusoïde régulière décalée de ψ par rapport à celle représentant $\frac{dz}{dt}$ et d'ordonnée maximum $e = 12{,}65$ m: s au delà de la valeur moyenne.

La valeur moyenne de $V\frac{dz}{dt}$ dans la période sera

$$\frac{1}{T}\int_0^T V\frac{dz}{dt}dt = \frac{be}{2}\cos\psi = \frac{2{,}6 \times 12{,}65}{2}\,0{,}955 = 15{,}7.$$

On vérifie alors immédiatement que la condition (114) de vol à voile se trouve remplie pour la valeur de b que nous avons choisie.

Pour passer des variations de V à celles de la pulsation v, il convient tout d'abord d'évaluer $\frac{dx}{dt}$ à l'aide de l'équation (117) qui s'écrit:

$$\frac{dx}{dt} = 1{,}16\cos t - 0{,}11\sin t - 0{,}386\sin 2t.$$

Nous trouvons par points le diagramme ci-après (Fig. 3):

t sec.	0	1	2	3	4	5	6	6,28
$\frac{dx}{dt}$ m: s. . .	1,16	0,18	−0,29	−1,04	−1,06	0,64	1,34	1,16

La condition de vol à voile étant réalisée, l'aire totale de cette courbe est évidemment nulle.

La formule (105) nous permettra finalement de connaître l'intensité réclamée à chaque instant à la pulsation du vent, cette intensité étant connue par la relation

$$v = V_m - V + \frac{dx}{dt}. \tag{122}$$

En remarquant que $V_m = 20$ m: s, et en nous reportant aux valeurs déjà obtenues pour V et $\frac{dx}{dt}$, nous obtenons le diagramme de la Figure 3:

t sec.	0	1	2	3	4	5	6	6,28
v m: s.	1,01	−12,72	−12,69	1,69	9,71	10,69	4,54	1,01

La courbe des variations de v ainsi obtenue est une sinusoïde à arches légèrement inégales, la plus grande valeur absolue de v étant 14 m: s, atteinte pour

$t=1{,}5$ sec. Une telle intensité est trois fois supérieure à celle que l'on peut espérer rencontrer couramment dans le vent.

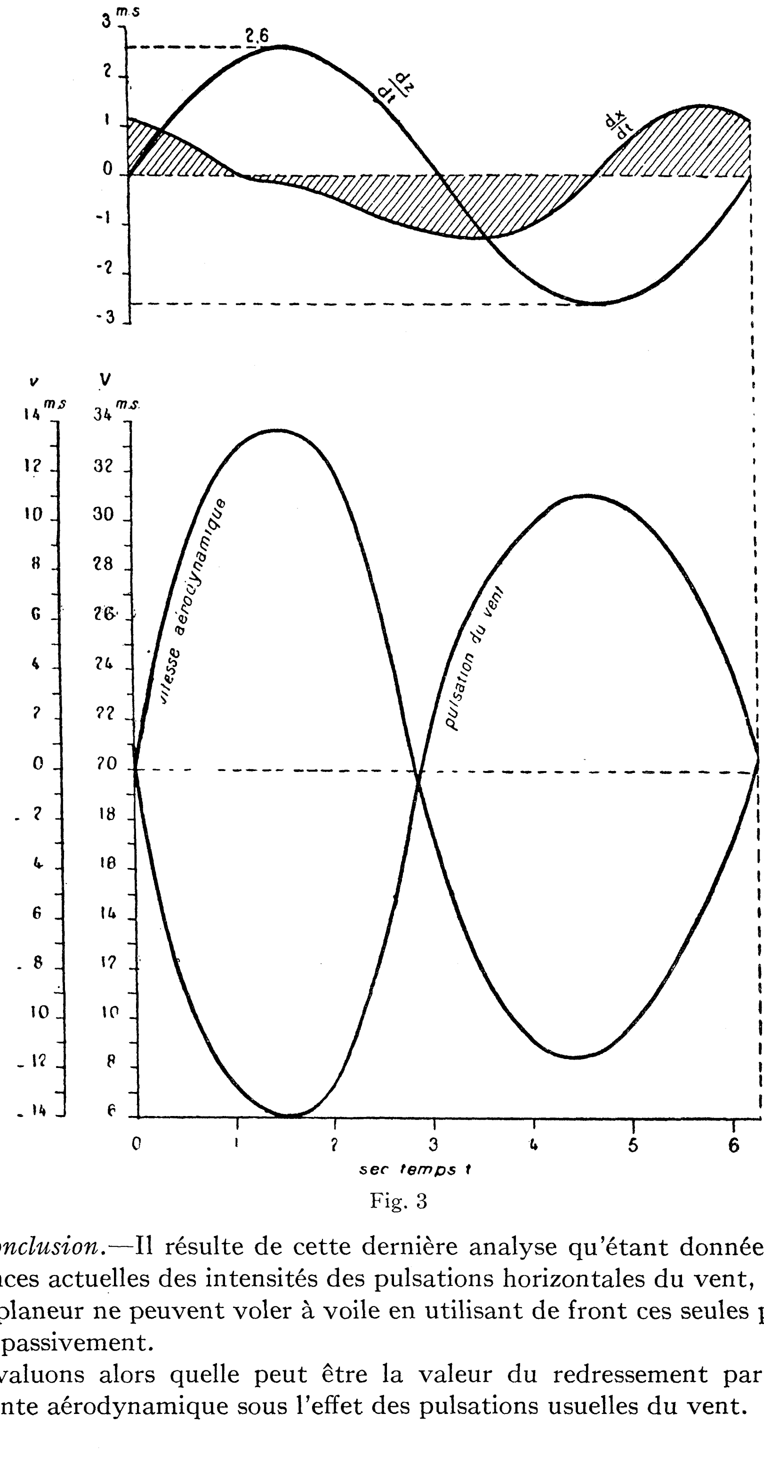


Fig. 3

Conclusion.—Il résulte de cette dernière analyse qu'étant données les connaissances actuelles des intensités des pulsations horizontales du vent, un oiseau ou un planeur ne peuvent voler à voile en utilisant de front ces seules pulsations reçues passivement.

Évaluons alors quelle peut être la valeur du redressement partiel de la résultante aérodynamique sous l'effet des pulsations usuelles du vent.

Soient X_m la valeur moyenne de la résistance qui subsiste, Z_m la valeur moyenne de Φ qui, ainsi que nous l'avons vu, est indépendante de l'intensité de la pulsation et égale à $\frac{2Pg}{Sa}$.

En admettant que la valeur maximum de $\frac{dz}{dt}$ soit proportionnelle à l'intensité de la pulsation, l'équation (111) montre que la diminution de la résistance est encore proportionnelle au carré de l'intensité de la pulsation du vent.

Si donc v'_M est l'intensité effective de la pulsation du vent, v_M l'intensité de la pulsation de même période qui permettrait le vol à voile, on peut écrire:

$$X_m = Z_m \left(1 - \frac{v'^2_M}{v^2_M} \right) \text{tang } \phi_0. \tag{123}$$

Mais la valeur angulaire du redressement se mesure par l'inclinaison ϕ_m, en arrière de la verticale, de la résultante aérodynamique moyenne, la tangente de cet angle n'étant autre que la résistance relative moyenne.

La relation (123) nous donne immédiatement

$$\text{tang } \phi_m = \frac{X_m}{Z_m} = \left(1 - \frac{v'^2_M}{v^2_M}\right) \text{tang } \phi_0. \tag{124}$$

C'est ainsi que, pour la période de 6,28 sec déjà envisagée, nous avons vu qu'il fallait que v_M atteigne la valeur de 14 m: s. Une pulsation de même période et de 5^m,50 d'intensité ne redresserait la résultante aérodynamique que du septième environ de la quantité qui serait nécessaire pour permettre le vol à voile.

L'énergie retirée de cette utilisation est donc peu appréciable.

Par ailleurs, les équations différentielles qui sont intervenues dans notre analyse ne sont pas linéaires.

Il en résulte que, théoriquement, si une pulsation est constituée par un certain nombre d'harmoniques, ces harmoniques ne doivent pas superposer simplement leurs effets.

Certaines approximations assez grossières permettent cependant de ramener ces équations à des types linéaires, ce qui fait que, pratiquement, la superposition doit s'opérer d'une façon satisfaisante.

Résumé Général

En résumé, une pulsation du vent peut toujours se décomposer, par rapport à un planeur, en trois pulsations élémentaires, la première verticale, le deuxième horizontale reçue perpendiculairement à sa route, et la troisième horizontale reçue de front.

Nous venons de montrer que les pulsations horizontales les plus intenses que l'on puisse espérer rencontrer dans le vent sont nettement insuffisantes pour permettre le vol à voile suivant le troisième procédé et ne peuvent, ainsi utilisées, communiquer au planeur une énergie appréciable.

Par contre, nous avons vu que les mêmes pulsations se révèlent d'une efficacité bien supérieure quand elles sont reçues latéralement par un planeur à ailes en M aplati.

C'est ainsi que, pour un planeur chargé à 10^{kg} au mètre carré, une pulsation horizontale de 5,50 m: s d'intensité et de 6 sec de période, reçue par le travers, redresserait la résultante aérodynamique de 65 pour 100 et seulement du septième environ de la quantité nécessaire quand elle est reçue de front.

Le planeur, pour profiter au mieux de l'agitation interne du vent, devrait donc s'orienter de façon à recevoir latéralement les pulsations horizontales les plus intenses.

Ainsi que nous l'avons vu, l'hodographe des pulsations n'est pas sphérique, mais a la forme d'un ellipsoïde dont le grand axe serait dirigé suivant le vent moyen, le plus petit axe semblant être le second axe horizontal.

Il en résulte que si un planeur, pour se placer dans les conditions les plus favorables, oriente son envergure parallèlement au vent, il ne recevra plus de front que des pulsations très faibles, n'apportant généralement pas le $\frac{1}{100}$ de l'énergie nécessaire à son vol.

Nous ne pourrions d'ailleurs prétendre que cette orientation soit la règle pour tous les oiseaux qui volent à voile, car l'effet favorable ne se trouve pas très atténué lorsque la pulsation principale est reçue obliquement par rapport au plan de symétrie du planeur.

C'est ainsi que si, dans un vent violent, la pulsation principale est inclinée à 60° sur l'axe du planeur, sa composante normale à l'axe n'est réduite que de 13 pour 100, et l'oiseau pourrait encore, dans ces conditions, puiser dans le vent la totalité de l'énergie dont il a besoin.

D'ailleurs, dans ce cas, l'oiseau peut, pour orienter sa route, recevoir alternativement cette pulsation tantôt d'un côté et tantôt de l'autre, ce qui expliquerait la marche en zigzags souvent observée du sol.

L'utilisation de toutes les pulsations verticales à courte période apporte enfin un complément d'énergie qui doit être considéré dans une explication complète du vol à voile.

Nous avons d'ailleurs montré qu'aucune gêne ni interférence réciproque ne pouvait se manifester dans l'utilisation simultanée des pulsations verticales et des pulsations horizontales reçues latéralement, ainsi que de tous leurs harmoniques.

Notons que les pulsations horizontales reçues de front ne peuvent non plus provoquer une interférence appréciable. Elles n'ont en effet aucune action sur l'utilisation latérale des pulsations horizontales puisque cette utilisation n'engendre aucun mouvement vertical du centre de gravité du planeur. D'autre part, les seules pulsations verticales efficaces sont à courte période, les pulsations reçues de front n'étant au contraire favorables que si leur période est suffisamment longue.

Cette simple sélection des périodes suffit, à elle seule, à empêcher toute interférence appréciable.

Cette superposition des effets de tous les harmoniques des pulsations, tant horizontales que verticales du vent, est l'une des particularités les plus remarquables du vol à voile dynamique.

L'irrégularité de l'agitation interne du vent peut laisser, en effet, supposer qu'un très grand nombre de tels harmoniques exercent leur action simultanée.

Si donc la discussion qui a été faite montre qu'une pulsation sinusoïdale, prise isolément, ne pourrait que dans des cas assez exceptionnels entretenir seule le vol à voile, il apparaît que la sommation des effets des pulsations en tous sens et de tous leurs harmoniques fournit l'explication rationnelle de la facilité d'évolution déconcertante des oiseaux voiliers dans le vent.

Notre analyse nous permet finalement de conclure que la sustentation sans battement des oiseaux dans le vent ne saurait plus être considérée comme mystérieuse, mais s'opère suivant des principes et des lois bien définis dans le cours de cette étude.

Toutefois si les oiseaux pratiquent avec une aisance si remarquable le vol à voile dynamique, il convient de ne pas oublier que la nature leur a donné la possibilité, grâce à la perfection et à la finesse de leur structure qu'aucune de nos machines n'a encore pu atteindre.

Nous nous permettrons, en terminant, de rappeler ici une phrase de M. Rateau dans une lettre qu'il nous avait adressée en mars 1924 à propos des recherches publiées ici:

«*Le problème du vol à voile commence à devenir bien clair.*»

THE DURATION AND LENGTH OF RUN REQUIRED BY SEAPLANES. AND FLYING BOATS "TAKING OFF" THE SURFACE OF THE WATER

By Mr. Alan Ferrier,

Flying Officer, Royal Canadian Air Force, Ottawa, Canada.

The problem which forms the subject of my paper to-day is one which, I think, will assume increasing importance as air traffic grows. It may be said that good performance in "taking off" is a most desirable, if not an absolutely essential, requirement for a seaplane, particularly in this country.

The immense forest and bush covered areas of Canada are, as everybody knows, dotted with innumerable lakes of all sizes and shapes, and though these areas are exceedingly inhospitable for land machines, they are, except for the long winter, a veritable paradise for seaplanes.

The map in Fig. 1 shows roughly the bush covered areas of Canada and also an imaginary future airline from coast to coast. This airline can be traversed by seaplanes and flying boats throughout, with the exception that the portion over the Rocky Mountains would, under present conditions, be hazardous.

Although nature has been so bountiful in her provision of alighting places, she does not always arrange to place a nice big lake just where one wants to alight. Let us see how this fact affects the problem of forest fire protection.

More and more work is being done every season to fight the forest fire, our most dreadful scourge. The success of detection and report of fires by air and radio telegraphy is an established fact, and now a sort of flying fire engine is being developed for the actual work of fire suppression. For such a machine, it is essential to alight as near a fire as possible, for it may take from two hours to half a day to travel four miles in some of this bush, and every hour is precious. But, and here comes the rub, it is equally essential to be able to "get off" again, and fires do not always occur near big lakes. It is obvious, therefore, that of two machines, one of which has good "take off" qualities and moderate air speed, and the other has poor "take off" and high air speed, the former is superior from the forestry point of view, because it can make use of more alighting places and can get closer to more fires, and the forestry point of view in Canada is a vital one.

Aspects of the Problem

Having, I hope, shown the need for the investigation of this problem, let us proceed to study it in an elementary manner.

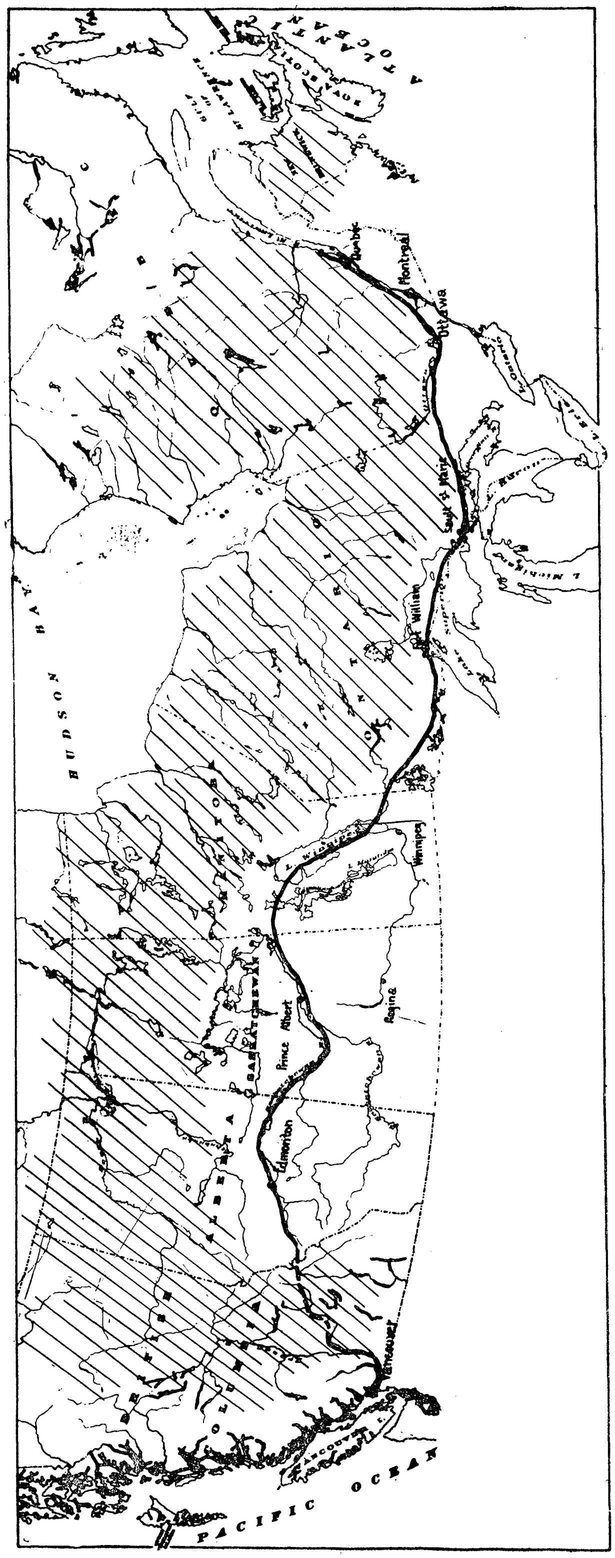

HUDSON BAY
PACIFIC OCEAN
Montreal
Ottawa
Winnipeg
Regina
Prince Albert
Edmonton
ALBERTA

Fig. 1

For a given minimum flying speed, the "take off" qualities of a seaplane or flying boat depend on the following:

(1) The thrust of the airscrew.

(2) The air resistance of the wings.

(3) The parasite resistance of struts, wires, engine, etc.

(4) Lastly, and very important, the water resistance and hydroplaning qualities of the floats or hull.

Airscrew

The thrust of an airscrew, which depends on airscrew efficiency, forward speed and engine horse power, has been the subject of considerable experimentation. Fig. 2 shows the curve of thrust of a typical airscrew turning at constant revolutions plotted on a speed base. Starting at about 4.5 lbs. per h.p. the thrust falls away at a gradually increasing rate. All airscrews of good design would have a similar thrust curve, that is to say, airscrews of fixed pitch. A good variable pitch airscrew has not yet been put on the market, but the ideal is one

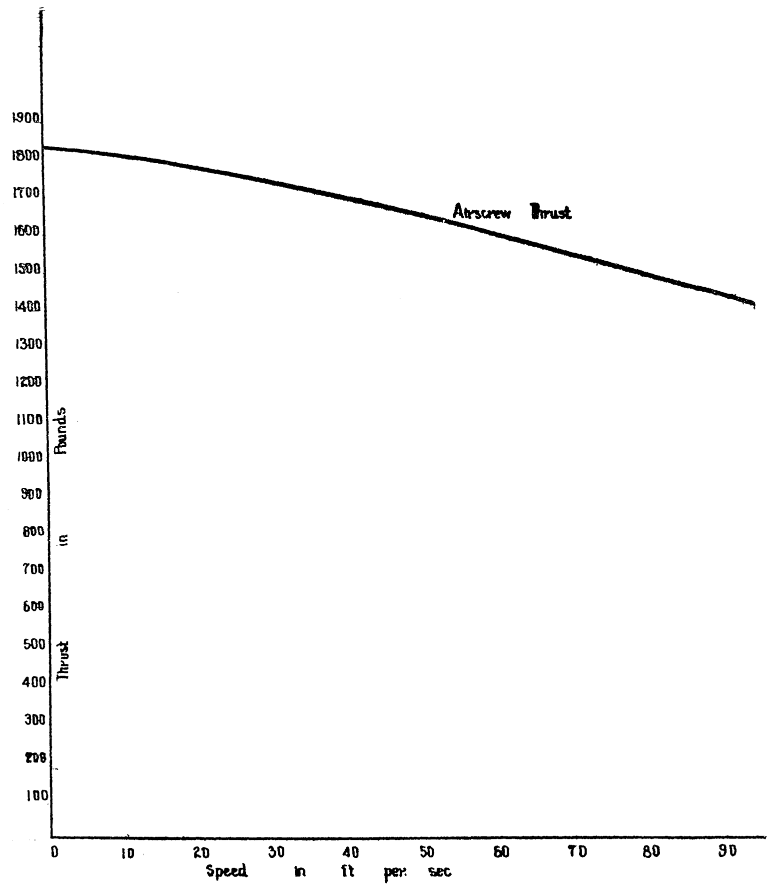


Fig. 2

whose pitch is so controllable that the maximum thrust corresponding to the best efficiency of the airscrew can be maintained at all forward speeds within the practical range. The thrust curve of such an airscrew would be a horizontal straight line.

Float or Hull Resistance

The water resistance and planing qualities of seaplane floats and flying boat hulls have also been the subject of very considerable experimentation. Fig. 3 shows two typical resistance curves on a base of speed, one for a two-float seaplane, and the other for a single float flying boat.

When the engine throttle is opened, a flying boat tends to dip her nose and the resistance rises very rapidly until a certain definite hump in the curve is reached. As the speed increases, the seaplane takes less of a plowing attitude, but the resistance continues to rise until it reaches a maximum at what is called the hump speed. Then the resistance falls off rapidly. This is when the machine is planing, that is to say, she is no longer supported on the water by flotation but by the dynamic action of the forces on her planing bottom.

It is to be noted that when the results shown in Fig. 3 were obtained, the models were forcibly held at a constant angle of attack from a little below the hump speed onwards, and the moment required to hold them thus was measured.

As the speed increases, of course, more and more of the load is taken by the wings, so that at the actual instant of "take off" the water resistance of the hull disappears. The curves shown take this into account, the load on the hull having been decreased in proportion to the square of the speed. The air resistance of the hull increases, but this is also taken into account.

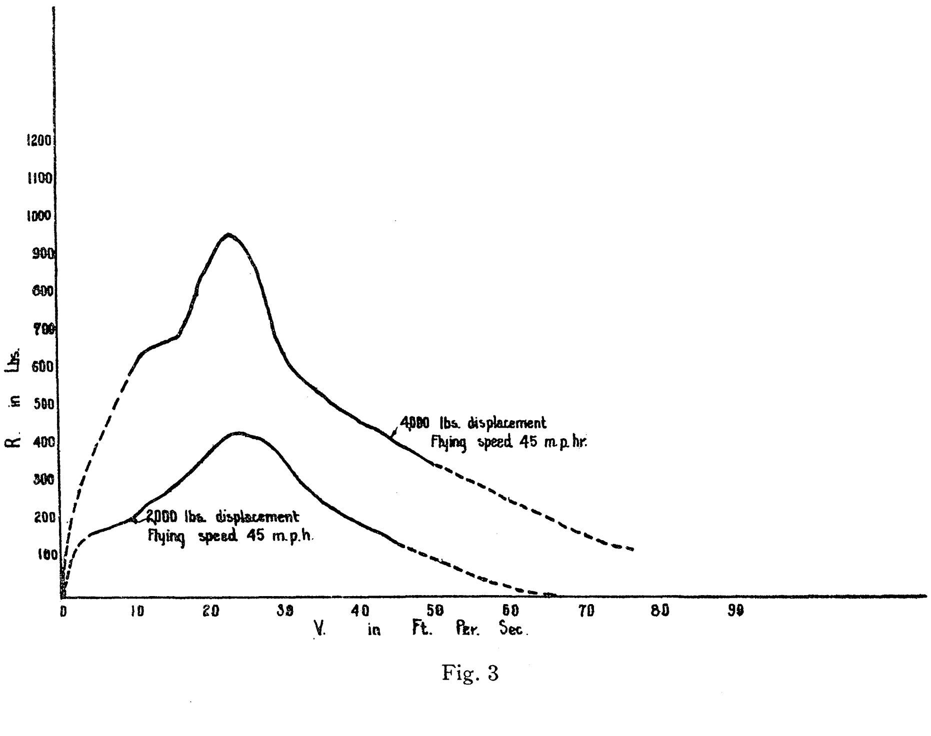


Fig. 3

It may be noticed that the curves are dotted at the beginning and end, and this is intended to mean that they have been exterpolated. In looking through a considerable number of papers on the subject, it was found that information on the resistance at low speeds was meagre and the same may be said of the resistance at speeds approaching flying speeds, but in the latter case, the curves are not so complicated and may be exterpolated with fair confidence.

The experimenters seem to have devoted most of their attention to resistance at and around the hump speed, and a very important speed it is, but it is hoped that in the future, accurate data on the resistance throughout the whole range of speed will be made available.

One can imagine the pioneer seaplane designer breathing a sigh of relief as he discovered that his curve of available thrust overtopped his curve of resistance at hump speed. The curves shown are average ones and represent typical resistance curves for machines of the displacement and minimum flying speed stated. For a rough estimate of "take off" these curves may be used with fair confidence for any machine of similar displacement and flying speed, for judging from the published results of numerous experiments by the Froude National Tank, it takes a very large, not to say freakish, variation from orthodox form to cause any great departure from the mean resistance curve. In support of this contention a slide kindly loaned by Squadron Leader Hume* shows $\frac{\text{Resistance}}{\text{Displacement}}$ curves for 5 hulls. It is remarkable how 4 out of 5 of the curves hang together. Another slide shows two more resistance curves with attitude of floats also recorded.

Planes and Parasite Resistance

So many data regarding these have been obtained and published that it is not necessary to dwell on the subject at any length. In the analysis which follows, it is assumed that the planes remain at a constant angle of attack and that the parasite drag coefficient and the wing drag coefficient can be lumped for simplicity's sake into one coefficient, and that the combined air resistance will vary as the square of the speed. It is admitted that this assumption is a source of error, but it will be seen later that in the method to be presented, the variation in parasite resistance may be taken into account.

There are available experimental data regarding the moments required to hold a machine at the optimum angle of attack. These moments may be translated into terms of lift and drag of the tail, and a correction made to the parasite resistance figure. This refinement will entail much additional labour, however, and with a really well designed hull, would not be of great importance. For it is now possible to design a hull or float which automatically takes up the required attitude, and in fact, some modern specifications call for this property as an essential in the design.

*Mathematical Analysis

The acceleration of an aircraft is dependent on the excess of thrust available over the total resistance at any instant, and this excess must, in any case, be

*Not reproduced in this paper.

sufficient to accelerate the craft up to the speed required for normal flight. The equation for the acceleration of a flying boat may be stated as follows:

$$\frac{dV}{dt}=\frac{g}{W}\left\{\phi(V)-k_x\frac{\rho}{g}AV^2-f(V)\right\}, \tag{1}$$

where $\phi(V)$ is a term for the thrust of the airscrew expressed as a function of the forward speed; k_x is the combined plane and parasite resistance coefficient, and A is the area of the wings. The weight of air $\frac{\rho}{g}$ is expressed in terms of slugs per cubic ft. and $f(V)$ is a term for the combined water and air resistance of the hull expressed as a function of the forward speed.

The most obvious way to get the time of run required to accelerate from rest to flying speed would be to put

$$T=\frac{W}{g}\int_0^{V_f}\frac{1}{\phi(V)-k_x\frac{\rho}{g}AV^2-f(V)}\,dV,$$

and integrate.

But on examining the hull resistance curves, it is obvious that even if it can be found, the function connecting resistance and velocity will be a most complicated one. The longer way round may be the shorter way home, so let us try the elementary way.

From the expression

$$V=u+ft, \tag{2}$$

that is final velocity = initial velocity plus acceleration multiplied by time,

on giving t an increment Δt we get

$$V+\Delta V=u+f(t+\Delta t) \tag{3}$$

$$\Delta V=f\Delta t,$$

therefore

$$\Delta t=\frac{\Delta V}{f}=\frac{\Delta V}{\frac{dV}{dt}}, \tag{4}$$

and we have a value for $\frac{dV}{dt}$ in equation (1).

The procedure from here in order to find the time to reach taking off speed is one of step by step integration, and the calculation is best done in tabular form.

Now in order to get at a way to determine the distance run in order to attain take off speed, we can start from the expression

$$V^2=U^2+2fS, \tag{5}$$

that is, the final velocity squared is equal to the initial velocity squared plus twice the acceleration times the distance.

From (5) we get

(6)
$$S=\frac{V^2-U^2}{2f}.$$

Now give V a little increment ΔV, then

(7)
$$S+\Delta S=\frac{(V+\Delta V)^2-U^2}{2f},$$

from which we get

(8)
$$\Delta S=\frac{2V\Delta V+\Delta V^2}{2\frac{dV}{dt}}.$$

Again the result required is arrived at by means of step by step integration.

Before going any further to illustrate the practical application of the results just obtained, it would be as well to point out how much more difficult the problem of "taking off" is with a seaplane than with a land machine.

With a land machine, the rolling resistance of the undercarriage is something that can be dealt with comparatively easily, but with the seaplane or flying boat, the water resistance of the hull depends on so many different factors that it defies pure mathematical analysis. In the first place, all theory breaks down in rough water. Then there is the question of suction behind the steps and the tendency of the water to cling to the sides of the hull, both of which tend to prevent the hull from unsticking. There are many other factors which are too numerous to mention in a paper of this scope.

Practical Application to a Particular Case

The following is an example worked out for a machine of 6000 lbs. displacement, with an engine of 400 h.p., and with a "getting off" speed of 70 feet per second, or nearly 48 miles per hour.

Fig. 4 shows curves of airscrew thrust, wing and parasite resistance and hull resistance, all plotted on a speed base. The height of the heavily shaded area represents hull resistance, and the height of the two shaded areas combined represents total resistance. The ordinates between the shaded areas and the thrust curve represent the excess of thrust over resistance at the corresponding speed, and if this be multiplied by g and divided by W, we get the $\frac{dV}{dt}$ of equation (1).

The Table shows in convenient form the carrying out of the step by step integration; it requires no explanation, but column No. 11 may be of interest, as it brings out forcibly the importance of cleanness of air structure design. Nearly half of the total run required takes place during the acceleration from 50 to 70 feet per second. The total run and not the time elapsed is the important item.

On consulting Fig. 4 again, it will be noticed that wing and parasite resistance are preponderant over this range, and also that the airscrew thrust drops more and more rapidly. This leads to the conclusion that, rather contrary to expectation, cleanness of air structure design is of as great importance as hull or float design if a short "take off" be required. It also emphasizes the need for the much sought after variable pitch airscrew.

The results obtained in the example shown are low, but they are possible under favourable conditions. In other words, the imaginary machine would be of exceptionally clean design and low "taking off" speed.

Another point brought out by study of the quantities in the last two columns of the table, is that one can expect a very quick "take off" from a machine with a low wing loading. But a low wing loading argues a large flying structure, with consequent increase in wing and parasite resistance. Therefore, as is usual in practical engineering problems, a compromise has to be arrived at between the machine with very low resistance and a high getting off speed, and a machine with higher resistance and a very low getting off speed.

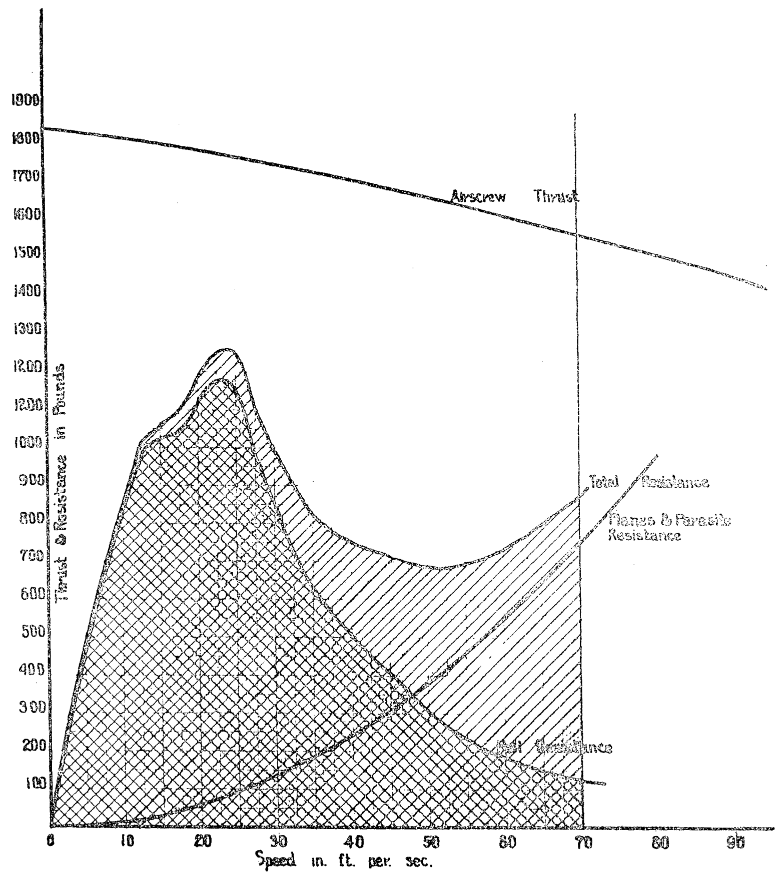


Fig. 4

TABLE

1	2	3	4	5	6	7	8	9	10	11	12
V ft./sec.	$\Delta\ V$ ft./sec.	Excess Thrust	$\frac{dv}{dt}=$ Excess Thrust $\times\frac{g}{w}$	$\Delta t=\frac{\Delta V}{\frac{dv}{dt}}$	$\Sigma\Delta t$	$2V\Delta V$	ΔV^2	$2V\Delta V + A V^2$	$\frac{\text{Col. 9}}{2\frac{dv}{dt}}=\Delta S$	$\Sigma\Delta S$	$\frac{\Sigma\Delta S}{S}$ as a %
1	2	1728	9.275	.216	.216	4	4	8	.43	.43	
3	2	1512	8.12	.246	.462	12	4	16	.99	1.42	
5	2	1332	7.15	.280	.742	20	4	24	1.68	3.10	
7	2	1165	6.26	.319	1.061	28	4	32	2.56	5.66	
9	2	1025	5.50	.364	1.425	36	4	40	8.64	9.30	
11	2	890	4.77	.419	1.844	44	4	48	5.04	14.34	2.53%
13	2	780	4.18	.478	2.322	52	4	56	6.70	21.04	
15	2	740	3.97	.504	2.826	60	4	64	8.06	29.10	
17	2	705	3.78	.529	3.355	68	4	72	9.53	38.63	
19	2	637	3.42	.585	3.940	76	4	80	11.70	50.33	
21	2	560	3.00	.666	4.606	84	4	88	14.66	64.99	11.45%
23	2	511	2.74	.730	5.336	92	4	96	17.50	82.49	
25	2	515	2.76	.725	6.061	100	4	104	18.08	101.29	
27	2	605	3.25	.615	6.676	108	4	112	17.2	118.49	
29	2	717	3.85	.520	7.196	116	4	120	15.6	134.09	
31	2	790	4.25	.470	7.666	124	4	128	15.05	149.14	26.4%
33	2	851	4.57	.438	8.104	132	4	136	14.88	164.02	
35	2	907	4.86	.411	8.515	140	4	144	14.80	178.82	
37	2	936	5.02	.399	8.914	148	4	152	15.15	193.97	
39	2	952	5.11	.391	9.305	156	4	160	15.65	209.62	
41	2	960	5.15	.388	9.693	164	4	168	16.30	225.92	40%
43	2	965	5.17	.387	10.080	172	4	176	17.00	242.92	
45	2	970	5.20	.385	10.465	180	4	184	17.70	260.62	
47	2	973	5.22	.383	10.845	188	4	192	18.40	279.02	
49	2	975	5.23	.383	11.231	196	4	200	19.10	298.12	
51	2	973	5.22	.383	11.614	204	4	208	19.9	318.02	56.3%
53	2	960	5.15	.388	12.002	212	4	216	21.0	339.02	
55	2	945	5.07	.395	12.397	220	4	224	22.1	361.12	
57	2	922	4.95	.404	12.801	228	4	232	23.4	384.52	
59	2	895	4.81	.416	13.217	236	4	240	25.0	409.52	
61	2	865	4.64	.431	13.648	244	4	248	26.7	436.22	77%
63	2	830	4.45	.450	14.098	252	4	256	28.8	465.02	
65	2	790	4.24	.472	14.570	260	4	264	31.1	496.12	
67	2	750	4.02	.498	15.068	268	4	272	33.8	529.9	
69	2	710	3.81	.525	15.593	276	4	280	36.7	566.62	

In conclusion, I wish to express my thanks to Squadron Leader Hume of the Royal Canadian Air Force for his very kind help to me in preparing this paper.

THE INFLUENCE OF MATHEMATICS ON THE DEVELOPMENT OF NAVAL ARCHITECTURE

By Mr. W. J. Berry,

Director of Naval Construction, The Admiralty, London, England.

The employment of ships for various purposes dates back to very early ages, but the science of Naval Architecture is of comparatively modern growth. Ships have advanced in size, speed, equipment and structural strength, but the progress from the primitive log or bundle of reeds used by the ancients to the 100-gun ship of the eighteenth century was effected wholly by methods of trial and error.

It was then impossible to predict the draught at which a new ship would float, or ensure that she would possess sufficient stability and satisfactory sea-going qualities. In consequence changes in design frequently proved to be the reverse of improvements, and it was commonly necessary to add girdling around the waterline of ships found defective in stability.

One of the first indications of the application of scientific principles to ship construction is found in Pepys' diary, which records that in 1666 "Mr. Deane . . . then fell to explain to me the manner of casting the draught of water which a ship will draw, beforehand, which is a secret the King and all admire in him; and he is the first that hath come to any certainty beforehand of foretelling the draught of water of a ship, before she is launched."

The method used by Mr. (afterwards Sir Anthony) Deane of calculating the displacement of ships is unknown; but it appears that about 1700 this was effected by dividing the body by equidistant sections, calculating the area of each and thence obtaining the displacement by some rough process of quadrature. There is, however, no record of any such calculations, and it is probable they were but rarely performed.

Shortly before 1700 two books on Naval Architecture appeared. One, by the Chevalier Renaud, was the outcome of a series of conferences held at Paris by order of Louis XIV, with the object of placing Naval Architecture on a more scientific basis. This book included a theory of resistance since known to be incorrect, which led to a controversy in which Huyghens, James and John Bernoulli took part. Renaud's principles were disproved, but no satisfactory theory of resistance was then substituted. The second book, by Paul Hoste, dealt with the stability, oscillations and resistances of ships; and although the theories given there were imperfect they were in advance of any previously propounded.

In 1746, however, the subject of stability was treated mathematically by Bouguer in his *Traité du Navire*. Bouguer gave three distinct methods of obtaining the displacement; and showed that the initial stability of a vessel depended on the relative positions of the centre of gravity and the metacentre. The book expresses certain erroneous ideas, particularly in regard to resistance and oscillations; but is notable as the first treatise in which the metacentric height was mentioned as the principal criterion of stability.

A further advance in the general theory was made on the appearance in 1775 of a work on the construction of ships by the Swedish Constructor—Admiral Frederick Henry de Chapman. This was translated into French by Vial du Clairbois, 1779, and into English by Dr. Inman, 1820. It contains the first published record of the use of Simpson's rules for approximate quadrature, and the calculations of displacement, centre of buoyancy, and metacentre given in the book closely resemble those made at the present day.

The importance of developing the theory of Naval Architecture was then fully appreciated in France; and a Committee appointed to examine Chapman's treatise presented a favourable report upon it. The defect of the book lies in the erroneous law of resistance formulated, in which, however, Chapman had merely reproduced the theory originally propounded by Newton and repeated by D'Alembert in 1744 and Leonard Euler in 1773. It assumed that the normal pressure on an element of surface was proportional to the product of the area, the square of the speed and the square of the sine of the angle between the surface and the direction of motion. With the usual notation the pressure varied as $A . V^2 Sin^2 \theta$ and the resistance resolved along the direction of motion $A . V^2 Sin^3 \theta$.

French scientists were doubtful of this law, and experiments were made 1763 and 1767 by the Chevalier de Borda, and in 1775 by d'Alembert, Condorcet, and the Abbé Bossut. The results did not confirm the law, but on the other hand no correct theory was established in its place. Towards the close of the eighteenth century the subject was further investigated by Romme and by several French Committees with inconclusive results. It was only in the latter half of the nineteenth century that the work of W. and R. E. Froude established the theory of resistance on a sound basis.

In contrast to the repeated efforts that were being made in France to encourage the scientific study of Naval Architecture very little interest in it had been taken in England since the time of Phineas Pett and Sir Anthony Deane (1600 to 1680). This was largely due to the prejudice of the constructors of the day, who, although good practical men were uninstructed in scientific matters, and inclined to view the application of all theories of ship construction with great mistrust. No steps were taken by the Government either to improve the scientific education of Naval constructors or to interest mathematicians in Naval Architecture; and ships continued to be built and designed by rule-of-thumb. As a result English ships of war were for a long period inferior in design to those of other nationalities; and such improvements as were effected followed frequently on the capture of a French ship which was taken as a model for subsequent construction in England. A good feature of English ships was their excellent workmanship; but this alone could not compensate for the effects of their inferior design.

Near the end of the eighteenth century interest in the subject was awakened by the establishment of a Society for the Improvement of Naval Architecture, which was in existence from 1791 to 1799. Few, if any, members of this Society were directly concerned with the construction of ships; but the Society, although short-lived, was indirectly instrumental in establishing the first British School of Naval Architecture. An important feature of the Society's work was its initiation of experiments on resistance*. These were carried out by Colonel Beaufoy, a member of the Society, largely at his own expense, and were improvements on those previously undertaken by French investigators. The resistances of bodies of various geometrical forms were determined, and the results established within limits the validity of model experiments—a principle that has since been found extremely useful and important.

Colonel Beaufoy also verified by experiments on models the theory of stability enunciated by G. Atwood, a mathematician, of which an account is contained in the Philosophical Transactions of the Royal Society 1796 and 1798. Atwood's method consisted of finding the transference of buoyancy from the wedge of emersion to that of immersion, from which he deduced the righting lever of stability. Atwood considered the metacentric method to be applicable only to evanescent angles of inclination, which is correct so long as exact results are required; but he did not realize that for ordinary ships the stability levers (*i.e.*, the arms of the righting couples) calculated from the metacentre are very nearly correct up to about 20° of heel. His papers are, however, of great interest and value; for he was the first to investigate stability at large angles (except Bouguer, who fell into certain errors in this respect), and the principles underlying his methods are still in use. The *Captain*, which capsized in the Bay of Biscay in 1871, is a tragic instance of a ship possessing ample metacentric height which was nevertheless deficient in range of stability.

Owing to the more general application of scientific principles, progress in Naval Architecture has been more rapid since 1800; and it will be convenient to terminate this brief account of the early history of the subject at that date, and to deal separately below with the development of its various branches.

Stability

In the earlier theories of ship stability attention was confined to inclinations about an axis fixed in direction. This case is the simplest and most important; nevertheless problems occur involving three-dimensional stability (*i.e.*, for rotation about any axis whatsoever), and it forms a necessary part of the complete theory of the subject.

The general case was fully investigated in 1822 by Charles Dupin in his *Mémoire sur la Stabilité des Corps Flottants*. He defined the locus of the centre of buoyancy, when a ship is inclined in all possible directions with constant displacement, as the surface of buoyancy; and showed that the stability depends entirely on the form of this surface and the position of G, the ship's centre of gravity, relative to it.

*V. Trans. Inst. Naval Arch. 1910 (A. W. Johns).

The tangent plane at any point of this surface is parallel to the corresponding waterplane, the line of action of the buoyancy being the normal to the surface. The stability of the ship is exactly the same as that of a solid shaped to the surface of buoyancy and allowed to roll on a smooth horizontal plane; the positions of equilibrium being determined by all the normals, *G.B.*, through *G* to the surface.

The stability at any position of equilibrium depends on the situation of *G* relative to the centres of curvature of the principal normal sections. If *G* is below the lower centre of curvature the ship is absolutely stable; if *G* is above the higher centre there is absolutely instability; but if *G* lies between the two centres, the ship is stable for certain inclinations and unstable for others, and the stability is termed "relative".

The principal radii of curvature of the surface are equal to the principal moments of inertia of the waterplane area divided by the volume of displacement; and the corresponding axes of inclination are perpendicular to the principal axes of inertia. The stability is therefore closely connected with the indicatrix of the surface of buoyancy, which is similar and similarly situated to the momental ellipse of the waterplane. It has been shown later that if a confocal ellipse be constructed, the squares of whose axes are proportional to the principal radii of curvature diminished by the distance *G.B.*, the resultant couple is represented in magnitude and direction by the diameter of this ellipse conjugate to the axis of inclination.

Dupin's results were confirmed by E. Guyou, who in his *Théorie du Navire* treated stability as a dynamical problem, and calculated the work done in inclining by estimating directly the gain in potential energy of ship and surrounding water. He first showed that for variations of draught without change in direction the energy is a minimum when the ship's weight is equal to her displacement, thus confirming the fundamental law of Archimedes. For varying inclinations at constant displacement, he proved that the potential energy is equal to the displacement multiplied by the perpendicular distance from *G* to the tangent plane to the surface of buoyancy. The locus of the foot of this perpendicular is a surface termed the "podaire", which constitutes a polar diagram of potential energy, or dynamical stability.

The normals from *G* to the podaire coincide with those to the surface of buoyancy and indicate the positions of equilibrium. At each such point the radius from *G* is stationary; the stability of the ship then depends on this radius being a maximum, minimum, or neither, the ship being respectively unstable, stable, or relatively stable. To render this clear, Guyou introduced the conception of a gradually contracting sphere of water, having its centre fixed at *G*. When the sphere has contracted sufficiently, portions of the podaire emerge forming small islands which correspond to absolute instability; later, small lakes remain, indicating positions of absolute stability. At an intermediate stage "passes" occur where two lakes or two islands just coalesce; and the stability is there "relative".

An interesting theorem follows immediately, viz., that the combined number of positions of absolute stability and instability exceeds that of the positions of relative stability by two. Guyou's method can also be applied to the oscillations of a ship; on adjusting the radius of the sphere to represent the total energy (kinetic+potential), it is evident that the inclination of the ship is limited by the water-edge, whilst the effect of resistance is gradually to contract the sphere and dry up the water.

Reverting to stability in a fixed direction, a solution of the problem had been provided by Bouguer (for small angles) and Atwood (for finite angles). By Atwood's method it was possible to calculate the righting levers at several inclinations, and to graph them on a base of angle. The resulting "curve of stability" is the standard method of exhibiting the large-angle stability of ships. The salient features are its initial slope (proportional to the metacentric height), its maximum ordinate and corresponding abscissa, the range or angle at which the ordinate vanishes, and its area (proportional to the dynamical stability).

A convenient method, based on Atwood's formula, of calculating the ordinates of the curve was given by F. K. Barnes in 1861; but his process was laborious, and it is now usual to calculate the levers at a series of constant angles for varying displacements, obtaining "cross-curves" from which the ordinary curves are readily deduced. This is done expeditiously by the aid of an integrating instrument capable of recording moments as well as areas. The sections of the ship around which the pointer is run in turn are spaced in accordance with a rule by Tchebycheff in which the ordinate multipliers are all equal. This rule reduces the number of readings necessary, and also ensures greater accuracy than could be obtained if Simpson's Rule were applied to the same number of sections. In this connection it is to be observed that a great variety of rules for approximate quadrature are employed in ship calculations, all of which are based on the integration of parabolic curves of various orders; in addition to Simpson's rules and combinations of them, the unequal-spacing rules proposed by Tchebycheff & Gauss are occasionally employed.

The subject of stability lends itself well to study by geometrical methods, and further investigations in France have proceeded on those lines. The loci of the centres of buoyancy, the centres of flotation, and the pro-metacentres (the evolute of the first locus) are all closely connected with stability; and in 1862 Reech obtained the following simple formulae for the co-ordinates $(x.y)$ of the centre of buoyancy: $x=\int \rho \cos \theta.d\theta$; $y=\int \rho \sin \theta.d\theta$; where θ is the angle of indication and ρ is the radius of curvature of the curve of buoyancy, which is readily calculable at any given waterline. From these co-ordinates the righting levers can be immediately deduced.

In 1870 an expression for the curvature of the curve of flotation was obtained by Émile Leclert; which result has a close bearing on the variation of metacentric height with the draught of the ship. The connection between the English and French points of view, *i.e.*, between the curve of stability and the locus of pro-metacentres, was clearly shown in 1871, by White and Johns who also traced the various loci for a variety of prismatic solids.

The effect of the admission of water to certain compartments of a damaged ship has recently received much attention; and a large number of calculations have been and are still being undertaken with a view to rendering ships as nearly as possible unsinkable, and reducing their liability to capsize when holed. This aspect of stability has an important influence in the design of both warships and merchant ships; and the regulations issued by the Board of Trade on the spacing of bulkheads are directly based on considerations of this nature. The precise methods of calculation and the approximations devised by Welch and others with the object of reducing the arithmetical labour, are of technical rather than mathematical interest.

The stability of a vessel carrying liquid having a free surface is a problem of practical importance in view of the large number of ships containing oil fuel in bulk. Up to small angles of inclination account can be taken of the mobility of the liquid by raising its centre of gravity to the centre of curvature of its path as the ship heels. There is an exact analogy between the actual G and virtual G of the fluid mass, and the centre of buoyancy and metacentre respectively of the ship's buoyancy. The virtual rise of G is proportional to the moment of inertia of the free surface of the liquid, *i.e.*, it depends on the extent of the free surface. A small quantity of liquid, *e.g.*, water or oil, carried in a shallow tank and possessing a large free surface may thus render a ship initially unstable, although in such a case the stability would be recovered at a small angle.

From the foregoing remarks it will be evident that a fairly complete mathematical theory on the stability of ships has now been developed. Naval Architecture has thereby benefited; for it is possible to ensure in a new design any degree of stability that experience has shown to be necessary.

Oscillation and Steering

Several attempts were made by the early writers on Naval Architecture to investigate the rolling motion of a ship in a seaway. They were all based on assumptions which have since proved incorrect; and it was not until 1861 that W. Froude propounded a theory of rolling which accounted for the observed facts and has since been generally accepted.

Froude first considered the motion of a small thin layer of water on the surface of a simple wave. The motion of the layer is known, and the forces acting upon it must account for (*a*) its horizontal and vertical accelerations, (*b*) its inclination to the horizontal, (*c*) its rate of distortion. Neglecting the last, which is very small, it is evident that if the layer of water be replaced by a small raft of the same size and mass, the forces on the raft, which are unchanged by the substitution, will tend to keep its deck parallel to the slope of the wave, and to preserve its displacement constant. The net effect on the raft is to alter the apparent direction and magnitude of gravity. Froude confirmed this experimentally by observing that the bob of a pendulum suspended from a raft remained relatively at rest when the raft was rocked among waves, *i.e.*, the string always placed itself perpendicularly to the wave slope.

A ship differs from the raft, inasmuch as it extends vertically and horizontally into the water, and this intrusion must in some degree disturb the natural motion of the waves. Froude did not attempt any exact investigation on the ship, but came to the conclusion, based on general considerations, that the forces on the ship would tend to keep her mast perpendicular to a modified wave surface which he termed the "effective wave surface". This resembled the actual surface of the wave in all respects, except that it had a smaller height and slope; it could be regarded as the sub-surface of the wave at a level which was undetermined but lay somewhere between the keel of the ship and the waterline.

Denoting the effective wave slope at any instant by ϕ, the inclination of the ship (θ) is virtually reduced to $\theta-\phi$, and the righting lever (GZ) is that corresponding to the angle $\theta-\phi$ in still water. Neglecting the effect of resistance, the differential equation of rolling motion can be immediately expressed. If, moreover, the angles are sufficiently small, $GZ=m\ (\theta-\phi)$ and the equation of motion takes the simple form

$$K^2\ddot{\theta}/g+m\theta=m\phi=m\alpha\ \sin\ (qt+\epsilon),$$

where m is the metacentric height, K the polar radius of gyration of the ship's mass, and α is the maximum slope and $2\pi/q$ the period of the wave, assumed simply sinusoidal.

The equation is of a well known type, and from its solution it can readily be deduced that the angle of roll (θ) is compounded of a "free roll" in the period of the ship—$2\pi K/(gm)^{\frac{1}{2}}$, and a "forced roll" in the period of the wave—$2\pi/q$. The amplitude of the latter is determinate, and it becomes indefinitely large when the periods of ship and wave synchronize.

Several points of considerable practical importance can be deduced from the foregoing investigation:

(*a*) The natural period of roll (*i.e.*, in still water) depends on K and m; it can be lengthened by winging weights or by reducing the metacentric height. Usually only the latter is feasible.

(*b*) Heavy rolling is caused when the periods of ship and wave are equal or nearly equal.

(*c*) Rolling is a minimum when the ship's period is much longer than the wave period. This consideration generally limits the metacentric height desirable in a ship, for excessive stability would lead to heavy rolling in a quick "uncomfortable" period.

(*d*) A moderate amplitude of rolling would also result if the ship's period could be made very short, and the vessel would then move like a raft, keeping her deck approximately parallel to the wave surface. This is generally impracticable so far as rolling is concerned, but is of importance in connection with pitching, which is governed by the same laws as rolling. The natural period of pitching is generally shorter than that of rolling: it can be reduced by keeping weights amidships, and the smaller it is the less the probability of the vessel shipping seas over bow or stern.

Froude also studied the laws of rolling resistance by means of still water experiments on ships and models. These were rolled artificially to as large an angle as possible, and the roll was then allowed to die out, the successive angles attained being observed. He deduced that the decrement of roll in a single swing or half-oscillation, say, from port to starboard, is represented by $a\theta+b\theta^2$ where θ is the mean of the initial and final angles in the swing and a and b are approximately constant. The existence of the empirical coefficients a and b can be accounted for by assuming that the former is due to the generation of waves and the latter is caused by frictional and eddy resistances varying as the square of the velocity.

Terms representing the resistance can be added to the equation of rolling motion; but in its general form no analytical solution can be obtained. Froude in 1875 devised a graphical method of solution which was applicable even when the stability lever GZ is not a simple function of θ. In the simpler case when the metacentric law holds good and GZ is proportional to θ, an approximate analytical solution can be obtained. In 1896 Bertin gave a simple method of calculating the successive maximum angles of roll, making the assumption usually adopted in France that the whole of the decrement can be expressed by the term $b\theta^2$.

If, on the other hand, the b term is neglected and the a term retained, the equation is linear and can be readily solved. It is found that the general effect of the resistance is to reduce and limit the magnitude of the "forced roll" near synchronism; it also causes the "free roll" to die away, although this does not materialize in an actual seaway, owing to irregularities in the waves. From this brief account of W. Froude's investigations, it appears that his theory of rolling, although incomplete and founded on assumptions that are difficult to justify entirely, has led to important practical conclusions with considerable influence on ship design. Subsequent experience has confirmed his results and so justified his hypotheses; nevertheless it is desirable on general scientific grounds that a more comprehensive theory, taking into account combined rolling, pitching, yawing and dipping oscillations should be established. In 1898 an investigation on these lines was undertaken by Kriloff with great mathematical skill and ingenuity; but it is not easy to deduce any general results from his analysis, though a number of particular cases were dealt with completely by him. Much attention has recently been devoted to the mathematical theory of the motion of aeroplanes; and it is to be hoped that the corresponding problem for ships may one day be as completely investigated.

An important addition to Froude's original theory was the discovery of the augmented rolling resistance caused by bilge keels. A comparison of rolling observations taken on a ship originally completed without bilge keels, and later fitted with them, was made by White in 1894. Both stillwater experiments and experience at sea showed that the bilge-keels were more efficacious in limiting rolling than had been supposed; and evidently the calculated pressures on them due to rolling had been under-estimated. A reason for the discrepancy was given in 1900 by G. H. Bryan, who showed that a ship when rolling induced an opposite counter-current of water, which would appreciably increase the pressure in the bilge-keel. Numerical results were obtained for several forms of bilge, and it was found that the increases of velocity and pressure were very large in a ship having a full midship section. The reduction in rolling due to bilge keels may also be partly due to the greater mass of the surrounding water set in motion by the roll; some of this water leaves the ship and does not return, and the consequent transference of energy causes a degradation of the rolling.

The interaction of rolling with other movements of the ship has been investigated in a few simple cases, a few of which may be cited. In 1864 W. J. M. Rankine showed that rolling is accompanied by forced dipping oscillations of half the rolling period. Their amplitude would be large only in ships having sloping sides near the waterline; and it is known that such vessels frequently roll "uneasily", due probably to their enforced rise and fall in the water. In 1905 A. W. Johns showed that resistance to rolling is considerably augmented by the ship's motion ahead; experience in all classes of ships, particularly vessels of high speed, such as destroyers, confirm this conclusion. In 1920 K. Suyehiro investigated mathematically and experimentally the motion of a vessel undergoing simultaneous rolling and pitching oscillations; he found that these motions were necessarily accompanied by yawing, *i.e.*, by oscillations in a horizontal plane, owing to the existence of gyrostatic terms in the equation of motion.

A valuable contribution to the theory was made in 1894 by R. E. Froude, who considered the general case of a ship rolling in a non-synchronous and irregular swell. He obtained the interesting result that if two identical ships entered simultaneously the same swell, the difference between their motions (due to differing initial conditions) was equal to the free roll of a similar ship in still water. Such a free roll would quickly degrade and disappear; so that in a short time the two original ships would have the same motion. It follows that the rolling of a ship in any swell is uninfluenced except at the start by her initial motion. This result considerably simplifies the problem of determining the greatest angle to which a ship will roll, and thence estimating her liability to capsize. In the same paper Froude showed how to estimate the maximum angle of roll in certain simple cases.

Before leaving the subject of rolling it should be remarked that a full application of the theory is dependent on our knowledge of the size and motion of the waves encountered by ships at sea. The observations recorded by Lieutenant Paris, Dr. Scoresby, Vaughan Cornish and others are very valuable in this connection; further and more exact data are, however, needed in order to estimate with any exactitude the minimum range of stability required to render a vessel safe under all conditions of service.

The influence of the form of rudder and stern on the steering and manoeuvring qualities of ships and the forces brought into action when a vessel is turning under helm has long been the subject of rough experimental investigations; but there were few accurate data available until recently. The absence of reliable information on the subject has led to a systematic series of model experiments, originally suggested by White, being undertaken at the National Physical Laboratory, Teddington, under the direction of G. S. Baker. Some valuable reports have already been published, and when the experiments have been completed by trials on ships, the collected data should throw some light on the problem of steering.

It appears improbable that this question can be dealt with by purely mathematical investigations, but such methods are of assistance in co-ordinating experimental data and deducing the forces on the ship which cannot be directly measured. The results of older trials have been analysed by Watts and Hovgaard, and interesting and useful information has thereby been obtained.

The relatively simpler case of airships has been mathematically investigated and a criterion expression obtained by which, on substituting coefficients obtained from model experiments, it can be ascertained if a form is a stable one when in flight. Experience with such vessels has, however, been so limited that it has been impracticable to confirm the results obtained from theoretical considerations. The scale effect, if any, remains undetermined.

Resistance and Propulsion

As stated in the first section of this paper the early estimates of the resistance caused by motion in a fluid were based on a law which was later disproved by the experiments carried out by Beaufoy and others. Further resistance experiments on ship-shaped forms were undertaken in 1834; from these Scott Russell deduced that resistance was caused by both viscosity and wave formation. He was, however, mistaken in his views on the character of the waves, and also in his theory that for favourable propulsion water-lines should be given certain mathematical forms.

In 1869 a Committee of the British Association recommended that resistance experiments should be carried out on full-sized ships; and W. Froude, who was a member of that Committee, also proposed a series of experiments on models with a view to comparing their results with those obtained in a ship.

Froude's suggestion was at first not agreed to, and even derided as being a waste of time; eventually with the support of Sir E. J. Reed it was approved by the Admiralty. This led to the establishment of the first model experiment tank at Chelston Cross, Torquay; and about this time W. Froude formulated a theory of resistance which is still accepted and is the foundation of the present methods of powering ships.

Froude divided the resistance into three parts, viz., that due to (*a*) skin friction, (*b*) wave making, (*c*) the formation of eddies. The last is small in all well-formed ships, and it is necessary to include it with the second; the term "residuary resistance" is sometimes applied to the combination of all resistances other than frictional.

It was assumed by Froude that the frictional and residuary resistances are independent of one another so that they can be estimated separately and added. This assumption greatly simplifies his theory of resistance; and although it has not been directly justified, experience has indirectly established its validity (at least approximately) by confirming the results obtained from it.

The laws of skin friction were investigated by Froude in a classical series of experiments made by him on planks, and he announced his results before the British Association in 1872 and 1874. He found that the resistance R of the plank could be represented by $f\rho A V^n$, where ρ is the fluid density, A the area of wetted surface, V the velocity of the plank, and f,n, empirical constants. The coefficient f depended on the roughness of the plank, and became smaller as its length was increased. This was attributed by Froude to the frictional wake or current of water dragged forward by the plank; the velocity of wake being greater towards the rear, the relative speed and friction were less. The index n also depended on the roughness of the planks; its value varied from 2 for rough surfaces to 1.825 or 1.83 for surfaces corresponding to a smoothly painted hull.

The longest planks tried by Froude were 50 feet in length. He estimated the coefficient f for the last foot of the plank as well as that for the whole length; by assuming that the former remained unaltered for all lengths greater than 50 feet, he enabled his results to be applied to full-sized ships. The validity of this method of extrapolation has, however, been questioned, and recent researches in viscosity present an alternative method.

From considerations of dimensions Rayleigh obtained the formula below for viscous motion uninfluenced by gravity or other external force:

$$R/\rho A V^2 = F(VL/\nu)$$

when ν is the coefficient of kinematic viscosity and L a linear dimension which can here be identified with the length of the plank.

By means of experiments on pipes Osborne Reynolds determined the form of the function $F(VL/\nu)$, finding that it can be represented by $k(VL/\nu)^{-m}$, where m has two different values according as VL/ν is great or small. When VL/ν is below a certain critical value the flow is purely viscous or laminar, and $m=1$; $R/\rho A \propto \nu V/L$, confirming Poiseuille's result. When VL/ν is greater than another higher critical value, the motion is turbulent and m is a small fraction; $R/\rho A \propto \nu V^{2-m}(L/\nu)^m$, or V^2 approximately. When VL/ν lies between the two critical values the motion is unstable, changing readily from viscous to turbulent and vice versa.

These results for pipes have been confirmed by various experiments on planks and other flat surfaces. In 1915 G. S. Baker* compiled a large number of available data taken from resistance experiments in air and water, including those by W. Froude, Zahm, Geber and himself, and plotted them in a form of $R/\rho A V^2$ on a base of VL/ν. Except for a few discrepancies the spots followed closely two curves: one nearly hyperbolic, applicable to small abscissa values, and another nearly straight, though slowly descending, for large abscissa values; the former represented viscous flow and gave $m=1$, the latter corresponded to turbulent flow, and for it $m=0.14$ approximately. Between the ends of the two curves the spots were irregularly placed, indicating instability of flow. Here, as in all other experiments on resistance and propulsion, the results for air accorded well with those for water, due allowance being made for the difference of density and (where necessary) viscosity.

The second (turbulent motion) curve includes the range covered by model experiments of ships, so that the law of frictional resistance should be $R \propto L^{-.14}\rho A V^{1.86}$. On comparing it with Froude's formula it appears that $f \propto L^{-.14}$ instead of being a linear function of the length; moreover the index n is slightly greater than the 1.825, which was later deduced by R. E. Froude from his

*Trans. North East Coast Inst. of Eng. & Ship., Vol. XXXII.

father's experiments. With full sized ships VL/v would be much greater than any values as yet realized in experiments, so that the exact law of frictional resistance for ships is uncertain. Either method probably gives a fairly close approximation to the correct value, but precise information is at present lacking. This is of importance, for frictional resistance (except at very high speeds) accounts for more than 70% of the total. A committee has recently been formed with the object of instituting frictional experiments on full-sized ships, so that more exact data should be available in the near future.

It has not hitherto been found possible to investigate the turbulent motion caused by friction even in the simplest cases by purely mathematical methods, but the nature of this motion has been investigated experimentally by Calvert (1893) and others. In the immediate neighbourhood of the moving surface—planks or ships—there is an eddying belt of water, which is succeeded further out by a wake or forward current whose velocity rapidly diminishes as the distance from the surface is increased and gradually increases towards the rear. The wake has an important influence on the working of the propellers, and affects the motion of any device, *e.g.*, a speed recorder, fitted in the water near the hull. It must be modified by the stream-line motion round the ships, so that the wake and the frictional resistance of a ship differ from those of a plank. This difference of resistance has been assumed negligible and is probably small; but the effect of form on frictional resistance appears to merit further investigation which would have to be mainly of an experimental character.

W. Froude, in his investigations on wave-making resistance, made no attempt to estimate it directly; he determined it for ships from experiments on models, using the law of comparison to connect the two results. This law is to the effect that if L is the linear ratio of ship to model, and the speed ratio is $L^{\frac{1}{2}}$, then the resistance ratio will be L^3. It can be proved very simply by a "dimensional" method, assuming that the waves are purely gravitational and that the influence of viscosity is negligible.

In order to verify this law the *Greyhound* was towed at sea and her resistance measured by a dynamometer; the results were compared with those deduced from experiments on a model. After deducting the frictional resistances (calculated separately) the agreement between the residuary resistances was found to be nearly exact. This experiment confirmed Froude's hypothesis, and established the utility of model experiments for the purpose of estimating horse-power. This method of powering is now employed in all cases where accuracy is required; and numerous experiment-tanks have been constructed which are in daily use both for testing new designs and for purposes of research.

The effect of varying form and proportion can also be rapidly estimated from model experiments; and a large mass of data has been accumulated from the systematic experiments of W. and R. E. Froude, D. W. Taylor, G. S. Baker, and others. These results enable the resistance of many ships to be calculated approximately without running a special model in a tank; and a system of coefficients has been devised for the purpose of facilitating such calculations and recording the results of model experiments.

The method of estimating power by model experiments permits a large variety of forms to be rapidly tested and the most economical one selected for a new design. Appreciable saving in fuel has been frequently obtained thereby at a trifling cost; and it is certain that the great advance in the speed and size of ships effected during the last 30 or 40 years could not have been realized without the aid of tests based on W. Froude's theory.

Wave resistance, unlike frictional resistance, has been calculated by purely mathematical methods in a few simple cases; and it will be convenient to review briefly the theoretical work that has been undertaken on the subject. As a preliminary to finding the wave resistance, it is necessary to determine the wave pattern accompanying a ship or other body advancing through the water. This was first done by Kelvin about 1887 for the case of a single pressure-point applied to the surface, and moved at uniform velocity along the water. He obtained the well-known result that the wave system is comprised between two lines inclined backwards at about 20° to the line of motion; it consists of (*a*) a series of transverse waves having almost straight crests extending to the bounding lines and spaced at regular distances $2\pi V^2/g$; (*b*) a series of diverging waves extending from the ends of the transverse crests towards the point of pressure, (*c*) a local disturbance near the point of pressure. The system extends infinitely far in the rear, but the heights of the waves there degrade so that only the first few waves are of importance.

From this result it is possible to obtain the waves due to a diffused travelling surface pressure. A similar but more interesting case was worked out in 1917 by T. H. Havelock, who found the wave pattern and the corresponding resistance of a moving sphere wholly submerged. He obtained the remarkable result that as the speed becomes greater the resistance increases up to a maximum, and then gradually diminishes to zero. A further case investigated by him was that of an elongated solid of revolution moving horizontally along its axis. The wave pattern resembles that made by two pressure points, one near each end of the solid; interference takes place between the component systems so that the resultant transverse waves are large or small according as the components are in or out of phase. This leads to corresponding fluctuations in the resistance, which are exhibited by humps and hollows in the curve of resistance plotted on a base of speed.

The case of a ship moving on the surface differs from that of the submerged solids above considered; it is possible, however, to replace the ship by a surface pressure distribution which is approximately equivalent except quite close to the ship, and thence to determine the wave pattern and the resistance due to it. The equivalent pressure distribution can be determined approximately by first replacing the actual vessel by a solid of revolution having the same immersed area at each section; the velocity of the stream-line flow, and thence the pressure are obtained by finding the source and sink distribution along the axis approximate to the form of this solid. A graphical method of effecting this was devised by D. W. Taylor, and the pressure curves for some actual ships were thereby obtained by W. McEntee in 1909, and G. S. Baker and J. L. Kent in 1913. They found that the curves generally consist of two crests at bow and stern, with a slight hollow between; and are generally similar to the curve for the elongated solid referred to above. The subsequent determination of the resistance is possible, though very laborious; for even the comparatively simple case worked out by Havelock demanded a difficult and tedious mathematical analysis. The main features of the resistance curve would, however, remain unaltered; *i.e.*, there would be fluctuations in the curve corresponding to interference between the two principal component wave systems.

The foregoing results of mathematical investigations are entirely confirmed by experiment. The wave pattern consists of two components due to bow and stern, each consisting of transverse and diverging waves. These interfere, and the consequent fluctuations of resistance were observed by W. Froude in a series of experiments on a model whose length was successively altered by varying the parallel middle body without altering the ends. At a speed of 14.4 knots the curve of wave resistance plotted on a base of length showed maxima at 220, 330 and 440 feet length, and minima at 165, 275 and 385 feet. Corresponding fluctuations occurred when resistance was plotted in a base of speed. A more detailed investigation was made in 1913 by Baker and Kent, and similar results were obtained.

As far as is known the only attempt to calculate the wave resistance of an actual ship by purely mathematical methods was made by J. H. Michell in 1898. The subject presents great difficulties, and it seems highly improbable that purely theoretical investigations will ever replace the experimental method now in use. Nevertheless the mathematical theories are interesting and instructive from a general scientific standpoint, and it is hoped that such researches will be continued. Attempts have been made, *e.g.*, by Havelock and Hovgaard, to construct semi-empirical formulae on a rational basis which would represent the wave resistance of a ship; but hitherto not much success has been obtained, which is not surprising in view of the complexity of the problem.

A considerable amount of attention has been given lately to the modification in total resistance produced when a ship is proceeding in a narrow or shallow channel. Scott Russell accidentally discovered that the effort of towing a barge in a canal was reduced when its speed was increased, and that the barge then travelled along the crest of a solitary wave with little resistance. The ideal two-dimensional case of a pontoon moving in a canal of the same breadth was dealt with mathematically at a considerable length by Kelvin, who obtained the forms of a number of pontoons whose wave-resistance was nil.

The more difficult case where the depth alone is limited has been investigated by Havelock, who evaluated the resistance caused by a symmetrical distribution of surface pressure advancing over shallow water. At moderate speeds the wave resistance is greater in shallow water, but at high speeds the opposite holds. This result is confirmed by trials carried out on ships, the only difference being due to the skin resistance which is always increased by limitations in breadth or depth of water. It is found that the total resistance and horse-power of ships at moderate speed relative to their size is increased by shallowness; that of ships at high speeds (*e.g.*, destroyers) at first increases up to a maximum, and then diminishes, so that the resistance in shallow water may be less than that in deep; at very high speeds there may be no increase, *i.e.*, the resistance is greatest in deep water.

The alteration in resistance is accompanied by one in the wave pattern which broadens out laterally; and the existence of shallow water is frequently revealed by the perceptible change in the form of the waves near the stern. This also is in accord with theoretical investigations.

The general effect of shallow water has now become of great importance on account of the large increase in size and speed of ships. It is, for instance, impossible for a modern large warship to obtain her full speed over large areas of the North Sea and English Channel. If, in the future, further advances are made in size and speed of war ships, the areas over which such ships can operate at top speed will be considerably restricted.

When a ship is driven by her own propeller, instead of being towed as assumed in the theory of resistance, the motion of the water and the forces on the ship are modified. In particular, at a given speed, the thrust exerted by the screw is usually greater than the resistance of the ship when towed; and the power developed by the machinery or transmitted by the shaft is approximately twice that required for towing. The difference between these powers is due partly to the interaction of propeller and ship, partly to losses in the propeller, and partly due to friction in machinery and bearings; their ratio is termed the propulsion coefficient. The prediction of the propulsive coefficient in a new ship is by far the most uncertain factor in the estimate of power required; the matter is largely of a technical nature and will not be pursued in this paper. The same remark applies to the process of propeller design which is based entirely on standard experiments carried out by R. E. Froude, D. W. Taylor, and others, which are not of great mathematical importance.

The principles underlying the action of a propeller are, however, of general interest; and although no theory which is even approximately complete has yet been put forward, a brief description follows of the theoretical investigations undertaken.

In the first theory due to W. Froude in 1878 the area of each propeller blade was divided into elementary strips and an expression obtained for the pressure on each strip from independent experiments on planes dragged obliquely through water.

By integration the thrust and torque on the whole propeller were obtained and the propeller efficiency deduced. The blades were assumed to act on undisturbed water, *i.e.*, not to interfere with one another—an assumption which cannot be correct for the broad-bladed propellers now made, though the error would probably be small in a typical air-propeller. The propeller efficiencies obtained from this theory nevertheless agreed fairly well with those deduced from experiment.

In 1889 R. E. Froude propounded a simple theory in which the propeller was regarded as an actuator, *i.e.*, the pressure was assumed to be suddenly increased within the propeller, the velocity increase being gradual. When the problem is reduced to one of steady motion the velocity of the water increases ultimately from V (the speed of the ship) to $V+v$, v being the speed of the sternward race, Froude showed that the increase of pressure takes place when the water has received half its acceleration, or at the velocity $V+\frac{1}{2}v$. From this it can be deduced that the propeller efficiency is $V/(V+\frac{1}{2}v)$, an expression given by McEntee.

In 1916* W. Burnside made a further investigation on a propeller regarded as an actuator and confirmed some of R. E. Froude's results; he also took into account in a general way the effect of the proximity of the stern of the ship. Such investigations of necessity omit certain important features of the problem, but it is probable that they give a fairly accurate idea of the motion of the water everywhere except quite near the screw; on general grounds it is desirable they should be continued.

Strength

The complete determination of the stresses brought on the structure of a ship in a seaway is an insoluble problem. Many of the external forces, *e.g.*, those due to the impact of a heavy sea are unknown, and moreover, the structural arrangements of an ordinary vessel preclude any exact calculation of the stresses induced. To distribute the material of the hull to the best advantage, and to eliminate unnecessary weight while ensuring adequate strength, some form of stress calculation is required; and it is found advisable in practice to deal with the question piecemeal, calculating the separate stresses due to the principal straining actions and making simplifying assumptions where necessary. With practical experience of ships at sea as a guide, this method gives satisfactory results on the whole; but the need for economy in weight, particularly in warships, is so great that further developments or applications of the theory of elasticity will give welcome assistance towards attaining nearer perfection in the distribution of structural material.

During the last 100 years the lengths of ships have progressively increased, and the proportions of modern ships are such as to render it certain that the principal strain and stresses are those due to longitudinal bending among waves. Reed was one of the earliest investigators of longitudinal strength, and in 1871 he showed that a ship might be considered structurally as a beam and obtained its bending moment by assuming the vessel to be placed on a trochoidal wave of her own length with either the crest or the trough of the wave amidships, these being the conditions which induce the greatest stress. As a basis of comparison, it is now customary to assume the height of the wave to be one-twentieth of its length. In 1871 and 1874 W. John, on the foregoing basis, estimated the stresses in a number of ships (using the ordinary theory of stress in beams) and in several cases found them very high. He cited instances of such ships receiving serious structural damage, and it was evident that the theory would be useful in preventing unduly weak ships being built. This has proved correct; and for many years past the structure of nearly all new ships (except comparatively short vessels) has been designed directly or indirectly from such calculations.

Various refinements have since been proposed which are of considerable theoretical interest. In 1883 W. E. Smith showed that when account was taken

*Proc. London Math. Soc., 1918.

of the difference between the actual pressures in a wave, and those due to the hydrostatic head, the calculated bending moment on a ship was modified; and in certain cases worked out by him it was considerably reduced. A ship in a seaway is subject to vertical oscillations, and the effect of the accelerations thereby produced, and the consequent modifications in the bending moments and stresses were calculated in 1890 by T. C. Read. In addition in 1896 A. Kriloff determined the effect of all possible oscillations on the stresses; his investigation provides a complete solution of the problem, though the calculations required to apply it are long and difficult.

None of these refinements of the original theory is employed in practice, and moreover they are not strictly necessary so long as the method is only used to compare ships of similar types. It is nevertheless important to ascertain if the calculated and actual stresses are in agreement, and this was undertaken in 1905 by J. H. Biles on the destroyer *Wolf*. The ship was supported on cradles in dry dock, so that the bending moments were readily obtained. The strains in various parts of the structure were measured by extensometers, and the stresses deduced from an assumed modulus of elasticity of the material, which was chosen so as to make the moment of resistance of the section equal to the known bending moment on it. The vessel was afterwards sent to sea in rough weather and the strain again measured. It was found that the ordinary theory of bending gave substantially correct results; but the modulus of elasticity of the structure was only about 10,000 tons per square inch, less than its value (13,000) for solid steel. Moreover the greatest stresses induced at sea were substantially less than those calculated for the ship on a standard wave. The theory was justified; but the stresses obtained by it are evidently over-estimated. This exaggeration becomes more pronounced still with long ships, since the height-length ratio of the assumed standard waves is greater than in actual waves. Allowance for this is made either by reducing the wave heights more nearly to observed dimensions, or by permitting high nominal stresses to be worked to.

The full application of the theory involves the determination of the stresses permissible. This presents no difficulty for parts in tension; but for stiffened plating under compression or subjected to shear, no entirely satisfactory theory has yet been evolved, and the methods of calculating the permissible stress are then to some extent empirical.

Next in importance to longitudinal strength is the resistance of the structure to transverse deformation. Stresses are induced in the transverse framing of a ship resting on the blocks in dock, or rolling in a seaway; a method of estimating them was devised by J. Bruhn in 1901 and applied by him to various types of ships. His process is based on the principle of least work, though it can also be derived from the theory of the flexure of thin rods. A ring consisting of frame, beam and plating is considered and its flexural rigidity calculated at various points along its perimeter. The bending moment at any point can be expressed in terms of the known external forces and a number of unknown internal reactions, and the deformation obtained by graphic integration. On expressing the fact that the total deformation round the ring is zero, equations are obtained, the solution of which gives the unknown reactions and thence the stress at any point. Bruhn's method has been used principally in connection with merchant ships and submarines, but it is also suitable for application to all ships in which the transverse bulkheads are not closely spaced.

Local stresses arise through many causes, among the most important being those in shell plating and framing due to the pressure of the sea. The maximum stress and deflection of plating subjected to lateral pressure has been investigated experimentally by J. Montgomery, W. Hovgaard, and others; and mathematically by T. G. Boobnoff (1902), who showed that thin plating resisted partly as a diaphragm, by pure tension, and partly as a beam by flexure. This question is of great interest, for experience has shown that bottom plating subject theoretically to high stress has nevertheless shown no signs of weakness on service. The explanation may lie in the permanent set produced by overstrain, which induces self-stress in the material, and partly relieves it against further strains.

The recoil forces due to gunfire are generally large, and the structure under a gun is specially designed to transmit and distribute such loads. The external forces caused by the firing pressures are known, but the stresses induced and the liability of the support to collapse can only be found very approximately. In this connection a theory of the strength of a short cylinder or cone loaded unevenly would be a useful guide for future designs of such structures. Further investigations on the theory of thin plating in general, whether plane or curved, will also be of valuable assistance to a naval architect in connection with the structural design of ships.

Conclusion

From the foregoing account of the history of scientific naval architecture, it is evident that its development and the resulting progress in ship design have been greatly dependent on the application of mathematical principles. The foundations of the subject have now been laid and the need for further theoretical investigations is less acute than it was 100 years ago. This improvement is due in large measure to the establishment of various Schools of Naval Architecture and the foundation (1860) of the Institution of Naval Architects and kindred Societies which have very materially aided progress in the science.

Portions of the theory are, however, still incomplete; and in this paper the directions in which further mathematical research is likely to prove useful have been indicated.

On account of the continual progress in naval architecture new types of ships are being evolved and new methods of propulsion devised. These changes will introduce fresh problems, many of which will require theoretical investigation; it is thus certain that in the future, as in the past, the close connection between naval architecture and mathematics will be maintained.

I wish to acknowledge my indebtedness to Mr. L. Woollard for the assistance he has rendered in the preparation of this paper.

THE EDUCATION IN MATHEMATICS OF STUDENTS OF NAVAL CONSTRUCTION

By Mr. L. Woollard,
Chief Naval Constructor, The Admiralty, London, England.

The preparation of a scheme of education suitable for a naval constructor presents some difficulties which do not arise with those of other professions. A student of law, medicine, theology, etc., receives first his general education, which is followed by specialized training for his vocation; the two are largely independent. A student in naval construction, on the other hand should be trained in the practical details of his profession at an early age when his mind is most receptive. The more advanced stages of his professional training, however, demand a large amount of theoretical knowledge; whilst a good general education also is a great advantage to him.

A compromise has to be effected in order to reconcile these conflicting considerations.

By selecting mathematics as the principal subject for his general education, and by teaching it in conjunction with practical and professional subjects, it has been possible to devise a combined system of general and professional education which has proved very satisfactory in its results.

Since the type of general education which should be given to engineers, and in particular, the scope of their mathematical studies, has frequently been the subject of controversy, a brief account of the mathematical course followed by students of naval construction at the Royal Naval College, Greenwich, may prove of interest to members of this Congress.

All students, prior to entering Greenwich, spend four or five years at a dockyard or other shipbuilding establishment; the training during this period is mainly practical, but it includes also a thorough preliminary education in mathematical and scientific subjects.

The course at Greenwich lasts three years, nine months of each year being spent at the College, and three months at a Dockyard on practical work. The average age of the students in naval construction commencing is about 21 years, and their number is usually 4 or 5 per year. The principal subjects taken are Naval Design and Construction, Mathematics (Pure and Applied), and Applied Mechanics; instruction is also given in Physics, Chemistry, and a foreign language. At the end of each session an examination is held, the subjects being those taught during the session. The final examination covers nearly the whole three years' course.

In the first year the following subjects are taken in the course on mathematics: Analytical Geometry, Differential and Integral Calculus, Statics and Dynamics, and (formerly) Electrostatics and Electrodynamics. Although the applied branches are not in the first year carried very far, sufficient instruction is given to establish all these subjects on a sound basis.

The second year's course comprises the Theory of Curves, Differential Equations, Analytical Solid Geometry, Theory of Equations, and Dynamics of a Particle and of a Rigid Body of two or three dimensions. Gyroscopic problems are included.

Hydrostatics, Analytical Statics, Flexural Rigidity of Rods and Hydrodynamics complete the course in the third year. The last subject is treated at considerable length; the course deals with all ordinary problems on irrotational motion of fluids, the motion of a solid in a fluid, and waves both in liquids and in gases (*i.e.*, Sound). Pure mathematics, *e.g.*, Green's Theorem and Partial Differential Equations, is included, but only as subsidiary to the Applied.

Degrees are not conferred on successful students, but professional certificates are granted on the results of the final examination, a percentage of 60 on the total marks being required for an ordinary or "second class" certificate.

From the foregoing it will be evident that the mathematical training received by students of naval construction at Greenwich is more extensive than that usually given to technical students; and is in fact equal, or nearly equal, to the standard realized in universities where mathematics alone is taken as a specialized subject.

The course undoubtedly includes rather more than the minimum necessary to enable an average naval constructor to perform his professional duties; though several who have to specialize in certain branches of ship design have found it necessary or desirable to extend their knowledge of the subject. The justification for the greater time devoted to mathematics lies in its general educational value, in the mental training it provides, and in the development of the faculty for exact reasoning and logical deduction. For these reasons a high standard of mathematics has been retained, although the continuous progress in the science of naval architecture is rendering it increasingly difficult to resist encroaching on the time devoted to mathematical subjects.

The best method of teaching mathematics and in particular of making it most useful to technical students is beyond the scope of this paper, being a matter for a professor of the subject rather than for a naval constructor. An important and admirable feature of the Greenwich course is the working out by the students of exercises based on the substance of the lectures. This is generally done under the supervision of the professor, who also personally delivers the majority of the lectures. An hour's lecture by the professor is followed by one or two hours on exercises designed to illustrate points which were not specifically dealt with in the lecture.

The Chair of Mathematics at Greenwich from 1885 to 1919 was held by Professor William Burnside, F.R.S.; and the course outlined above was evolved by him as the most suitable one for naval constructors. With Professor Burnside's original work on mathematics, and his development of certain branches, notably the Theory of Groups, members of this congress will be familiar. The instruction given at Greenwich has greatly benefited by being in the hands of so eminent an authority, who, in addition to possessing a complete knowledge of his subject, was an admirable and patient teacher.

It is desired to record an appreciation of the ability and painstaking labour which he devoted to rendering a naturally difficult subject attractive to technical students. This was done without sacrificing any of the rigour appropriate to the subject; and it is with the object of showing that a sound mathematical education can be successfully combined with an efficient technical training that this paper has been written.

Many of the Greenwich students have served in the most responsible positions of the public service and of private shipbuilding enterprises, and they comprise the great majority of naval architects of any standing in this country. All of them have freely expressed their gratitude for the sound mathematical instruction they received at Greenwich and for its usefulness in their subsequent professional careers.

MATHEMATICS IN INDUSTRIAL RESEARCH

By Mr. A. P. M. Fleming and Mr. R. W. Bailey,
Research Department, Metropolitan-Vickers Electrical Company, Limited, Manchester, England

Introduction

Industrial research or research directed specifically to industrial needs and progress differs in many respects from pure science research, notably in the origin and character of its problems. In the former case a difficulty or need often arises suddenly and a practical solution, although possibly incomplete, is demanded with minimum delay. Frequently progress upon important design or manufacturing operations is suspended pending a solution of the problem encountered and usually the manufacturer is justified from economic considerations in accepting a solution which enables work to proceed until a more complete settlement of his difficulties is yielded by further experience and research. Thus progress in the industrial arts aided though it be by the advantages of well-directed research is usually a step by step movement in which of necessity strides are short but as continuous as possible. In pure science research, however, which is concerned with the establishment of fundamental principles and knowledge, speed and the necessity of a quick solution, if only an interim or partial one, are not conditions, and generally time is available for more complete and rigid methods. Industrial research, when most widely developed, also embraces these more fundamental activities and in some degree provides for them, but in this case its mathematical needs are the same as in pure science research, and they are satisfied by mathematical physicists and mathematicians of like bent and training. No special problem is presented here, but it is in connection with the major activities of industrial research, closely associated with immediate progress, that its special mathematical needs arise.

The authors do not contemplate a special form or arrangement of mathematics for industrial research. The introduction of what is now known as practical mathematics and the adaptation of mathematics to suit the needs of an art, as for example mathematics for engineers, has not proved an unmixed blessing. While, if we may digress, a capable teacher can and does impart real mathematical knowledge, in teaching such subjects the practice frequently leads to unsatisfactory results and a lamentable absence of mathematical ability. The authors, therefore, are averse to adding to the list of mathematics for this or that, but they hope that by drawing attention to the conditions and special needs of industrial research these will be of interest to mathematicians, and some of the many so ably qualified may bear these in mind and extend and develop methods of treatment which are within the possibilities of research workers and provide the implements they require.

Character of Industrial Research

Industrial research, even in one industry, may range from the investigation of a works manufacturing process to a highly theoretical enquiry of a technical character. When, therefore, industry in general is considered the mathematical needs, as may be imagined, are exceedingly diverse and range from little more than simple arithmetic and algebra to the formation and solution of many innocent looking but particularly intractable differential equations. Generally, of course, the more nearly an industry approaches a science the greater are its mathematical needs. Probably the engineering industry is the most advanced and extensive in this respect and this contribution is written with the conditions of this industry chiefly in mind.

The complexity of most practical problems is such that a rigid mathematical solution is rarely possible. This is so even in what appear to be simple cases and perhaps the greater part of engineering theory is based upon hypotheses and conditions which although not rigidly accurate are sufficiently well representative of facts or which are susceptible of suitable correction to conform with the results of experiment. A typical example in this connection is afforded by the branch of mechanics embraced by hydraulics. Here, as is well known, certain assumptions are made regarding the phenomena of flow of fluids and formulae are based thereon, suitable coefficients being employed as shown by experiment to be necessary. It will be appreciated that an exact treatment of many of the problems, which are regarded in hydraulics as straightforward, would be found impossible of solution by a rigid application of the principles of hydrodynamics. Nevertheless hydraulic engineers are able by employing the former to design with confidence, and the many examples of hydraulic engineering associated with the utilization of water power testify to the practical utility of these principles. In electrical engineering knowledge is more precise and theory is correspondingly exact, but even here many cases are encountered where recourse must be had to experience on account of the impossibility or extreme complexity of a mathematical solution. Broadly speaking, the conditions are similar in industrial research problems where mathematics has to be applied, and rarely can it be said that a rigid and complete solution is attempted or necessary.

As may be surmised, investigation will comprise (*a*) mathematical analysis, (*b*) mathematical analysis in conjunction with experiment, (*c*) experiment entirely, and frequently all three will be utilized in whatever order consideration and experience indicate will be most satisfactory. In some cases knowledge of a phenomenon or other data may be insufficient for a mathematical analysis to be advisable and recourse must be had to experiment to provide the information; on the other hand, where knowledge and data are sufficient to permit a mathematical analysis having a sufficient approximation to actualities initial mathematical analysis is often the most satisfactory first line of attack.

Sufficient has been said to indicate that in industrial research problems involve such complexities if rigid treatment is contemplated that approximations have to be made and moreover a partial solution leading to a complete one later is of more value if obtained quickly than the latter if much delayed. This being so, judicious approximation, except when rigid methods are readily applied, should be a guiding principle in the use and development of mathematics in connection with industrial research. This view may not at first commend itself to mathematicians, but it is necessary to realize that to the research worker mathematics is a tool and its utility in common with that of other implements employed is judged by its results.

The use of mathematics in industrial research is restrained by the complexity of the problems and phenomena dealt with, and also, of course, by the limitations of the mathematical capacity of the investigator. On account of the former, alternative methods of attack, such as by experiment, the use of models, and experimental work utilizing other physical phenomena known or found to follow similar laws to those involved in the problem often provide a solution where a mathematical one would be impossible. Thus the development of science of aeronautical design in reference to airships and aeroplanes has been largely based upon model experiments employing the principle of dynamic similarity and in the case of two dimensional stress in elastic solids the optical apparatus and methods devised by Coker yield results for parts of almost any contour whereas so far only simple cases have been solved mathematically. Much work is now done with the aid of models wherein geometrical, dynamical or equational similarity with the full scale part is ensured or differences allowed for, and as a result the demand for mathematical analysis is minimized. Concerning the limitations imposed by the mathematical knowledge of the investigator, while in some cases skill in making judicious approximations considerably extends the range of utility of one's mathematical equipment, there are cases where something further is required. To overcome this handicap the staffs of some research organizations include assistants of special mathematical training and ability. It is difficult, however, to obtain the service of mathematicians who also possess sufficient practical insight and bent to co-operate most effectively with other expert investigators, and there is a tendency for the mathematician to become an operator solving mathematical difficulties encountered by the investigator without an appreciation of the physical quantities of the work. Where the investigator, however, combines with his special knowledge a considerable aptitude in mathematical analysis, the maximum advantage is obtained from the latter and the best results ensue. If a research staff does not include a mathematician and in some cases even when it does, problems may arise which call for expert mathematical treatment, and in these cases authorities of the highest standing are consulted, not always, it may be said, with success, but with the satisfaction that, if unsuccessful, other methods must be tried. The ingenuity of the investigator is then taxed, and sometimes by a combination of graphical and analytical methods and approximations a solution is reached which is sufficiently accurate within the range of quantities of practical account to be useful. Such solutions, however, usually suffer in not being general, and moreover the work entailed is long and tedious and being repetitive in character due to the successive approximations necessary to secure a sufficiently reliable result it sometimes occupies weeks. Instances of this kind demonstrate the practical importance of approximate, and to the purist possibly irregular, methods of treatment, but also they emphasize the need for developing rapid methods of solution by successive approximations.

The following examples are given to illustrate how readily in practice problems are encountered presenting considerable mathematical difficulty and where in the absence of a solution by analytical methods recourse is had to experiments or rough approximations to obtain practical data.

1. Comparatively thin high tensile steel rings of cylindrical form having a length usually less than the diameter are employed to support the end connections of the electrical conductors in the rotors of high speed electric generators. The true stress distribution in the ring caused by the centrifugal force of the conductors, due to tension and bending, can be obtained when the stress distribution due to two equal and opposite forces acting on a diameter at any distance from the middle of the ring is determined. The case of uniform radial loading in any plane either circumferentially or axially is readily solved, but the general one of single loads presents difficulties, and so far as the authors are aware has not been solved. In practice it has been necessary to make assumptions and to be guided by the results of experiments.

2. In studying certain problems in vibration and whirling of shafts it was desirable to ascertain the amplitude of the displacement and its phase relation with the disturbing force when the frequency of the latter increases from zero and passes through the natural frequency of the system acted upon. The usual treatment of vibration problems employs the general equation $A\ddot{x}+B\dot{x}+Cx=f(t)$ where $f(t)$ is of the form $F \sin qt$, q being a constant. The solution of this equation is well known and is straightforward. In the case under consideration, however, q is a function of t and also F when the whirling of shafts is considered. Attempts to solve the equation taking q as a linear function of t and as a sine function were unsuccessful.

3. In electrical apparatus for alternating current, particularly transformers, the magnetic circuit is built up of many thin sheets of steel. Owing to the alternation of the magnetic flux energy is lost as heat in the iron, being conducted to the exposed surface of the mass of iron. The heat flow is complicated by the fact that the thermal resistivity in the direction across the laminae is about twenty times that in the plane of the laminae. Taking the case of an assemblage of a number of rectangular laminae forming a rectangular solid it is required to calculate the temperature at any point within the solid given the external temperature, the rate of energy loss per unit volume, assuming that all exposed surfaces dissipate heat at the same rate per degree above the surrounding temperature.

A solution has not been found to this problem and data have to be obtained by measurement of temperature using thermocouples placed at different positions in the body of laminations.

4. It was required to find the rate at which oxygen diffuses through still water in a vertical pipe and a number of tests had been made by apparatus consisting of an upper vessel connected to a lower one by a large bore glass tube. The oxygen concentration of the water in the upper vessel was known and maintained approximately constant and the contents of the lower vessel were analysed for oxygen concentration after a given time, the initial concentration being zero.

The derivation of data was complicated by the fact that the capacity of the conduit was comparable with that of the lower vessel. Taking the rate of diffusion as proportional to the oxygen concentration gradient the problem and its equations are similar to those for the case of heat flow in a conducting medium temperature and heat corresponding with oxygen concentration and content respectively. An attempt to make a mathematical analysis proved so formidable that it was decided to repeat the series of experiments with a much larger lower vessel such that the oxygen taken up by the water in the conduit could be neglected.

5. In the case of oil immersed switches used for electrical purposes the conductor may be regarded as of U form, the two limbs of the U projecting well above the steel tank which for the present purpose may be regarded as an open rectangular box. The magnetic properties of the steel of the tank are defined by its $B-H$ curve. It is required to calculate the loss of energy in the tank due to eddy current set up in the sides and bottom thereof by a given alternating current in the conductor.

It will be understood that important decisions generally depend upon the solution of such problems as the foregoing, and this should add zest to the work. If assumptions are made their influence should be closely estimated so that an idea of the maximum error caused thereby is known. This is an important aspect of mathematics in industrial research and one which at first is not always fully appreciated by mathematicians and undue nervousness is sometimes displayed when it is realized that action is to be taken upon the result of a mathematical analysis. This state of mind does not arise from the accuracy of the work, but generally some doubt regarding the validity or influence of assumptions.

There appear to be two directions in which the authors consider mathematicians can render further assistance to industrial research, namely, (1) by giving special attention to the development of approximate methods of solution of tedious and difficult or at present impossible mathematical operations which, by a process of successive approximations can possibly be made to yield results sufficiently accurate for practical purposes and (2) by incorporating work of the character embraced by (1) in the curriculum so that mathematicians so trained will be of greater service to industry.

THE TEACHING OF MATHEMATICS FOR ENGINEERING STUDENTS

By Professor Sir James Henderson,
Royal Naval College, Greenwich, England.

Mathematics is a science comprising a number of subjects, many of which may be made the study of a lifetime, and the science attracts men of a certain type of mind, who are interested in the subject for its own sake, and who take up teaching as an attractive means of livelihood, which permits them to follow up their forte. A youth who shows special ability in languages, mathematics, engineering, music or drawing usually tries to make that particular subject his life's vocation.

Mathematics differs from other subjects in that throughout the ages it has gradually entered into our educational system and elementary mathematics is considered to be an excellent mental training for everyone. Whether a similar mental training could be obtained by the study of some other subject has been argued elsewhere, but need not be considered here, as no engineer desires to displace mathematics from its high position in the educational system.

The great difficulty and the cause of most of the misunderstandings arises from the fact that mathematics, besides being a serious study, or a hobby, or relaxation, or a means of livelihood, to one class of men having a certain type of mind is also a tool employed by another class of men having, generally speaking, a type of mind of a different character. Although some engineers would make good mathematicians and *vice versa*, still, as a general rule, the two types of mind are quite distinctive. The engineer is interested in practical problems relating to the application of physical principles in the service of man, and is ready to employ any and every tool to solve the particular problem in which he is engaged, whereas the mathematician is frequently interested in the pure science alone, and not at all in its application.

Now in other phases of life, the man who makes the tool is not considered to be the best exponent of the art of using it, and since the use of a tool may take years to learn, it follows that the longer the learning takes, the less the maker of the tool will know about its practical application. In music, for example, the maker of the musical instruments is not considered to be an expert on their use either alone or in orchestration, although he may occasionally be consulted by orchestra conductors as to the best method of producing a certain effect. He would not as a rule be chosen to lecture to a class of budding musicians studying orchestration, yet the selection of a mathematician to teach engineers the methods of applying that science in the practical problems of life, may appear to some engineers as a somewhat parallel case.

The reasons for the widespread opinion among engineers that there is something wrong with the teaching of mathematics are due to one or other of the following: (1) a want of confidence in their own ability to apply the few mathematical tools they possess, (2) the feeling that these mathematical tools are frequently insufficient to solve the practical problems which they have to deal with from time to time, (3) the unpractical nature of the solutions of many so-called practical problems presented by mathematicians, or (4) an exaggerated idea of the power of higher mathematics to deal with the complicated phenomena of nature and the feeling that if only they possessed the mathematical knowledge themselves they would be better equipped as engineers.

The solution of all the difficulties lies in education of both mathematicians and engineers and a better understanding between them. The want of confidence in applying the elementary mathematical knowledge possessed by the average engineering student could be removed by more frequent exercise of that application. As a general rule the mathematical teacher expounds a new subject to his students and is content with working one or two examples on its application which may have no connection with engineering practice. That may be quite sufficient for students who are going to be teachers and hope to pass on the knowledge they acquire to others in the form in which they receive it. It is altogether insufficient for engineers, however. Consider, for example, the elementary subject of logarithms. Thousands learn to use logarithms as a tool, either by tables or the slide rule, who have no knowledge of the meaning of the word or the manner in which the tables are compiled. So it is with other mathematical functions, such as those compiled in navigational or ballistic tables. The aptitude in the application of these functions only comes by much exercise, and the engineer wants to use all his mathematical tools with the same facility with which he uses his logarithmic tables. The teacher of mathematics will argue that he has no time to give to such exercise if he is to cover his curriculum. The obvious answer to that is that it is better to have few tools and to know how to use them well, than a chest of tools which have never been used. The engineer in looking through his mathematical note books, which represent his mathematical tool box, sees whole books of notes of which he has never made any use. The mathematician will maintain that it is impossible to estimate the mental training he has obtained from such study, but why the mental training in mathematics to-day should be very largely the same as that given three generations ago is what the engineer does not appreciate. There is no tool used more frequently by engineers than the elementary differential and integral calculus, but the percentage who leave the Universities with that experience in its use which creates confidence, is extremely small, whereas with the ideal training they ought to use it with the same facility as they use the slide rule. Surely the mathematical curriculum extending as it does over 7 years, say between the ages of 13 and 20 years, could be revised by a united effort on the part of mathematicians and engineers to improve the present position.

I turn now from the ordinary student to the brilliant or research student who, after he has gained his experience will probably be frequently involved in problems of an abstruse character and is desirous of getting all the help he can from the more intricate mathematical tools. The more he knows of mathematics and the limitations of these mathematical tools, the better. It is not necessary that he should know Fourier mathematics, elliptic functions, Bessel's functions, Zonal or Spherical harmonics and all the other branches of higher mathematics as the mathematician knows them, but he ought to know what these are, and he ought to be able to recognize, with the aid of a handbook if necessary, any differential equation which he derives in the course of his work, so that he is in a position to lay his hand on the solution if one has been obtained or to know that no solution is possible except by graphical methods or successive approximation. It may be argued that he has only to turn up a text-book on differential equations but these books are written by mathematicians for mathematicians and deal extensively with the methods of obtaining a solution in which the engineer is not interested. When the engineer buys a machine tool the maker of that tool provides a handbook on its use. When are mathematicians going to take sufficient interest in the use of the tools they produce to furnish a similar handbook describing their use and their limitations? Such a book would be of the greatest value to engineers and when it is available, advanced students should be exercised in its use by a series of practical problems involving the use of various branches of higher mathematics. In electrical engineering, examples in the use of Fourier mathematics, Zonal and Spherical harmonics are easy to produce. In the oscillations and vibrations of structures and of instruments, the application of differential equations of the second, third and fourth order frequently occur and up to the eighth order are possible. In such a book it will be as important to include the equations which the mathematicians cannot solve as those which he can. In dealing with the cases of vibration for instance, the mathematician assumes the damping proportional to the velocity and in that case a simple solution is possible, but in nature, the damping proportional to the square of the velocity is much more frequent and no simple solution is then possible.

The engineer has to deal with problems of vibration very frequently and a thorough straightforward discussion of the whole subject would be most helpful to him including the conditions under which a harmonic impulse may excite the fundamental vibration, and also those under which the beats in two high frequency impulses will excite a vibration of beat frequency. Such cases do occur in practice and have not been fully explained by the mathematician; in fact the latter case has been stated to be impossible without any qualification as to conditions.

It is most important that the discussion of natural phenomena should not be subservient to any mathematical theory and that if experiment and theory do not agree the theory should be suspected. A knowledge of the limitations of mathematical tools is therefore of the greatest value to engineers.

Some mathematicians in dealing with practical problems have been unable to resist the temptation to ignore altogether certain terms in the equation of motion if thereby a neat solution became possible. There can be no objection to such an approximation if the approximation is admitted, and if the influence of the omitted term on the result is discussed where possible, but to ignore a term entirely and to present the solution as that of the practical problem only serves to widen the breach between the mathematician and the engineer when they should be working together to surmount the mountain of problems which still remains to be solved. The engineer cannot wait for the mathematician, he has to obtain a solution somehow, by that most expensive of all methods, trial and error, if there is no alternative.

Mathematicians have frequently looked askance at engineering science because of its frequent use of trial and error methods. I have known examples in which such a view was justified, but as a general rule the view is based upon a partial aspect of the problem, and the empirical method is employed because of the deficiencies of the mathematical tools or because of our ignorance of the working conditions.

The two heads under which the present discussion may be summarized are: (1) improvement in the teaching of mathematics in our schools and universities to enable the engineering student when he leaves the university to have a working knowledge of the application of the differential and integral calculus and differential equations of the more simple type; (2) the provision of a handbook by mathematicians for the use of engineers describing the use of the higher mathematical tools.

Under the first head I do not believe that any great benefit is to be obtained from the adoption of methods known as calculus dodging. The confidence in applying the mathematical tools required by the engineer can only be acquired by a student after considerable effort and very extensive application and the more frequent and varied the applications, the better. Collaboration between the mathematical and the engineering teacher could give these applications the practical form which will often arrest the attention of an engineering student who would show little interest in a problem of pure mathematics. Something will probably have to be sacrificed from the seven years mathematical curriculum of the average student which could probably be decided by a combined conference of mathematicians and engineers.

Under the second head, the engineer wants from mathematicians a candid admission of the limitations of the tools which they have to offer for the solution of the practical problems with which he has to deal. That will be a first step towards a combination. Let no formula be presented as a solution of a practical problem without the assumptions and limitations upon which it is based being stated clearly and let these assumptions be discussed and justified and let the influence of any limitations be analysed. Such candid discussion is bound to lead to further advance because many young mathematicians are restrained from tackling certain problems by the idea that they have been already solved, whereas the solutions are so limited by assumptions as to be worthless in practice, and the assumptions on which they are based are frequently hidden in the text.

On a subject of this kind one's opinions are naturally based largely upon the individual training and experience so that perhaps a few words of explanation of my personal experience will not be out of place.

Although I write as an engineer, I have a fair knowledge of higher mathematics and its applications. During the five years in which I had the privilege as a student and research scholar of attending the lectures given by Lord Kelvin in his higher Natural Philosophy class in the University of Glasgow, in various branches of mathematical physics, amplified by study at all the classes both compulsory and optional given by Professors Jack and Gibson on pure and applied mathematics, followed by a year's study under Helmholtz, Kundt, Planck and Fuchs in Berlin, I obtained a facility in the use of mathematical tools in the solution of physical problems which I have found very useful in practice. Twenty years spent in trying to solve practical problems of importance to the British Navy have confirmed an early impression which I learned from Kelvin that the number of practical physical problems in which mathematics offers a rigorous solution is very limited. Kelvin sometimes used graphical methods when analytical ones were impossible and these are frequently the only possible methods, but they are subject to considerable limitation, and when the influence of variations in the coefficients of an equation have to be studied they become very cumbrous.

I have been fortunate in having as my mathematical colleague Professor Wm. Burnside, a mathematician of great ingenuity and skill, who has assisted me considerably by devising methods of dealing approximately with otherwise untractable equations when the relative magnitudes of the various terms imposed by the practical conditions can be prescribed. This is a branch in which the mathematician can be of great use to the engineer.

I hope that these rambling remarks will be sufficient to act as an introduction to a good discussion, and that the outcome may be a better understanding between two types of scientific men who can be mutually helpful in advancing our knowledge of natural phenomena.

THE TEACHING OF THE ELEMENTS OF THE THEORY OF ELASTICITY TO ENGINEERING STUDENTS

By Professor E. G. Coker,
Engineering Laboratory, University College, London, England.

The mathematician may view the subject of the theory of elasticity as a branch of Applied Mathematics open for development in any direction he may choose or find possible with his students, but this is rarely practicable for a professor of engineering, who is charged with the duty of providing the elements of a theory which may be applied with safety to the design of the various kinds of structures required in the practice of his profession, and although the aim of both kinds of teaching is to some extent the same, yet the mathematician is more rightly concerned with the rigorous application of mathematics to obtain exact solutions of elastic problems which he may choose at will, while the engineer, recognizing the immense difficulties of the practical problems, not of his own choosing, has generally to abandon attempts at rigour of treatment and be content with a sufficient degree of approximation to the truth, which will enable him to deal with his complicated problems with safety by the application of a judicious blend of mathematical reasoning and practical judgment tempered by experiment.

The responsibilities which engineers have to take in practice are so great and the consequences of a mistake in calculation may be so disastrous, that the initial assumptions of theoretical calculation are usually subjected to severe scrutiny by experimental means before applying their conclusions, and this is the more necessary since the steel which the engineer usually employs may at some future time be overstressed and it will then not conform to elastic calculations in its new phase of plastic behaviour.

It has also to be borne in mind that the majority of young engineers, while possessing the right kind of ability for future success in their profession, are rarely endowed with the intuition of the mathematical mind, which seems to need no concrete illustrations and finds the symbolism of mathematics all sufficient.

It is mainly with the former type that we have to deal and it is very helpful to provide illustrations and experiments in the laboratory, to give a reality and meaning to the subject.

Like many others it has been my practice for many years to employ materials capable of great strains with little residual stress to illustrate such elementary matters as the position of the neutral axis, and the distribution of stress in a beam under uniform bending moment, such as you now see in the accompanying photographs, where changes in the ruled lines can be measured and the distribution of stress can be plotted from such measurements.

In a similar fashion it is not difficult when the strains are large for students to determine approximately the stresses in discontinuous members of plate form when loaded in their own plane, by ruling circles on one face and measuring the amounts and directions of the principal strains at a sufficient number of places. Excellent work of this kind has also been accomplished with steel members by several observers, but the skill required to measure small strains in metals is very great, and the labour so excessive that it is not feasible to introduce such work into a scheme for the practical instruction of engineering students preparing for a first degree, although the results of such research work are extremely valuable as lecture illustrations.

A means of rapidly obtaining a physical interpretation of the mathematical reasoning underlying the elementary theory of elasticity as usually taught in schools of engineering, is for students to examine the stresses produced in transparent models (when viewed in polarized light) of the beams, columns, curved members, cylinders and other bodies, which are dealt with from a mathematical standpoint, as it is found that they readily acquire the few principles of the optical theory of photo-elasticity, which are necessary, and that if such experiments are arranged to interleav with the mathematical treatment, it imparts a reality to the latter which is of benefit.

It is proposed here to describe a few typical experimental illustrations, which have been found useful, especially when employed at intervals in the manner described above, and not as here following one another rapidly in a brief paper.

Using circularly polarized white light it is necessary to show that pure tension stress applied to a member in a dark field causes that member to glow with a uniform colour depending on the stress intensity, and in the example now shown by natural colour photographs it can be observed, when commencing with no stress at all and a dark field of view, that the colours run through a cycle in which the colours correspond approximately to the following scale of stress, viz.: white, 550 lbs. per sq. in., yellow 700 lbs. per sq. in., purple red 950 lbs. per sq. in., and blue 1,000 lbs. per sq. in., and repeat themselves very nearly for multiples of these stresses, whether they are pure tension or pure compression. As tension diminishes the thickness and compression increases it, due account must be taken of this for great accuracy of observation, although this is a matter not especially necessary to discuss here.

Having now formed an idea of the colour scale in terms of stress, we can now pass on to consider a few cases where simple tension and compression stresses occur together, and illustrate many of the problems of simple beams associated with the well known relations $p/y = M/I$. Thus in the photograph now shown a beam of rectangular cross section is under uniform bending moment only, and its neutral axis is plainly visible at the centre and marked by a black band which we have already learned to associate with absence of stress when simple tension or compression occur alone. On each side there are parallel bands of colour, which not only show that the stress at each cross section is the same, but that it is nearly linear across the beam, although not quite.

We may leave this latter discrepancy for the moment until we come to consider the effects of the curvature which this type of loading produces.

It is perhaps of importance to students to point out at this stage that the limitations imposed on the use of the above simple formula for bending stress are many and important, for we cannot deviate from the mathematical conditions imposed by the usual proof without causing alterations in stress distribution to arise, which are often surprisingly large. Homogenity is one of these conditions. We cannot for example drill a small hole anywhere in such a beam, except at the neutral axis, without altering the stress conditions so greatly that our formula becomes a very inaccurate expression for the stress intensity in the neighbourhood of the discontinuity, and in the example now shown of such cases you observe that the stresses have now apparently become exceedingly complicated, as indeed they are. As a matter of fact it is quite easy to produce models of practical cases of riveting in plate-girders, which defy any attempt at calculation.

We may also try what is the effect of altering the cross-section, say by models of Tee and I shaped cross sections, and prove experimentally that the neutral axis of a straight beam passes through the centre of area of the cross section when under uniform bending moment.

It is also of interest to examine experimentally the stress distributions in a member loaded eccentrically to illustrate an important practical combination of varying simple stress due to a pull or push combined with a couple. Here for example is a plate tension member subjected to a pull P in a line coinciding with one edge, so that we have for any point of the cross section a stress p given approximately

$$(1) \qquad p=\frac{P}{A}\pm\frac{M}{I}y$$

and substantially verified by an inspection of the colour bands, with the further interesting results that the neutral axis is found to move slightly towards the centre of curvature of the beam bent by this eccentric application of load, while the passing of the elastic limit of the material is made evident by the non linear stress relation found at the highest loads.

Here again the curvature of the beam produces an important deviation, which can be further emphasized by illustrations of the effects produced by the uniform bending of beams of considerable curvature and examining the stress distribution to show that the stress distribution is no longer linear, but can be derived from a simple solution of the general equation $\nabla^4\chi=0$. This latter is independent of θ here and therefore can be easily shown to have a solution

$$(2) \qquad \chi=\alpha r^2+\beta r^2\log r+\gamma\log r+\delta$$

leading to the stresses

$$(3) \qquad \begin{cases} \widehat{rr} = 2\alpha + \beta(2 \log r + 1) + \gamma/r^2, \\ \widehat{\theta\theta} = 2\alpha + \beta(2 \log r + 3) - \gamma/r^2, \end{cases}$$

and connected to the uniform bending moment M by the relation

$$(4) \qquad M = \chi_i - \chi_o$$

where the subscript i refers to the internal boundary and the subscript o to the outer boundary.

From these three relations it is a useful exercise for a student to obtain the values of the constants in terms of the contour radii and the value of M.

In a comparison of the theoretical expressions for the stresses due to bending moment, it is now necessary to draw attention to the fact that the theory shows that $\widehat{rr}$ and $\widehat{\theta\theta}$ are here principal stresses and are represented in the colour picture by their difference, and it is useful to indicate the optical effect of the combined action of principal stresses p and q by a few simple experiments on a pair of tension or compression members set at right angles and pulled or pushed in various ways to show that the combined colour effect is always proportional to the difference $(p-q)$ of the stresses. One of these is shown by a colour photograph and draws especial attention to the case where $p=q$ in magnitude and sign resulting in the black field being restored. This latter experiment in fact indicates a useful method of evaluating the stress picture point by point by comparing it with a stressed tension or compression member, either set at one side in the field of view, or across the field of view at the point concerned and in the direction of one principal stress. If then r is the stress on this latter member, which produces a black field,

$$r = p - q$$

at that point.

For a comparison of the results of calculation and experiment it is sufficient to show here that $\widehat{rr} - \widehat{\theta\theta} = p - q$ and we can, for example, readily verify that the black band now shown in the accompanying photograph is at the place where calculation indicates that $\widehat{rr} - \widehat{\theta\theta}$ is zero, but that this is not a neutral axis since along this arc the principal stresses merely balance their optical effect. We have in fact a condition, as the theory indicates, of radial stress $\widehat{rr}$ and cross stress $\widehat{\theta\theta}$ both non-linear, resulting in a circular black band which intersects the cross sections at points which lie nearer to the inner contour than do the points where $\widehat{\theta\theta}$ is zero, a fact which can be readily made clear by measuring the lateral strains corresponding to

$$\widehat{rr} + \widehat{\theta\theta}$$

so that with these latter measurements and the results obtained optically $\widehat{rr}$ and $\widehat{\theta\theta}$ can be separated.

This affords a useful exercise for comparing the results of the theory with experiment, since it is not difficult to show from equations (3) and (4) that

$$\widehat{rr} = \frac{2M}{c}\left\{\frac{r_i^2 r_0^2}{r^2}.\log\frac{r_0}{r_i} + \log r(r_0^2-r_i^2) - (r_0^2\log r_0 - r_i^2 \log r_i)\right\}$$

$$\widehat{\theta\theta} = -\frac{2M}{c}\left\{\frac{r_i^2 r_0^2}{r^2}\log\frac{r_0}{r_i} - (1+\log r)(r_0^2-r_i^2) + (r_0^2\log r_0 - r_i^2 \log r_i)\right\}$$

where

$$c = (r_0^2 - r_i^2)^2 - 4r_0^2 r_i^2 \cdot \left(\log\frac{r_0}{r_i}\right)^2$$

thereby bringing out clearly the non-linear character of the stress distribution in radial sections of a curved beam and its complex character.

For students capable of following the analysis, it is of interest to compare the theories of stress distribution in hook shaped bodies with actual measurements, and also bodies like piston rings under forces applied at their outer contours by the pressure of the cylinder wall, while more difficult cases present themselves in the various forms of chain links used in practice, and shown under stress in the accompanying photographs, where we approach the existing limits of theoretical and experimental knowledge. Examples of thick cylinders consisting of simple or built-up type afford important practical cases, usually dealt with by calculation only, but which can be readily verified by photo-elastic experiment, as is shown by an illustration of a cylinder under internal pressure in which the stresses are found to agree with calculation. It is in fact an advantage of a theoretical course, illustrated by experimental data of this and like kinds, that an engineering student learns to appreciate that in almost every aspect of his future professional work on the stresses in structures and machines, he is attempting the solution of problems of the highest degree of complexity, which necessitate extreme caution in the manner in which he applies his elementary processes of stress calculation, as these latter are rarely adequate.

As another illustration of a useful character, we may take the subject of shear stress where for example it is easy to show that in a shallow beam of rectangular cross section the shear stress $\widehat{xy}$ is expressed by the usual relation

$$\widehat{xy} = \frac{S}{Iz}\int_{y}^{y_{max}} yzdy$$

and this is verified by the optical effects observed in the very short shallow beam shown, but not in a deep beam of the same span. It is also verified in the shallow encastre beams, of the type of the next example, at the points of inflection, where it may also be noted that not only are these points of inflection indicated by this form of distribution, but their position is also marked by black patches indicating the termination of the colour bands at the upper and lower contours, and therefore suggesting a means of testing the theory of continuous beams under complicated systems of loading by a comparison of the actual positions of these points with their calculated positions, a question of considerable importance in aeroplane and airship structures.

**Curved Beams, Rings and Chain Links*, by E. G. Coker, Honorary Member's Lecture, The Junior Institution of Engineers, 1922.

It would clearly be possible to give a much larger number of illustrations of the type already indicated in this paper to show that pictures of stress difference have an educational value for the training of students in the elements of the theory of elasticity, and some years' experience of the methods described have tended to confirm this, and also to show that it is just as useful for the professional engineer to study the effects of stress in models by this means and to check laborious structural and other calculations, or even to use experimental results so obtained in cases where theoretical calculation is very difficult or perhaps impossible.

List of Optical Illustrations, mostly Colour Photographs.

Slide	Description
1.	Plain Tension Member under no stress.
2.	Plain Tension Member, white, stress 500 lbs. per sq. in. approximately.
3.	Plain Tension Member, yellow, stress 700 lbs. per sq. in. approximately.
4.	Plain Tension Member, purple red, stress 950 lbs. per sq. in. approximately.
5.	Plain Tension Member, blue, stress 1000 lbs. per sq. in. approximately.
6.	Beam under pure bending moment.
7.	Beam with hole under pure bending moment.
8.	Stress distribution in a beam with two holes.
9.	Tension member under load along one edge.
10.	Tension member stress distribution.
11.	Two compression members crossed.
12.	Curved beam under pure bending moment.
13.	Circular link under stress.
14.	Circular link, stress distribution.
15.	Elliptical link under stress.
16.	Elliptical link, stress distribution.
17.	Thick cylinder under internal pressure.
18.	Thick cylinder, stress distribution.
19.	Shallow beam under shear showing parabolic distribution  .
20.	Deep beam under shear showing  distribution.
21.	Encastre beam showing shear distribution at points of inflection.

ABSTRACTS OF COMMUNICATIONS

SECTION IV

EFFECTS OF VARIATIONS IN HOOKE'S LAW ON IMPACT, THE THEORY OF BEAMS, AND ELASTICITY

By Professor E. R. Hedrick,
University of California, Southern Branch, Los Angeles, California, U.S.A.

It is well known that actual experiments indicate that many substances do not follow Hooke's Law with any reasonable degree of approximation. (See E. R. Hedrick, Engineering News, vol. 74, No. 12, Sept. 16, 1915, pp. 542-543). The same conclusions have been reached by others, particularly by Bach, and they seem to be accepted as standard for precise engineering practice in Germany. On the other hand, little seems to have been done in carrying out the consequences of the revised forms of these laws in the many other problems of engineering that depend upon Hooke's Law. The present paper is an effort to point out that some of these consequences can be carried out without serious difficulty, particularly in the case of problems involving impact, and in the case of some problems on beams. The difficulties encountered in other problems, particularly in connection with the theory of elasticity, are mentioned, and it is shown that some of these problems can be handled by means of elliptic and hyperelliptic integrals.

CALCULATION OF LONG TRANSMISSION SYSTEMS

By Professor T. R. Rosebrugh,
University of Toronto, Toronto, Canada.

A set of four linear equations in real quantities may be derived expressing the voltage squared, current squared, power and reactive power (z, w, x, y) at each end of the line in terms of similar quantities at the other end.

That is, there exists a quaternary linear homogeneous transformation with real coefficients between the two ends relative to these quantities. Its determinant is unity.

In addition $wz = x^2 + y^2$ and $w'z' = x'^2 + y'^2$.

When descriptive of a circuit with bi-terminal symmetry there are four real and independent parameters, otherwise there are six, and in either case many relations subsist between the 32 coefficients of the two sets besides those for the moduli.

P, Q and E^2 (x, y, z) define the position of a point and corresponding state of the circuit preferably to P, Q and E which are not homogeneous.

Any three independent conditions are sufficient to fix the state-point (x, y, z) in the model. In particular, quadratic equations serve for each of the 56 ways in which three of the eight quantities x, y, z, w, x', y', z', w' may be given.

A single condition relative to the circuit will determine a quadric surface or a plane locus and two a conic or a straight line.

THE BINARY LINEAR SUBSTITUTION OF DETERMINANT UNITY IN PROBLEMS OF GENERAL DYNAMICS, ACOUSTICS AND ELECTRICITY

By Professor T. R. Rosebrugh,
University of Toronto, Toronto, Canada.

Wherever, as in many cases of General Dynamics, including Acoustics and Electricity, there is an axi-symmetric determinant, whose coefficients are of the form $a_{rs}D^2+b_{rs}D+c_{rs}$ and when all the forces are zero but two, then one pair consisting of force and its co-ordinate is obtained from another like pair by a binary linear substitution in which if $\alpha, \beta, \gamma, \delta$, (rational functions of the time-differentiator D in general meromorphic) be the coefficients the determinant $\alpha\delta-\beta\gamma$ is unity.

The inverse substitution will have a rearrangement of the same four coefficients, namely, δ, $-\beta$, $-\gamma$, α, hence a new proof of Lord Rayleigh's Theorem of Reciprocity (which as ordinarily given is that theorem which is associated with the coefficients of the second position, namely, β and $-\beta$) and also three other similar theorems.

With sinusoidal forces, the binary substitution reduces to complex-number terms and has been applied by the writer and others in transmission line and network theory.

As a transformation it is in general non-commutative.

The constants of the matrix readily determine the constants of either the "general T" or the "general Π" circuit effectively equivalent to the actual arrangement and vice versa.

Products of matrices give the matrix for a tandem group.

The terms of the matrix for a parallel group of transmission lines are simply obtained and have been previously given by the writer.

The n-th power of the matrix may be put in a form involving hyperbolic functions.

Simple algebraic processes give the 2^n power of the matrix where n is integral and positive or negative.

A GENERALIZATION OF RAYLEIGH'S RECIPROCAL THEOREM

By Mr. J. R. Carson,
American Telephone and Telegraph Company, New York, N.Y., U.S.A.

Rayleigh's Reciprocal Theorem, in a slightly modified form, may be stated in the language of electric circuit theory as follows:

If a set of equi-periodic sinusoidal electromotive forces $V'_1, \ldots, V'_n$ in the n branches of a network produce a set of currents $I'_1, \ldots, I'_n$ and a second set of equi-periodic forces $V''_1, \ldots, V''_n$ produce a second current distribution $I''_1, \ldots, I''_n$ then,

$$\Sigma I_j'' V_j' = \Sigma I_j' V_j''.$$

The proof of this theorem, as given by Rayleigh, is applicable only to a quasi-stationary system of linear currents: that is, to a dynamical system in which the currents are the Lagrangian velocities. The practically important question as to whether a corresponding theorem holds in the case of radio transmission, where the assumptions of quasi-stationary systems break down, led to the derivation of the following theorem:

Let a distribution of impressed periodic electric intensity $F' = F'\ (x, y, z)$ produce a corresponding distribution of current intensity $u' = u'\ (x, y, z)$, and let a second distribution of equi-periodic impressed electric intensity $F'' = F''\ (x, y, z)$ produce a second distribution of current intensity $u'' = u''\ (x, y, z)$, then

$$\int (F'.u'')dv = \int (F''.u')dv,$$

the volume integration being extended over all conducting and dielectric media. F and u are vectors and the expression $(F.u)$ denotes the scalar product of the two vectors.

The proof of the theorem, which is limited to the case of non-magnetic media, starts with Maxwell's equations. By means of the retarded potentials, A and Φ, an integral equation in the current density is set up. As a consequence of the symmetry of the nucleus of the equations it follows at once that

$$\int (u'.G'')dv = \int (u''.G')dv,$$

where $G = F - \nabla\Phi$, and Φ is the retarded scalar potential. An application of the relation $div\ u = -(1/c)\ (\partial/\partial t)\rho$ and of Green's Theorem leads finally to the desired result

$$\int (u'.F'')dv = \int (u''.F')dv.$$

As an example of the practical importance of this theorem, it enables us to deduce the transmitting characteristics of an antenna from a knowledge of its receiving properties.

A NEW THEORY OF LONG DISTANCE RADIO-COMMUNICATION

By Professor G. W. O. Howe,
University of Glasgow, Glasgow, Scotland.

By treating the earth and the Heaviside layer as the conductors of a transmission line, a formula is developed for the transmission of electromagnetic waves around the earth over long distances. The calculated results are shown to be in agreement with observations, but necessitate the introduction of an arbitrary empirical factor which calls for further investigation.

DERIVATION OF THE DIFFERENTIAL EQUATIONS OF MOTION OF A PROJECTILE REGARDED AS A PARTICLE

By Professor W. H. Roever,
Washington University, St. Louis, Missouri, U.S.A.

The projectile is initially regarded as a satellite, and as such is subjected (1) to the earth's gravitational attraction and (2) to the resistance of the earth's moving atmosphere. The laws of relative motion determine the additional forces needed to account for its motion with respect to the earth. Thus the differential equations of motion of the projectile are found to involve the potential function (W) of the earth's weight field of force, the function (F) expressing the retardation due to resistance of the moving air, and the angular velocity of the earth's rotation. Replacing W by some of its approximations and F by some experimentally determined functions (methods for finding which are indicated), differential equations of motion are obtained which take into consideration weight, air resistance, wind, rotation of earth, convergence of verticals, curvature of layers of constant air density and also other influences, but not rotation of the projectile due to rifling of the gun. Such differential equations, involving functions and parameters representing various influences, must be obtained before differential corrections, based on variations of such functions, or parameters, can be found for computing the separate or combined effects of these influences.

A METHOD OF COMPUTATION FOR SOUND-RANGING DATA

By Dr. T. R. Wilkins,
Brandon College, Brandon, Canada.

A method is described which locates the asymptotes of the hyperbolas used in sound ranging without calculation. An automatic instrument is described embodying this method. The device also makes the temperature correction automatically.

DESIGN IN GUN CONSTRUCTION

By Professor H. C. Plummer,
Artillery College, Woolwich, England.

The principles of gun design are briefly discussed, and after reference to the evolution of the modern gun a direct comparison is instituted between a wire gun taken as typical with a gun built of solid tubes.

NEW AERODYNAMICAL CONCEPTIONS AND FORMULAE

By Dr. W. F. Gerhardt,
Flight Research Branch, Air Service, Dayton, Ohio, U.S.A.

This paper is intended to give an account of some new aerodynamical conceptions and formulae, suggested by the results of certain experiments, which cover the general case of an airplane cellule of N planes, and which permit of new and important designing conclusions, especially with regard to planes suitable for commercial aeronautics.

It begins with a formulation of the general problem of the development of the airplane wing cellule, and lays down briefly the important properties that such a cellule should possess. The general conclusion is drawn that if possible, these maxima of performance should be attained with as little mechanical adjustment as possible. A brief summary of the now known methods of extending these limits is presented, from which it is evident that before further conclusions are drawn the case of N planes must be discussed. To do this some simplifications of the theory are required.

Accordingly the general elements of the two dimensional flow about an aerofoil are presented, first for the infinite and then for the finite aerofoil. In this development is introduced the conception of the equivalent flow stream area, which is shown to be a basic expression of the whole phenomenon and to simplify the formulae evaluating the induced drag. With this conception it is comparatively easy to derive the fundamental relations for the case of N planes, and to indicate how the details of the vortex theory are applied to get the constants of the second order of magnitude in the relations.

In consequence the main desiderata, landing speed, load capacity, economy, etc., are shown to be dependent on the weight per unit flow stream area, and thus on the dimensions of the cellule in plan and front elevation, but quite negligibly on the *section* in side elevation.

Finally these formulae are checked, and the designing principle illustrated in the experiments on the "Cycleplane" the first man powered airplane to fly with the normal human output.

ARITHMETIC SOLUTION OF ENGINEERING PROBLEMS, WITH SPECIAL REFERENCE TO HYDRAULICS

By Professor R. W. Angus,
University of Toronto, Toronto, Canada.

A large number of problems met with in engineering deal with complicated phenomena very difficult to express mathematically except by approximate empiric equations, having variable coefficients and exponents. The use of the integral calculus in such cases involves formulas so difficult and long that their frequent employment is impracticable, and not a few cases occur in common practice where integration of the differential expressions is not possible.

Such problems occur frequently, as in the science of hydraulics, where one wants to study cases of unsteady motion in pipes and open channels, to know the effect of water hammer, or the rise in a river at various points due to floods. It is also frequently desired to know the history of the changes taking place, which involves plotting curves on a time base.

All of these problems,even those involving the most complicated empiric equations, may be solved by relatively simple arithmetic, usually involving a trial and error method in the solution, and the processes being carried out step by step with small intervals as abscissae. Each step is checked on itself and the result is a curve representing the net result of the change and also the history connected therewith, with great accuracy.

The method is not new, but has not been used as much as its value warrants, and is presented on account of its many practical applications.

MATHEMATICS IN INDUSTRIAL RESEARCH

By Dr. George A. Campbell,
American Telephone and Telegraph Company, New York, N,Y., U.S.A.

Selling Mathematics to the Industries; Mathematics in Electrical Communication; Industrial Mathematics as a Career; Training for Industrial Mathematics.

WHAT THE ENGINEER EXPECTS OF THE MATHEMATICIAN

By Professor C. F. Jenkin,
University of Oxford, Oxford, England.

The author suggests that mathematicians should give more help to engineers than they have done in the past. Some of the younger mathematicians are giving the help asked for; examples of the problems they are solving are quoted.

It is suggested that mathematicians should not wait to have questions put to them, because the engineer is often not aware that their problems can be solved mathematically, but should survey the science of engineering and volunteer assistance.

PHOTOGRAPHIC ENGINEERING SUPPLEMENT TO SECTION IV

Views of Developments in the Niagara System of the Hydro-Electric Power Commission of Ontario, visited by the Members of the Congress on invitation of the Commission, August, 14, 1924.

THE HYDRO COMMISSION AND NIAGARA DEVELOPMENT

THE Hydro-Electric Power Commission of Ontario was appointed by the Provincial Government in 1906 to act as Trustee for the municipalities of the Province in a co-operative scheme under which available water powers were to be utilized and electrical energy supplied "at cost" to the municipalities.

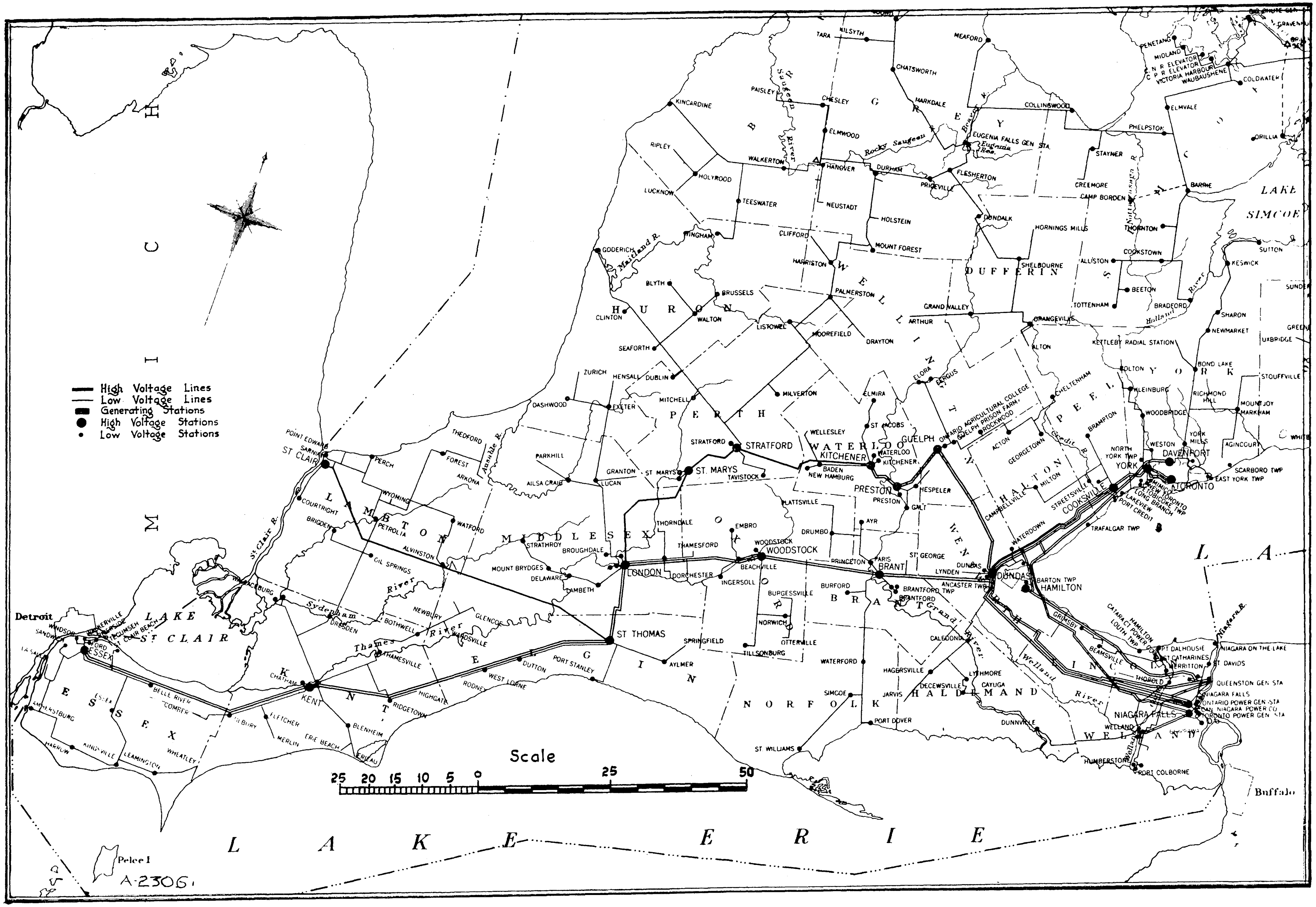
The Commission commenced to distribute power in 1910. From 3,500 horse power distributed at the end of that year the amount distributed increased to 700,000 horse power in 1924, to 1,000,000 horse power four years later, and provision has been made for an increase to 1,400,000 horse power in the near future.

Each municipality assumes responsibility for that portion of capital employed in delivering electrical energy to it. Each municipality sells electrical energy to its own local consumers at rates and under conditions approved by the Commission. The total capital investment is about $285,000,000.

Eighty per cent. of the electrical energy utilized for domestic service costs the consumer less than two cents per kilowatt-hour. Eighty-five per cent. of the electrical energy utilized for commercial light service costs the consumer less than three cents per kilowatt-hour. Seventy-five per cent. of the electrical power distributed by municipal systems and utilized for power service costs the consumer less than twenty-five dollars per horse power per annum. The Commission is also rapidly extending the benefits of cheap electrical energy to rural districts.

The number of hydro-electric generating plants operated by the Commission is twenty-two. Three of these having an aggregate capacity of over 850,000 horse power supply the "Niagara System." These are known as the Toronto, the Ontario, and the Queenston-Chippawa power plants. The Toronto power house is situated on the margin of the Niagara river just above the falls. The Ontario power house stands on the edge of the river in the gorge below the falls. The Chippawa-Queenston power house is located further down the river at the foot of the rapids near Queenston. It is the largest single hydro-electric plant in the world, its present capacity with nine units installed being about 550,000 horse power. Provision has been made for a tenth unit.

The Queenston-Chippawa development diverts water from the Niagara river above the falls and rapids by a special intake structure built in the Niagara river at the mouth of the Welland river at Chippawa. The water thus diverted passes along the deepened and enlarged channel of the Welland river for a distance of four miles, for which distance the direction of flow of the river is reversed. The water then enters the canal proper and traverses the Niagara peninsula for a distance of 8.75 miles, passing through an earth section, then into the rock-cut section of the canal, through a control-gate—an electrically-operated, roller-sluice gate of 48 feet clear span. The canal is 48 feet wide and lined with concrete. The depth of the water is from 35 to 40 feet, and at one point the floor of the canal is more than 140 feet below ground level. The canal terminates in a forebay, which is a triangle-shaped enlargement of the canal situated near Queenston at the edge of the Niagara gorge. Along the edge of the gorge, 320 feet above the river surface, is built the screen-house. Leaving the forebay the water passes beneath the screen-house, through wide, screened orifices, into the huge steel penstocks or tubes 14 to 16 feet in diameter which lead the water down the face of the cliff to the turbines of the power house. The turbines operate under a head of from 293 to 305 feet and rotate at a speed of 187.5 revolutions per minute.



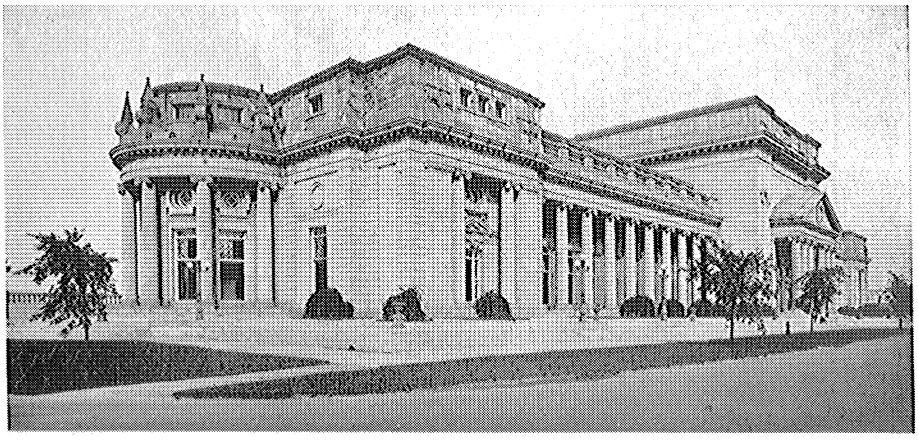
Transmission lines and stations in the Niagara System of the Hydro-Electric Power Commission

Reproduced by courtesy of Hamilton-Maxwell Inc.

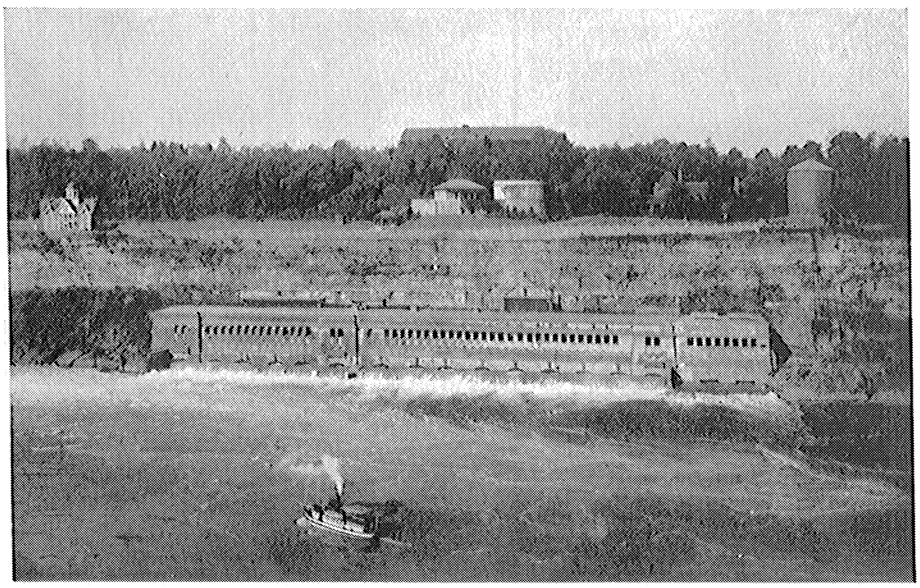
Niagara Falls as viewed from the air

Reproduced by courtesy of Hamilton-Maxwell Inc.

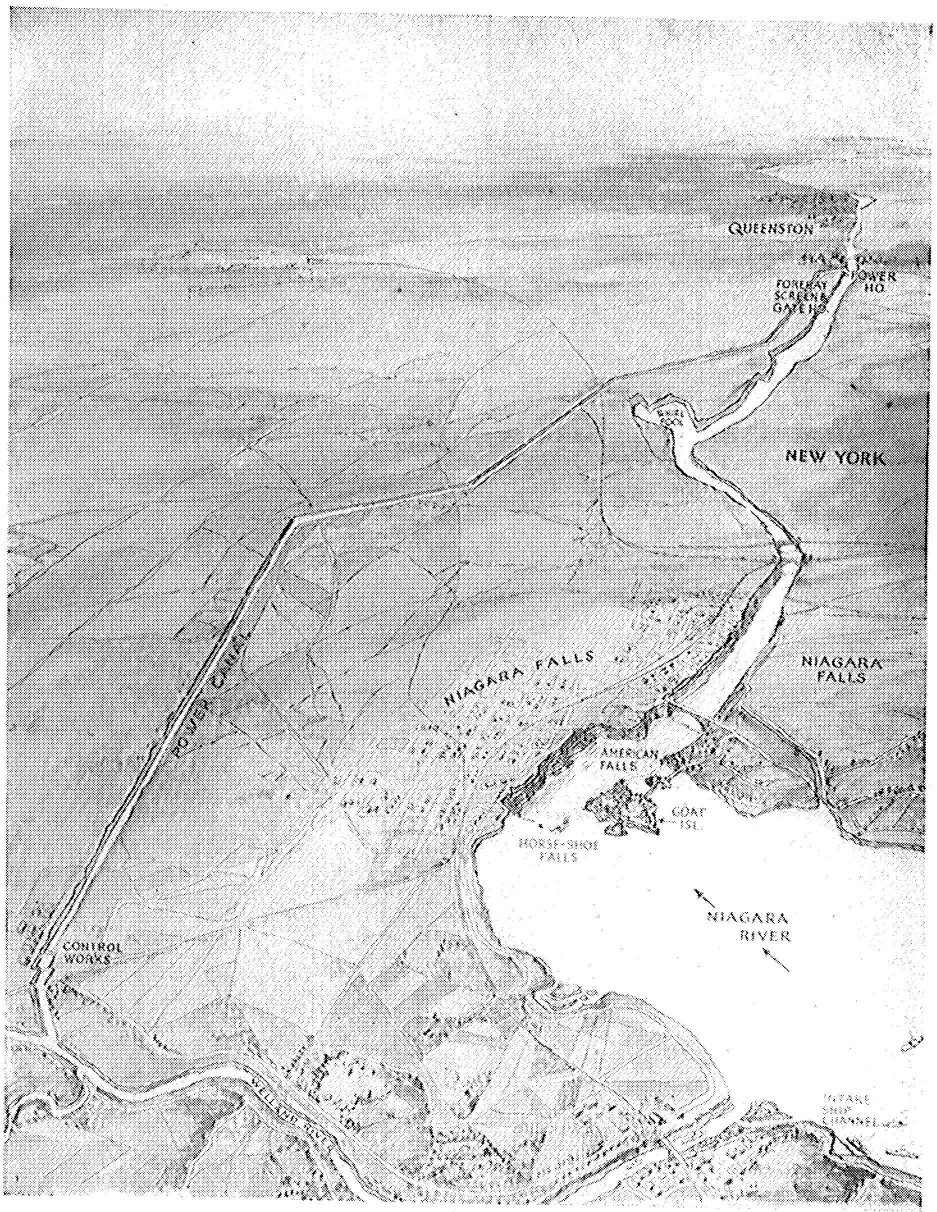
The Toronto power house here stands in the foreground. On the extreme right is seen the intake of the Ontario power house



Toronto power house on margin of river above the falls. Capacity 147,000 h.p. The turbines are located in wheelpits blasted through the solid rock to a depth of 150 feet below the surface.

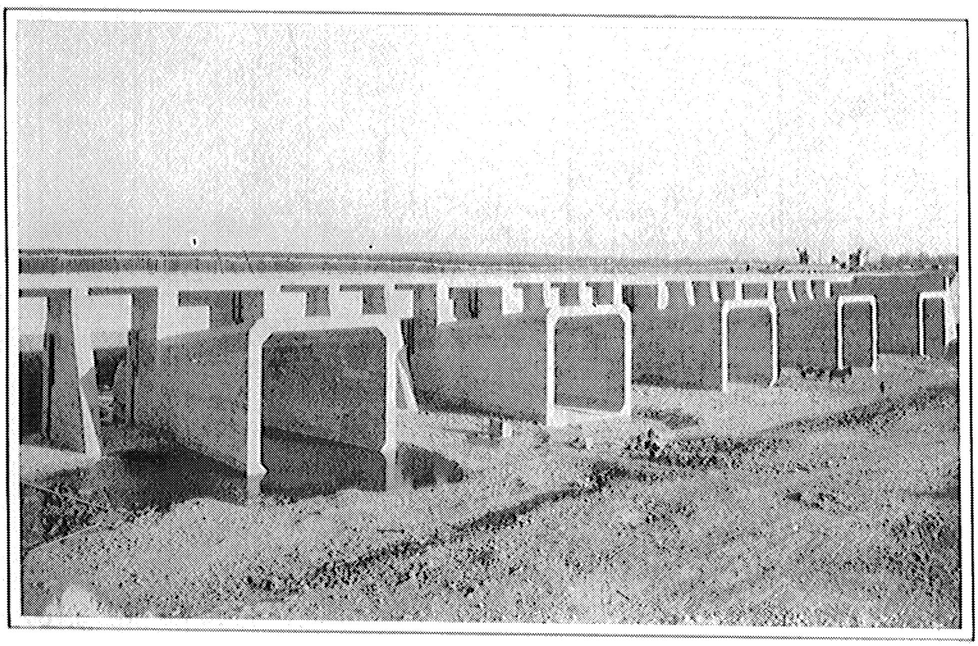


Ontario power house in gorge below the falls. Capacity 183,000 h.p.

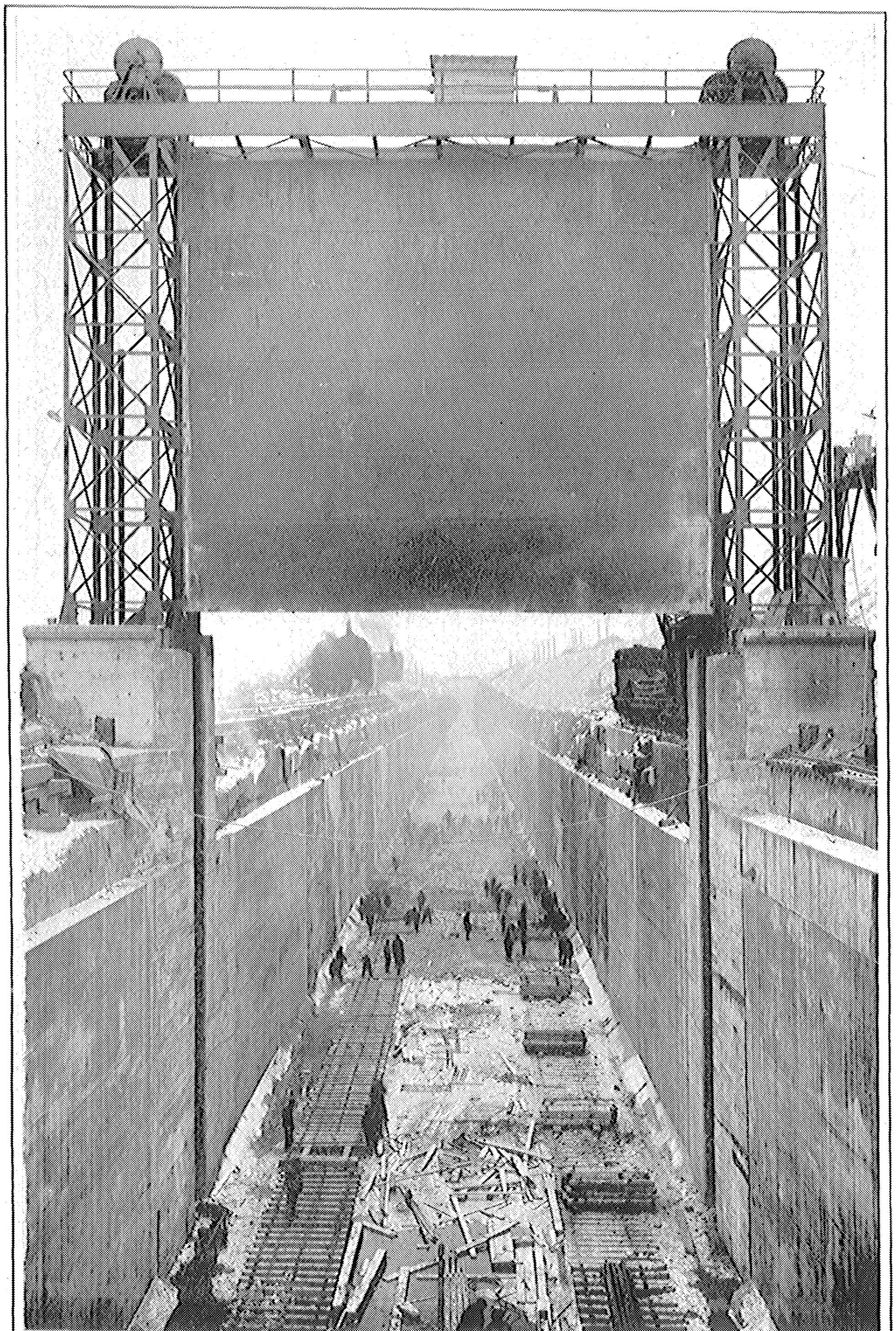


Diagrammatic bird's-eye view of Queenston-Chippawa power development. The view shows intake from Niagara river above the falls with Welland river section in foreground leading water to control works at upper end of canal which stretches to power house near Queenston, where water is returned to the lower Niagara river.

Entrance to Welland River section of power canal, showing intake works under construction and village of Chippawa.



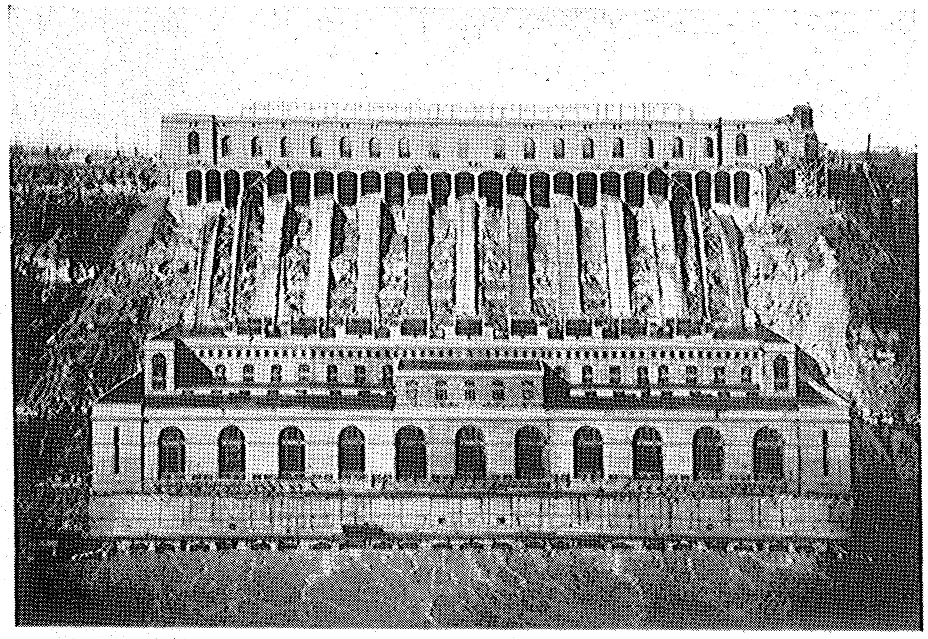
Intake tubes from forebay side before admitting water.



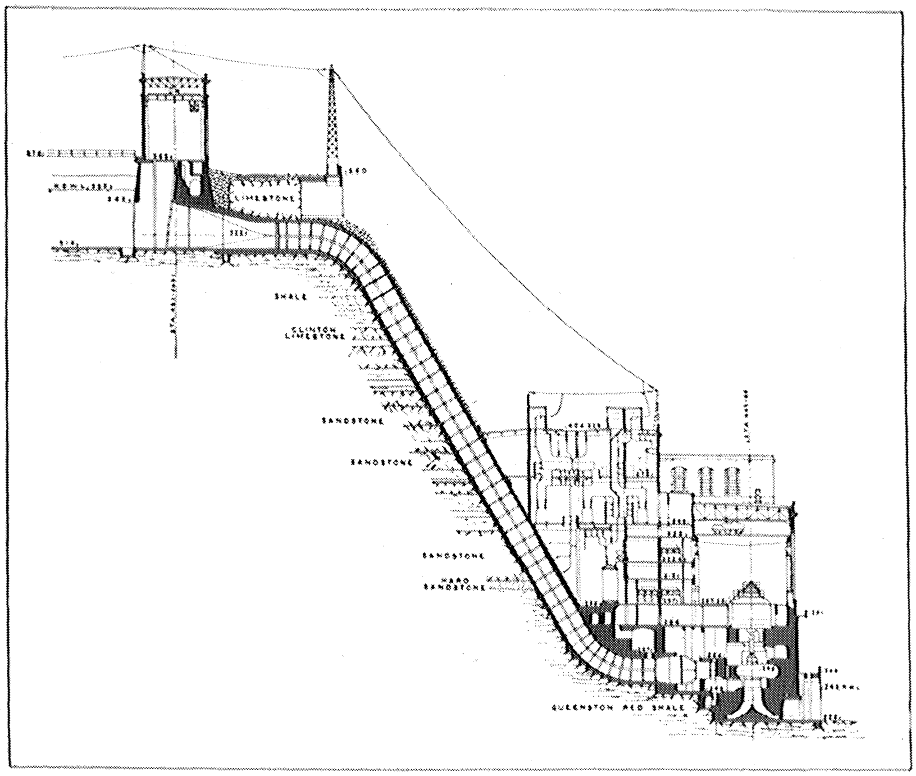
Electrically-operated control gate at entrance to rock-cut section of canal. The concrete facing of the walls of the canal are 35 to 40 feet in depth. The canal is 48 feet in width. Workmen are clearing the canal preparatory to admitting water.

Reproduced by courtesy of Hamilton-Maxwell Inc.

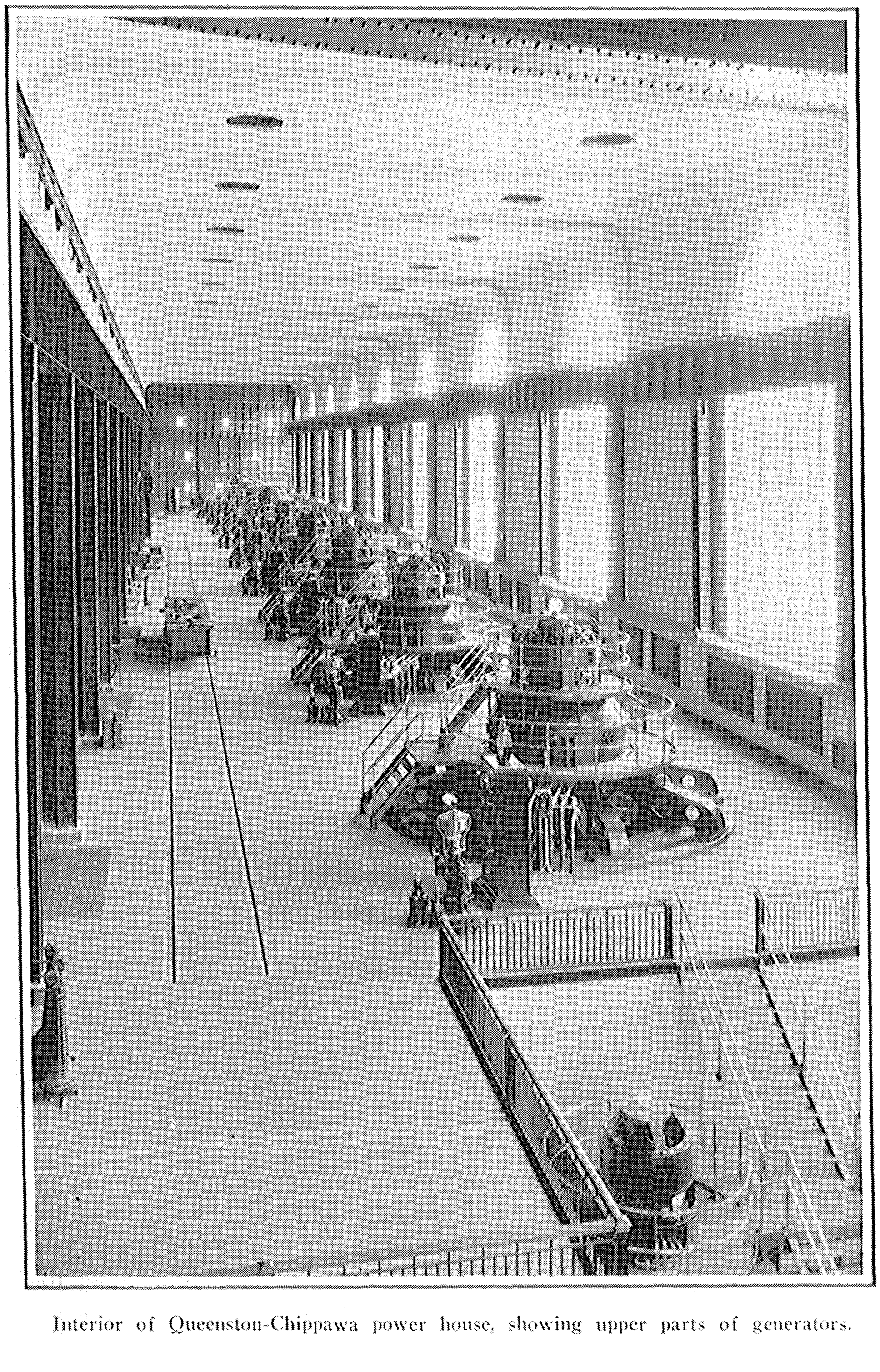
Lower end of canal with forebay and screen house. Beneath the surface at the widening of the canal is located a device for correcting the vortex motion of the water.



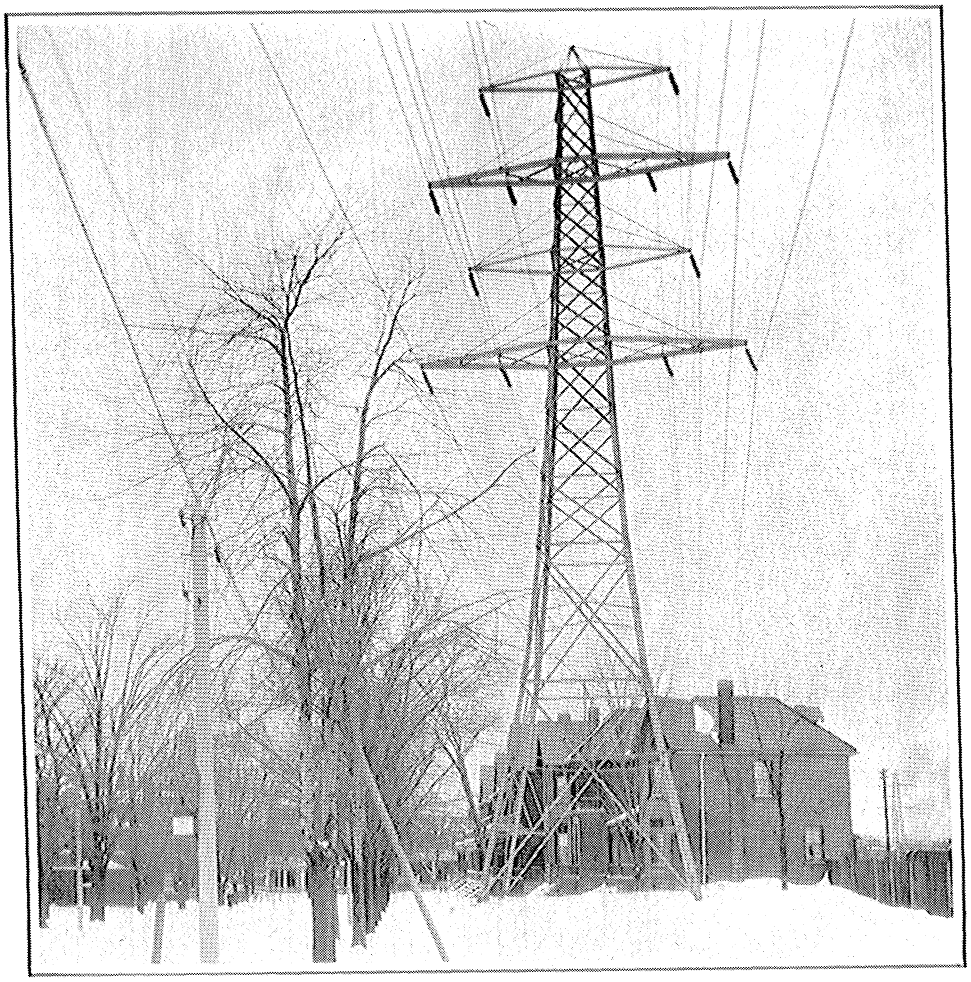
Queenston-Chippawa power house with nine units installed. It is located in the gorge of the Niagara river, the screen house standing on the top of the cliff and concrete covered steel pipes leading the water down the face of the cliff to the turbines. The power house stands on the edge of the river, is about 560 feet long and reaches half-way up the cliff.



Cross section through screen house and power house, showing pipe line down face of cliff with transmission line above.



Interior of Queenston-Chippawa power house, showing upper parts of generators.



Type of transmission line tower. High tension lines carry current of 110,000 volts. Current of double this voltage will in the near future carry power to Toronto from the falls of the Gatineau river in the Ottawa watershed, a distance of about 240 miles.

COMMUNICATIONS

SECTION V

STATISTICS, ACTUARIAL SCIENCE

A PROBLEM IN KEYNES'S TREATISE ON PROBABILITY

BY PROFESSOR EDWIN B. WILSON,
Harvard School of Public Health, Boston, Massachusetts, U.S.A.

On page 351 of *A Treatise on Probability*, the author, J. M. Keynes, states the following general problem: "If an event has occurred x times in the first y trials, its probability is $(r+x)/(s+y)$; determine the *a priori* probability of the event's occurring p times in q trials. If the *a priori* probability in question is represented by $\phi(p, q)$, we have

$$\phi(p, q)=\frac{r+p-1}{s+q-1}\,\phi(p-1, q-1)+\frac{s+q-1-r-p}{s+q-1}\,\phi(p, q-1). \tag{1}$$

I know of no solution of this, even approximate. But we may say that the conditions are those of supernormal dispersion as compared with Bernoulli's conditions. That is to say, the probability of a proportion differing widely from r/s is greater than in Bernoullian conditions; for when the proportion begins to diverge it becomes more probable that it will continue to diverge in the same direction."

It is my purpose to discuss the solution of this problem.

First let me explain the set-up Keynes gives. It is supposed that s trials have been made of which r have been successful so that the probability judging from experience thus far is r/s. If then y more trials are made of which x are successful the probability based on the total experience is $(r+x)/(s+y)$. This represents the hypothesis and we desire the probability of p successes in q trials subsequent to the initial r successes in s trials. How Keynes obtained his solution as a functional equation may be surmised. We have $\phi(p, q)$ as the chance of p successes in q trials. Suppose $q-1$ trials have been made and we are about to make the qth. If we are to have p successes we must have one of these two mutually exclusive possibilities after the $q-1$ trials: Either there were p successes to be followed by a failure, or there were $p-1$ successes to be followed by a success. Now $\phi(p-1, q-1)$ is the probability in the first alternative at the end of the $q-1$ trials. There have then been altogether $s+q-1$ trials and $r+p-1$ successes so that $(r+p-1)/(s+q-1)$ is the chance of success on the qth trial and the resultant probability of the first alternative is the product of this fraction by $\phi(p-1, q-1)$ as written. On the second of the alternatives the probability at the end of $q-1$ trials is $\phi(p, q-1)$; the total of trials is $s+q-1$ and the total of failures is $s+q-1-r-p$, so that the chance of failure is the quotient of the latter by the former and the resultant probability of the second alternative is as written. The sum of these is $\phi(p, q)$.

The exact solution of the difference equation (1) is easy to write down. It is

$$\phi(p, q) = \frac{q!}{p!(q-p)!} \frac{(r+p-1)!\ (s+q-1-r-p)!\ (s-1)!}{(r-1)!\ \ (s-r-1)!\ \ (s+q-1)\ !}. \tag{2}$$

That this satisfies (1) is seen by simple substitution. The solution was obtained not by divination but by the simple expedient of applying the reasoning Keynes used on his preceding page. This is as follows. We start with $S=r$ successes and $F=s-r$ failures and probabilities for success and failure on the $s+1=S+F+1th$ trial as $S/(S+F)$ and $F/(S+F)$ respectively. Now the chance of success or failure in the following q trials is calculated by using the total experience up to that stage. Consider q successes in q trials. The chance is

$$\frac{S}{S+F}\cdot\frac{S+1}{S+F+1}\cdots\cdots\frac{S+q-2}{S+F+q-2}\cdot\frac{S+q-1}{S+F+q-1}. \tag{3}$$

Consider $q-1$ successes followed by 1 failure. The result is

$$\frac{S}{S+F}\cdot\frac{S+1}{S+F+1}\cdots\cdots\frac{S+q-2}{S+F+q-2}\cdot\frac{F}{S+F+q-1}.$$

Consider 1 failure followed by $q-1$ successes. The result is

$$\frac{F}{S+F}\cdot\frac{S}{S+F+1}\cdots\cdots\frac{S+q-3}{S+F+q-2}\cdot\frac{S+q-2}{S+F+q-1}.$$

Wherever we place the single failure we have the same sequence of numbers in the denominator and the same numbers, but in different sequence, in the numerator. There are q mutually exclusive possibilities so that the value of the probability of $q-1$ successes and 1 failure, however distributed, is

$$q\,\frac{S.S+1\ldots(S+q-2).F}{S+F.S+F+1\ldots S+F+q-1}.$$

The reasoning extends itself immediately to p successes and $q-p$ failures. For each hypothesis as to the arrangement of the successes and failures, the denominator of the fraction is always the same, viz., the sequence $S+F$ to $S+F+q-1$ in continued product. The numerator will contain the numbers S, $S+1$, $\ldots$, $S+p-1$ and F, $F+1$, $\ldots$, $F+q-p-1$ arranged variously depending upon the particular hypothesis as to the succession of successes and failures. The binomial coefficient [$q!$ divided by $p!(q-p)!$] gives the number of different arrangements. Hence

$$\phi(p, q) = \frac{q!}{p!(q-p)!}\ \frac{S.S+1\ldots S+p-1.F.F+1\ldots F+q-p-1}{S+F.S+F+1\ldots S+F+q-2.S+F+q-1}. \tag{4}$$

This exception for the notation adopted is identical with (2).

In a footnote to p. 351 Keynes remarks of his preceding paragraph, which contains the solution of the simple case $p=q$ given above as (3); "This paragraph is concerned with a different point from that dealt with in Professor Pearson's article *On the Influence of Past Experience on Future Expectation*, to which it bears a superficial resemblance." He refers forward to his discussion of this contribution by Pearson where he proceeds to berate in no moderate terms the general conduct of Pearson's work and even the illustrative examples given. That which appears to have blinded Keynes with rage is Pearson's use of "Laplace's Rule of Succession" and the "equal distribution of ignorance"—to him two exceedingly irritating red rags. I use the term *blinded* with intent. For it seems to me that Pearson has solved in one way and in a reasonable way the general problem which Keynes stated and of which he knew no solution even approximate. The reference to Pearson is Phil. Mag., Ser. 6, vol. 13, 1907, 365-378. Pearson's problem is this: Given n trials with p successes and q failures; to find the *a priori* chance that in a second trial of m instances there will be r successes and s failures. This sounds a good deal like the same problem. It is difficult at any rate to see how two authors could have stated two problems with only "a superficial resemblance" in such exceedingly similar terms. Let us compare the results.

Pearson's answer* in our original notation of (2) is

$$\psi(p, q) = \frac{q!}{p!\,(q-p)!}\,\frac{(r+p)!\,(s+q-r-p)!\,(s+1)!}{r!\quad(s-r)!\quad(s+q+1)!}. \tag{5}$$

Hence

$$\psi(p, q) = \frac{r+p\,.\,s+q-r-p\,.\,s\,.\,s+1}{r\,.\,s-r\,.\,s+q\,.\,s+q+1}\,\phi(p, q).$$

The solution is therefore different; but a comparison between (5) and (2) shows that (2) goes over into (5) identically if in (2) we replace r by $r+1$ and s by $s+2$, that is

$$\psi(p, q; r, s) = \phi(p, q; r+1, s+2).$$

Speaking in terms of hypotheses this means that Pearson's result can be had from the assumption that after r successes in s trials the probability of a success is $(r+1)/(s+2)$ in place of r/s as Keynes assumes. This shows that for all practically legitimate cases in statistics or probability, *i.e.*, when r, s, p, q are none of them small numbers, the solutions of the two problems are interchangeable, and in any case that the exact solution of Keynes's problem can be read off that of Pearson's by the simple expedient of changing r and s respectively into $r+1$ and $s+2$. Further Pearson has discussed in the very reference cited the distribution of the probabilities with respect to their supernormal and skew characteristics which are matters of unsolved concern to Keynes and the solutions given by Pearson even if the problems are only superficially the same would hold for Keynes's problem with only insignificant changes. It sometimes happens that two very different problems have the same mathematical solution.

Now neither Pearson's solution (5) nor Keynes's as implied in (1) and elaborated in (2) and (4) makes complete sense when r and s are small. For example, if you have made no trials so that $r=s=0$, Keynes assumes that the chances on the first trial are $r/s=0/0$ and $(s-r)/s=0/0$ for success and failure

*Keynes, p. 380, quotes at length from Pearson, but with obviously careless errors within the quotation marks. I take the solution from Pearson, *loc. cit.*

respectively. Those are sensible answers—we know nothing of the probabilities before any trials if we are to judge the probabilities on the empirical basis. Pearson, however, gives the probabilities of success and failure as $(r+1)/(s+2)=1/2$ and $(s-r+1)/(s+2)=1/2$. This is not so good; it is the equal distribution of ignorance; it is not the best we can do, for Keynes has done better. On the other hand suppose one trial made with success so that $r=s=1$. Then Keynes gives as probabilities on the 2nd trial 1/1 for success and 0/1 for failures; we are certain to succeed if we have succeeded and sure to fail if we have failed. That is very bad; it not only violates common sense, it is counter to moral tenets as expressed in: "If at first you don't succeed, try, try again." Pearson on the other hand gives 2/3 for success and 1/3 for failure which is as reasonable as can be expected and if taken with its standard deviation or probable error is entirely reasonable. In fact his earlier estimate of 1/2 for the chance of success on one trial is not really bad if that estimate be taken with its standard deviation or probable error and be written as* 0.50 ± 0.34 which means that the true probability is as likely to lie between 0.16 and 0.84 as outside those limits, a statement certainly somewhat equivalent to Keynes's 0/0. In a similar way on the results 2/3 or 1/3 in the second case we have a probable error (computed by the Bernoulli formula) of 0.225 so that having succeeded once we assume the chance of success to be 0.667 ± 0.225 or to lie as often between 0.442 and 0.892 as outside those limits. Now the total range of the probability is only 1.00 and the probable error allows a range of 0.45 half the time so that the definiteness of 2/3 and 1/3 as calculated by Pearson is not great; one success in one trial may well determine the true probability as between 0.442 and 0.892 as often as outside those limits because as a matter of fact if we only extend the limits 0.05 we shall cover half the permissible range for any probability. The situation is totally different in this case with Keynes's solution. He has the probabilities 1 and 0 for success and failure; the standard deviation and probable error are alike equal to 0,—and by no possibility can a single trial foretell for all the future with absolute certainty what will happen. Most of Keynes's comments about the unreasonableness of Pearson's solution (p. 382) are more true of his own solution (if he had got it) than of that of the great leader of biometry.

I cannot refrain from feeling with Keynes that statisticians frequently overdraw their conclusions. Keynes (p. 351) cites an example from Czuber, he could have given a more flagrant case from C. S. Peirce†; he cites (p. 382) an example from the article by Pearson which goes on as follows: A sample of 100 of a population shows 10 cases of a certain disease, what percentage may reasonably be expected in a second sample? The solution Pearson gives is that the range 7.85% to 13.71% will contain half the cases and Keynes remarks that it appears unreasonable that so much of a conclusion could be drawn from such restricted data. Very likely, but Keynes own formulation (p. 351) would if he had solved the problem have led him to the same answer except that he

*If p is the probability, the standard deviation in n trials is $\sqrt{p(1-p)/n}$, and the probable error is 0.7645 of the standard deviation when n, np and nq are large.

†Peirce, *Chance, Love, and Logic*, p. 101, reprinted from Pop. Sci. Monthly, April, 1878.

would have had to start with 11 cases in a sample population of 102 instead of 10 cases in 100, and the conclusion could scarcely gain much in assurance by so slight a change. It is difficult to assign a definite and complimentary reason for Keynes's misunderstanding both his own problem and Pearson's solution, but I am tempted to think that possibly he confused the different significance of the following two problems: (1) Having taken one sample from a population what will be the distribution of other actual samples fairly drawn from the same population? (2) Having taken one sample from a population what will be the distribution of other hypothetical samples made up from a game of chance based upon the first sample? The former is a problem in actual empirical statistics, the latter one in the mathematical theory of probabilities and the answers may be very different. It was, as I understand the matter, precisely to bridge over this gap between theory and application that Lexis introduced his ratio $L=\sigma_{obs}/\sigma_B$ between the dispersion σ_{obs} actually observed in actual samples and the theoretical dispersion $\sigma_B=(npq)^{\frac{1}{2}}$ that would arise in a Bernoulli series with probabilities as defined in the sample. It is a fact that the Lexian ratio often differs significantly from 1, and that statistical inferences based on a single honestly drawn sample are wide of the mark, not because the sample is not fair but because different samples do differ among themselves more than the Bernoulli theory would indicate. But the way to an explanation of the hypernormal dispersion is not through the solution of Keynes's functional equation, as he seems to think may be the case, because Pearson had solved that equation in his berated contribution of 1907 and had shown that the hypernormality introduced was really not large.

ON A CERTAIN LAW OF PROBABILITY OF LAPLACE

BY PROFESSOR H. L. RIETZ,
University of Iowa, Iowa City, Iowa, U.S.A.

1. INTRODUCTION

The frequency distribution of the sum of n elements, each being a real number taken at random from a given interval 0 to $a(a>0)$ of uniform distribution, naturally appeals to the student of mathematical statistics as an elemental frequency distribution. The law of probability involved in this distribution was considered fairly early in the history of probability theory. In fact, Laplace developed a formula for the distribution as a limiting case of the de Moivre problem. In giving expression to this law of probability, Laplace* makes use of an interesting function which we may write in the form

$$(1) \qquad f(x)=\frac{1}{a^n(n-1)!}\left\{x^{n-1}-\binom{n}{1}(x-a)^{n-1}+\binom{n}{2}(x-2a)^{n-1}-\dots \right. \\ \left. +(-1)^{n-1}\binom{n}{n-1}[x-(n-1)a]^{n-1}\right\},$$

in which each parenthesis $(x-ga)$ with exponent $(n-1)$ is assigned the value 0 when the number $x-ga$ in the parenthesis is not positive. Throughout this paper $\binom{n}{r}$ means as usual the number of combinations of n things taken r at a time.

With $f(x)$ thus defined, Laplace found that $f(x)dx$ gives to within infinitesimals of higher order the probability that the sum of n elements such as are described above will fall into the interval from x_1 to x_1+dx. Although the law of probability in question was developed and applied by Laplace it has apparently received very little attention for many years. It is not unlikely that this lack of attention is due to certain dfficulties surroundng the function $f(x)$ given in (1). In the present paper we shall derive formula (1) by a geometric method involving mathematical induction. This method has the special advantage of giving simple geometric pictures of the law of distribution for different values of n.

*Laplace: *Théorie Analytique des Probabilités,* Troisième Edition (1820), pp. 257-263. Cf. E. Czuber: *Wahrscheinlichkeitsrechnung,* I (1914), p. 66.

2. The Frequency Distribution of the Sum of Numbers Taken at Random from Uniformly Distributed Values Between 0 and $a (a>0)$

Consider first the case of the sum of two numbers each taken at random from values between 0 and a. To find the relative frequency or probability with which sums fall into the interval x_1 to x_1+dx, we make the usual assumption of geometrical probabilities that the chance that a point taken at random on an assigned segment of a line is proportional to the length of the segment. We next employ the facts that, for the interval $0 \leqq x \leqq a$, the first number may be selected arbitrarily from values between 0 and x, $x_1 \leqq x \leqq x_1+dx$, and that we may then with certainty select a second number such that the sum falls into the interval dx. Hence, the frequency for sums in dx is

$$kxdx$$

where k is a constant.

That is,

$$y = kx \tag{2}$$

may be taken as the frequency or probability curve for the distribution of sums of two numbers when the sum $x \leqq a$.

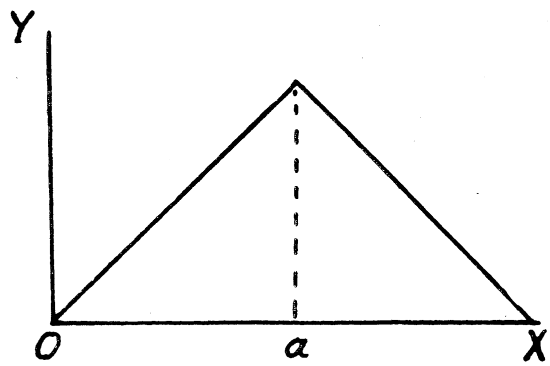


Fig. 1

Similarly, for the interval $a \leqq x \leqq 2a$, the first number selected may be taken arbitrarily as any number in the interval $x-a$ to a, $x_1 \leqq x \leqq x_1+dx$, and the second number may then be selected with certainty so that the sum falls into dx. That is,

$$y = k[a-(x-a)] = k(2a-x) \tag{3}$$

is the frequency or probability curve for the distribution of the sum x for the interval a to $2a$.

The equations (2) and (3) represent either frequency or probability depending on the determination k. If (2) and (3) are to represent probability curves, k is to be so chosen that the area under the curve (area of triangle, Fig. 1) is unity. Thus, the area

$$ka^2 = 1.$$

Then

$$k=\frac{1}{a^2}. \tag{4}$$

On the other hand, if (2) and (3) are to represent freqnency curves where a^2 is the total frequency represented by the area, we have

$$ka^2=a^2.$$

Then

$$k=1. \tag{5}$$

Unless specially stated to the contrary, we shall use frequency curves rather than probability curves in what follows. For our purposes, it is important to note that equations (2) and (3) are equivalent to the single equation

$$y=f(x)=x-\binom{2}{1}(x-a), \tag{6}$$

in which we assign the value zero to the parenthesis $(x-a)$ when $x-a$ is not positive.

Consider next the sum of three numbers each taken at random from values between 0 and a. To find the frequency of sums in the interval x_1 to x_1+dx, we use the facts that, for the interval $0\lesseqgtr x\lesseqgtr a$, we may select the sum of the two numbers* arbitrarily from the areas $\frac{x^2}{2}$ below x in Fig. 1, and that we may then with certainty select the third number so that the sum of the three falls between x_1 and x_1+dx. Hence the frequency in the interval x_1 to x_1+dx is

$$\frac{x^2}{2}dx,\ x_1\lesseqgtr x\lesseqgtr x_1+dx,$$

and the frequency curve is

$$y=\frac{x^2}{2} \tag{7}$$

shown as OF in Fig. 2.

When $a\lesseqgtr x\lesseqgtr 2a$, the frequency with which the sums of three numbers fall into x_1 to x_1+dx is given by selecting the sum of two of the numbers arbitrarily from the area of Fig. 1 between $x-a$ and x and adding the third number selected with certainty. The region of arbitrary selection has the area $\frac{3a^2}{4}-\left(x-\frac{3a}{2}\right)^2$.

Hence, the frequency curve FGF' of Fig. 2 is

$$y=\frac{3a^2}{4}-\left(x-\frac{3a^2}{2}\right)^2 \tag{8}$$

When $2a\lesseqgtr x\lesseqgtr 3a$, the frequency in question is given by the area of Fig. 1 from

*Here we make use of the usual assumption of geometrical probability that the frequency is proportional to the area and that the total area is a^2

$x-a$ to $2a$. That is, the frequency curve $F'L$ of Fig 2 is

$$y=\frac{(3a-x)^2}{2}. \tag{9}$$

It is easily verified that equations (7), (8), (9) giving segments of the curve in Fig. 2 are equivalent to the single equation,

$$y=f(x)=\frac{1}{2!}\left[x^2-\binom{3}{1}(x-a)^2+\binom{3}{2}(x-2a)^2\right], \tag{10}$$

if we assign the value zero to $(x-a)$ and $(x-2a)$ when $x-a$ and $x-2a$ are not positive.

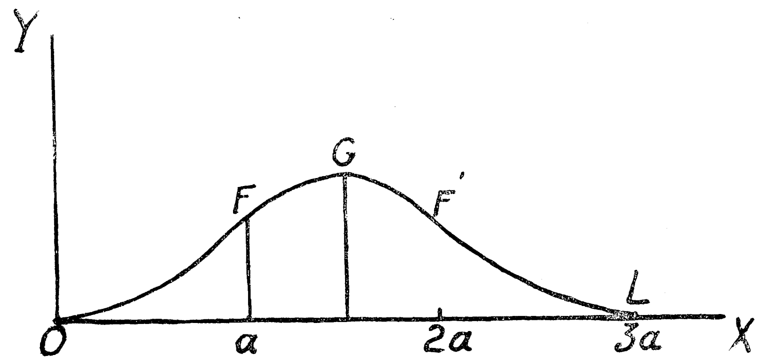


Fig. 2

Theorem. The frequency distribution of the sum of n elements each of which is a real number taken at random from a uniform distribution of range a is given by

$$\begin{aligned} y=f(x)=\frac{1}{(n-1)!}\Big\{x^{n-1}-\binom{n}{1}(x-a)^{n-1}+\binom{n}{2}(x-2a)^{n-1}-\dots \\ +(-1)^{n-1}\binom{n}{n-1}[x-(n-1)a]^{n-1}\Big\}, \end{aligned} \tag{11}$$

in which a term in parentheses with exponent $n-1$ is assigned the value zero when the expression in parentheses is not positive.

In (6) and (10) we have established this theorem for $n=2$ and $n=3$. Let us assume it to hold for $n=t$ and prove it for $n=t+1$. Thus, we assume that the sums of t numbers each taken at random from a range a are distributed so that the frequency in dx is given to within infinitesimals of higher order by

$$\begin{aligned} ydx=\frac{1}{(t-1)!}\Big\{x^{t-1}-\binom{t}{1}(x-a)^{t-1}+\binom{t}{2}(x-2a)^{t-1}-\dots \\ +(-1)^{t-1}\binom{t}{t-1}[x-(t-1)a]^{t-1}\Big\}dx. \end{aligned} \tag{12}$$

Introduce an additional element taken at random from the range a. Consider the distribution of the sum of the $t+1$ numbers in any interval $ga \leqq x \leqq (g+1)a$, where g is a positive integer such that $0 \leqq g \leqq t$. The frequency with which sums of the $t+1$ numbers fall between x_1 and x_1+dx can be found by taking values (sums of t numbers) at random from the area under the curve (12) of the distribution of sums of t numbers from $x-a$ to x and by then forming the required sum of $t+1$ numbers with certainty by the addition of a suitable number taken from the interval 0 to a.

This area in question from (12) is equal to

$$\frac{1}{(t-1)!}\int_{x-a}^{ga}\left\{x^{t-1}-\binom{t}{1}(x-a)^{t-1}+\binom{t}{2}(x-2a)^{t-1}-\dots+(-1)^{g-1}\binom{t}{g-1}[x-(g-1)a]^{t-1}\right\}dx$$
$$+\frac{1}{(t-1)!}\int_{ga}^{x}\left\{x^{t-1}-\binom{t}{1}(x-a)^{t-1}+\binom{t}{2}(x-2a)^{t-1}-\dots+(-1)^{g}\binom{t}{g}[x-ga]^{t-1}\right\}dx$$

(13)

$$=\frac{1}{(t-1)!}\left[\int_{x-a}^{x}\left\{x^{t-1}-\binom{t}{1}(x-a)^{t-1}+\dots+(-1)^{g-1}\binom{t}{g-1}[x-(g-1)a]^{t-1}\right\}dx+(-1)^{g}\binom{t}{g}\int_{ga}^{x}[x-ga]^{t-1}dx\right]$$
$$=\frac{1}{t!}\left\{x^{t}-\binom{t+1}{1}(x-a)^{t}+\dots+(-1)^{g-1}\binom{t+1}{g-1}(x-(g-1)a)^{t}+(-1)^{g}\binom{t+1}{g}[x-ga]^{t}\right\}.$$

Since we have considered any interval ga to $(g+1)a$ in deriving (13), we have that the frequency in dx of the sum of the $t+1$ numbers taken at random from 0 to a is given to within infinitesimals of higher order by

$$(14)\quad y\,dx=\frac{1}{t!}\left\{x^{t}-\binom{t+1}{1}(x-a)^{t}+\binom{t+1}{2}(x-2a)^{t}-\dots+(-1)^{t}\binom{t+1}{t}(x-ta)^{t}.\right\}dx.$$

But y in (14) is the same as y in (11) when $n=t+1$. Hence, the theorem is proved.

NOTE ON A SHORT METHOD OF COMPUTING TERMS AND SUMS OF TERMS OF THE ASYMMETRICAL BINOMIAL

By Dr. George F. McEwen,

Physical Oceanographer, Scripps Institution of Oceanography of the University of California, La Jolla, California, U.S.A.

This paper has been prepared for the investigator who finds it necessary to use the asymmetrical binomial in numerical applications, and wishes to make the computations with the least labour. It contains nothing that the trained statistician would not readily derive, but aims to bring final results together in a directly applicable form to meet the needs of the increasing number of investigators in various fields who might not make the transformations and substitutions required in order to avoid making the tedious computations involved when using elementary methods. There should be a means at hand to take the place of the familiar Bernoullian formula $.67449\sqrt{\frac{pq}{n}}$ for the probable error of a proportion, when the conditions that must be fulfilled in order that it may be legitimately used, are not realized.

The importance of the binomial series in statistical problems depends upon the well known law: If the probability that an event will happen in a single trial is p, and the probability that it will fail is $q=1-p$, then the probabilities of no failure, exactly one failure, exactly two failures, etc., and finally of all failures in n trials are given by the successive terms of the binomial expansion

$$(p+q)^n=p^n+np^{n-1}q+\frac{n(n-1)}{1\times 2}p^{n-2}q^2+\frac{n(n-1)(n-2)}{1\times 2\times 3}p^{n-3}q^3+\ldots+q^n,$$

$$r=1,\ 2,\ 3,\ 4,\ \ldots,\ (n+1)$$

where r = the number of the term. We can think of the terms of this expansion as frequencies of a frequency distribution. Short methods for computing either single terms or the sum of any number of terms in the general case where p and q are not equal and n has any value however great, can be based upon Pearson's type III frequency curve which bears the same relation to the asymmetrical binomial that the normal curve bears to the symmetrical binomial (Pearson 1906-7 and Greenwood 1913). The transformation required to obtain formulae directly applicable to computations pertaining to the binomial series, although not difficult, would not be made by many investigators who would profit by the results, presented here without proof.

Pearson's type III equation is

$$y = y_0 e^{-P_1\left(\frac{x}{a_1}\right)}\left[1+\frac{x}{a_1}\right]^{P_1}.$$

The rth term of the binomial expansion can be computed from this equation as follows (McEwen 1921)

$$P_1 = \frac{4npq}{1-4pq}-1,\ a_1 = \frac{P_1(p-q)}{2},\ y_0 = \frac{P_1^{(P_1+1)}\,e^{-P_1}}{a_1\Gamma(P_1+1)} = \text{approximately}^{*}\,\frac{1}{\sqrt{2\pi npq}},$$

$$x = r - nq - 1 + \frac{p-q}{2}.$$

The computation is facilitated by using Pearson's Table XXVI (Pearson 1914, pp. xiv and 37) which gives the values of the quantity in square brackets in the equation

$$\log_{10}y = \log_{10}y_0 - P_1[\log_{10}(1+X) - X\log_{10}e]$$

where $X = \dfrac{x}{a_1}$.

By means of the same frequency curve and its integral, the Incomplete Gamma Function, $I(u,\ P_1)$ tabulated in Table I (Pearson 1922, pp. 1-115) the sum of any number of terms can be computed as follows.:

The sum of the first r terms of the binomial expansion is approximately the tabular value of the function $I(u,\ P_1)$ where

$$u = \frac{P_1}{\sqrt{P_1+1}}\left\{\left[1-\frac{1}{a_1}\left(nq-\frac{p-q}{2}+\frac{1}{2}\right)\right]+\frac{r}{a_1}\right\}.$$

The P_1 of this paper is to be used for the p of his table and the a_1 of this paper is to be used for his a, (page vii). The function $\dfrac{P_1}{\sqrt{1+P_1}}$ is tabulated in Table IV (Pearson 1922, pp. 153-162). The approximation to the sum is about the same as that of the expression $\dfrac{1}{2}\left[1-\dfrac{2}{\sqrt{\pi}}\displaystyle\int_o^t e^{-t^2}dt\right]$ to the sum of the first r terms of the symmetrical binomial $\left(\dfrac{1}{2}+\dfrac{1}{2}\right)^n$ where (McEwen 1921) $t = \dfrac{n-2r+1}{\sqrt{2n}}$.

*An idea of the accuracy of the approximate expression is given by the value of the factor $\left(\sqrt{1+\dfrac{1}{P_1}}\right)\left(1-\dfrac{1}{12P_1}+\dfrac{1}{288P_1^2}+\ldots\right)$ by which the approximation $\dfrac{1}{\sqrt{2\pi npq}}$ should be multiplied. Also another exact expression for y_0 is $\dfrac{1}{\sqrt{npq}}\,\chi(P_1)$. The function $\chi(P_1)$ is tabulated in Pearson's Table IV (Pearson 1922, pp. 153-162).

Literature Cited

Pearson, Karl. 1906-07: *On the Curves which are Most Suitable for Describing the Frequency of Random Samples of a Population.* Biometrika, Vol. 5, pp. 172-175.

Pearson, Karl. 1914: *Tables for Statisticians and Biometricians.* Cambridge University Press, lxxxiii+143 pp., 5 r. in text.

Pearson, Karl. 1922: *Tables of the Incomplete* Γ *Function.* His Majesty's Stationery Office. London, xxxi+164.

Greenwood, M. 1913: *On the Errors of Random Sampling in Certain Cases not Suitable for the Application of a "Normal" Curve of Frequency.* Biometrika, Vol. 9, pp. 69-90.

McEwen, G. F. 1921: *Rapid Methods of Approximating to Terms in a Binomial Expansion.* Quarterly Publications of the American Statistical Association, Vol. 17, pp. 606-621.

ON A DISTRIBUTION YIELDING THE ERROR FUNCTIONS OF SEVERAL WELL KNOWN STATISTICS

BY MR. R. A. FISHER,
Statistical Department, Rothamsted Experimental Station, Harpenden, England.

1. THEORETICAL DISTRIBUTIONS

The idea of an error function is usually introduced to students in connection with experimental errors; the normal curve itself is often introduced almost as if it had been obtained experimentally, as an error function. The student is not usually told that little or nothing is known about experimental errors, that it is not improbable that every instrument, and every observer, and every possible combination of the two has a different error curve, or that the error functions of experimental errors are only remotely related to the error functions which are in practical use, because these are applied in practice not to single observations but to the *means* and other *statistics* derived from a number of observations.

Many statistics tend to be normally distributed as the data from which they are calculated are increased indefinitely; and this I suggest is the genuine reason for the importance which is universally attached to the normal curve. On the other hand some of the most important statistics do not tend to a normal distribution, and in many other cases, with small samples, of the size usually available, the distribution is far from normal. In these cases tests of *Significance* based upon the calculation of a "probable error" or "standard error" are inadequate, and may be very misleading. In addition to tests of significance, tests of goodness of fit also require accurate error functions; both types of test are constantly required in practical research work; the test of goodness of fit may be regarded as a kind of generalized test of significance, and affords an *a posteriori* justification of the error curves employed in other tests.

Historically, three distributions of importance had been evolved by students of the theory of probability before the rise of modern statistics; they are

Distribution	Due to	Date
Binomial expansion,	Bernoulli	c. 1700,
Normal curve,	Laplace, Gauss	1783,
Exponential expansion,	Poisson	1837;

of these the series of Bernoulli and Poisson, although of great importance, especially the latter, are in a different class from the group of distributions with which we are concerned, for they give the distribution of *frequencies*, and are consequently discontinuous distributions.

2. Pearson's χ^2 Distribution

In 1900 Pearson devised the χ^2 test of goodness of fit. If $x_1, x_2, \ldots, x_{n'}$, are the observed frequencies in a series of n' classes, and $m_1, m_2, \ldots, m_{n'}$ the corresponding expectations, then the discrepancy between expectation and observation may be measured by calculating

$$\chi^2 = S\frac{(x-m)^2}{m}\ .$$

The discrepancy is *significant* if χ^2 has a value much greater than usually occurs, when the discrepancy measured is that between a random sample, and the population from which it is drawn. To judge of this we need to know the random sampling distribution of χ^2. This distribution Pearson gave in his paper of 1900. The distribution for large samples is not normal; it is independent of the actual values of $m_1, \ldots, m_{n'}$; but it includes a parameter which, according to Pearson's original exposition, was to be identified with n', the number of frequency classes. Consequently in Pearson's original table, and in the fuller table given soon after by Elderton, the table is entered with the parameter n', which can take all integer values from 2 upwards.

More recently, it has been shown that Pearson neglected, as small quantities of the second order, certain corrections, which in fact do not tend to zero, but to a finite value, as the sample is increased. These corrections are, in fact, not small at all. In consequence of this, most of the tests of goodness of fit made in the last 25 years require revision. The important point is, however, that the distributions found by Pearson still hold, if we interpret n', not as the number of frequency classes, but as one more than the number of *degrees of freedom*, in which observation may differ from expectation. For example, in a contingency table of r rows and c columns, we ought not to take

$$n' = cr$$

but

$$n'-1 = (c-1)\ (r-1)$$

in recognition of the fact that the marginal totals of the table of expectation have been arrived at by copying down the marginal totals of the table of observations. For instance in a 3×5 table we should put $n'=9$, and not $n'=15$. In a 2×2 table $n'=2$, not $n'=4$.

One consequence of this is that it is more convenient to take $n=n'-1$, representing the number of degrees of freedom, as the parameter of the tables; in fact, to number the tables from (say) 1 to 50 instead of from 2 to 51. The real importance of n is shown by the fact that if we have a number of quantities $x_1, \ldots, x_n$, distributed independently in the normal distribution with unit standard deviation, and if

$$\chi^2 = S(x^2),$$

then χ^2, so defined, will be distributed as is the Pearsonian measure of goodness of fit; n is, in fact, the number of independent squares contributing to χ^2. The mean value of χ^2 is equal to n.

The χ^2 distribution is the first of the family of distributions of which I will speak, and like the others it turns up more or less unexpectedly in the distributions of a variety of statistics. In a noteworthy paper in 1908, "Student" investigated the error curve of the Standard Deviation of a small sample from a normal distribution, and with remarkable penetration he suggested a form for this error curve which has proved to be exact. The relation of this curve with that of χ^2 is close; if x stand for any value of a normal sample, $\bar{x}$ for the mean, and σ for the standard deviation of the population, then

$$\chi^2 = \frac{S(x-\bar{x})^2}{\sigma^2} = \frac{ns^2}{\sigma^2}$$

where n, the number of degrees of freedom, is one less than the number in the sample, and s^2 is the best estimate from the sample of the true *variance*, σ^2.

Another example of the occurrence of the χ^2 distribution crops up in the study of small samples from the Poisson Series. In studying the accuracy of methods of estimating bacterial populations by the dilution method I was led to the fact that the statistic

$$\chi^2 = \frac{S(x-\bar{x})^2}{\bar{x}}$$

when x is a single observation, and $\bar{x}$ the mean of the sample, is distributed wholly independently of the true density of the population sampled; for ordinary values of $\bar{x}$, but not for very small values, it is also distributed independently of $\bar{x}$, and its distribution is that of χ^2 with n one less than the number in the sample.

A similarly distributed index of dispersion may be used for testing the variation of samples of the binomial and multinomial series. The case of the binomial is interesting to economists, in that it leads at once to a test of the significance of the Divergence-Coefficient of Lexis. In fact, the method of Lexis was completed, and made capable of exact application, from the time of the first publication of the table of χ^2. I do not think, however, that this has been observed, either by exponents of the method of Lexis and his successors, or by exponents of the test of goodness of fit.

3. The General z Distribution

The most direct way of passing from the χ^2 distribution to the more general distribution to which I wish to call attention is to consider two samples of normal distributions, and how the two estimates of the variance may be compared. We have two estimates s_1^2 and s_2^2 derived from the two small samples, and we wish to know, for example, if the variances are significantly different. If we introduce hypothetical true values σ_1^2 and σ_2^2 we could theoretically calculate in terms of σ_1 and σ_2, how often $s_1^2 - s_2^2$ (or $s_1 - s_2$) would exceed its observed value. The probability would of course involve the hypothetical σ_1 and σ_2, and our formulae could not be applied unless we were willing to substitute the observed values s_1^2 and s_2^2 for σ_1^2 and σ_2^2; but such a substitution, though quite legitimate with large samples, for which the errors are small, becomes extremely misleading for small samples; the probability derived from such a substitution would be far from exact. The only exact treatment is to eliminate the unknown quantities σ_1 and σ_2 from the distribution by replacing the distribution of s by that of $\log s$, and so deriving the distribution of $\log s_1/s_2$. Whereas the sampling errors in s_1 are proportional to σ_1, the sampling errors of $\log s_1$ depend only upon the size of the sample from which s_1 was calculated.

We may now write

$$n_1 s_1^2 = \sigma_1^2 \chi_1^2 = \sigma_1^2 S_1(x^2)$$

$$n_2 s_2^2 = \sigma_2^2 \chi_2^2 = \sigma_2^2 S_2(x^2)$$

$$e^{2z} = \frac{s_1^2}{s_2^2} = \frac{\sigma_1^2}{\sigma_2^2} \cdot \frac{n_2 S_1(x^2)}{n_1 S_2(x^2)}$$

where S_1 and S_2 are the sums of squares of (respectively) n_1 and n_2 independent quantities; then z will be distributed about $\log \frac{\sigma_1}{\sigma_2}$ as mode, in a distribution which depends wholly on the integers n_1 and n_2. Knowing this distribution we can tell at once if an observed value of z is or is not consistent with any hypothetical value of the ratio σ_1/σ_2.

The distribution of z involves the two integers n_1 and n_2 symmetrically, in the sense that if we interchange n_1 and n_2 we change the sign of z. In more detail, if P is the probability of exceeding any value, z, then interchanging n_1 and n_2, $1-P$ will be the probability of exceeding $-z$.

Values of special interest for n_1 and n_2 are ∞ and unity. If n_2 is infinite, S_2/n_2 tends to unity and consequently we have the χ^2 distribution, subject to the transformation

$$e^{2z} = \frac{\chi^2}{n}, \; n_1 = n$$

or

$$z = \tfrac{1}{2} \log \frac{\chi^2}{n};$$

similarly if n_1 is infinite, we have the χ^2 distribution again with

$$z_p = -\tfrac{1}{2} \log \frac{1}{n} \chi^2_{1-p}$$

the curves being now reversed so that the P of one curve corresponds to $1-P$ of the other.

In the second special case, when $n_1 = 1$, we find a second important series of distributions, first found by "Student" in 1908. In discussing the accuracy to be ascribed to the mean of a small sample, "Student" took the revolutionary step of allowing for the random sampling variation of his estimate of the standard error. If the standard error were known with accuracy the deviation of an observed value from expectation (say zero), divided by the standard error, would be distributed normally with unit standard deviation; but if for the accurate standard deviation we substitute an estimate based on n degrees of freedom we have

$$t = \frac{x\sqrt{n}}{\sqrt{S(x^2)}}$$

$$t^2 = \frac{nx^2}{S(x^2)} = e^{2z} \text{ if } \begin{cases} n_1 = 1 \\ n_2 = n \end{cases}$$

consequently the distribution of t is given by putting $n_1 = 1$, and substituting $z = \frac{1}{2} \log t^2$.

The third special case occurs, when both $n_1 = 1$, and $n_2 = \infty$, and as is obvious from the above formulae, it reduces to the normal distribution with

$$z = \tfrac{1}{2} \log x^2.$$

In fact, one series of modifications of the normal distribution gives the χ^2 distributions, a second series of modifications gives the curves found by "Student", while if both modifications are applied simultaneously we have the double series of distributions appropriate to z.

Like the χ^2 distribution, the distribution found by "Student" has many more applications beyond that for which it was first introduced. It was introduced to test the significance of a mean of a unique sample; but as its relation to the z distribution shows, it occurs wherever we have to do with a normally distributed variate, the standard deviation of which is not known exactly, but is estimated independently from deviations amounting to n degrees of freedom. For example, in a more complicated form it gives a solution for testing the significance of the difference between two means, a test constantly needed in experimental work.

An enormously wide extension of the applications of "Student's" curves flows from the fact that not only means, but the immense class of statistics known as regression coefficients may all be treated in the same way; and indeed must be treated in the same way if tests of significance are to be made in cases where the number of observations is not large. And in many practical cases the number is not large; if a meteorologist with 20 years records of a place wishes to ask if the observed increase or decrease in rainfall is significant, or if in an agricultural experiment carried out for 20 years, one plot has seemed to gain in yield compared to a second plot differently treated, Student's curves provide an accurate test, where the ordinary use of standard errors or probable errors are altogether misleading.

The more general distribution of z like its special cases, crops up very frequently. I found it first in studying the error functions of the correlation coefficient. If the correlation, let us say, between pairs of n brothers, is obtained by forming a symmetrical table, we obtain what is called an *intraclass* correlation. If r is such a correlation, let

$$r = \tanh z, n_1 = n - 1, \; n_2 = n,$$

then this transformation expresses the random sampling distribution of r in terms of that of z, when n is the number in the sample.

It was the practical advantages of the transformation that appealed to me at the time. The distribution of r is very far from normal for all values of the correlation, (ρ); and even for large samples when the correlation is high; its accuracy depends greatly upon the true value of ρ, which is of course unknown, and the form of the distribution changes rapidly as ρ is changed. On the other hand the distribution of z is nearly normal even for small values of n; it is absolutely constant in accuracy and form for all values of ρ, so that if we are not satisfied with its normality we can make more accurate calculations.

The distribution shows a small but constant bias in the value of z, when r is derived from the symmetrical table; this bias disappears if instead of starting from the correlation as given by the symmetrical table we approach the matter from a more fundamental standpoint. Essentially we estimate the value of an intraclass correlation by estimating the ratio of two variances, the intraclass variance found by comparing numbers of the same class, and the variance between the observed means of the different classes. From n classes of s observations each, we have $n(s-1)$ degrees of freedom for the intraclass variance, and $n-1$ degrees of freedom for the variance of the means. From the definition of the z distribution it will obviously be reproduced by the errors in the ratio of two independent estimates of the variance, and if we estimate the variances accurately the bias in the estimate of the correlation will be found to have disappeared; it was, in fact, introduced by the procedure of forming the symmetrical table.

The practical working of cases involving the z distribution can usually be shown most simply in the form of an analysis of variance. If x is any value, $\bar{x}_p$ the mean of any class, and $\bar{x}$ the general mean, n the number of classes of s observations each, the following table shows the form of such an analysis:

ANALYSIS OF VARIANCE

Variance	Degrees of Freedom	Sum of Squares	Mean Square
Between classes.......	$n_1=n-1$	$sS_1^n(\bar{x}_p-\bar{x})^2$	s_1^2
Within classes.........	$n_2=n(s-1)$	$S_1\ (x-\bar{x}_p)^2$	s_2^2
Total.............	$ns-1$	$S_1^{ns}(x-\bar{x})^2$	

The two columns headed Degrees of Freedom and Sum of Squares must add up to the totals shown; the mean squares are obtained by dividing the sums of squares by the corresponding degrees of freedom, then $z=\log s_1/s_2$ may be used to test the significance of the intraclass correlation, or the significance of its deviations from any assigned value.

4. MULTIPLE CORRELATIONS

The same method may be used to solve the long outstanding problem of the significance of the multiple correlation. If y is the dependent variate, and $x_1, x_2, \ldots, x_p$ are p independent variates, all measured from their means, and if the regression of y on $x_1, \ldots, x_p$ is expressed by the equation

$$Y=b_1x_1+\ldots+b_px_p$$

such that the correlation of y with Y is R, then R is termed the multiple correlation of y with $x_1, \ldots, x_p$, and the ingredients of the analysis of variance are as follows:

Variance	Degrees of Freedom	Sum of Squares	
Of regression formula.......	p	$S(Y^2)$	$=nR^2\sigma^2$
Deviations from regression formula.................	$n-p-1$	$S(y-Y)^2$	$=n(1-R^2)\sigma^2$
Total.................	$n-1$	$S(y^2)$	$=n\sigma^2$

For samples from uncorrelated material the distribution of R may therefore be inferred from that of z, the actual curve for n observations and p independent variates being

$$df=\frac{\frac{n-3}{2}!}{\frac{n-p-3}{2}!\,\frac{p-2}{2}!}(R^2)^{\frac{1}{2}(p-2)}(1-R^2)^{\frac{1}{2}(n-p-3)}d(R^2)$$

which degenerates when $p=1$, into the better known distribution for a single independent variate

$$df=\frac{2\,\frac{n-3}{2}!}{\frac{n-4}{2}!\,\sqrt{\pi}}(1-R^2)^{\frac{n-4}{2}}dR,$$

a distribution first suggested by "Student"[7], which has been established since 1915.

5. The Correlation Ratio

The distribution, for uncorrelated material of the correlation ratio η, is clearly similar to that of the multiple correlation, R, and resembles the case of the intraclass correlation when the number of observations varies from class to class.

A number of values of the variate y are observed for each of a series of values of the variate x; n_p is the number observed in each array, $\bar{y}$ the mean of the observed values and $\bar{y}_p$ the mean of any array; the variance of the variate y may be analysed as follows, there being a arrays:

Variance	Degrees of freedom	Sum of Squares	
Between arrays............	$a-1$	$S\{n_p(\bar{y}_p-\bar{y})^2\}$	$=N\eta^2\sigma^2$
Within arrays.............	$S(n_p-1)$	$S(y-\bar{y}_p)^2$	$=N(1-\eta^2)\sigma^2$
Total................	$S(n_p)-1$	$S(y-\bar{y})^2$	$=N\sigma^2$

η^2 is thus distributed just like R^2. The transformation is

$$\frac{\eta^2}{1-\eta^2}=\frac{n_1}{n_2}e^{2z}=\frac{S_1(x^2)}{S_2(x^2)}.$$

If the observations are increased so that $n_2 \rightarrow \infty$, then

$$n_2 \frac{\eta^2}{1-\eta^2}$$

tends to be distributed in the χ^2 distribution corresponding to $(a-1)$ degrees of freedom, while for the multiple correlation with p independent variates

$$n_2 \frac{R^2}{1-R^2}$$

tends to be distributed in the χ^2 distribution with p degrees of freedom. These are two examples of statistics not tending to the normal distribution for large samples.

More important than either of these is the use of the z distribution in testing the goodness of fit of regression lines, whether straight or curved. If Y stand for the function of x to be tested representing the variation of the mean value of y for different values of x, let there be a arrays and let q constants of the regression formulae have been adjusted to fit the data by least squares, then the deviations of the observations from the regression line may be analysed as follows

Variance due to	Degrees of freedom	Sum of Squares	
Deviation of array mean from formula...........	$a-q$	$S\{n_p(\bar{y}_p - Y_p)^2\}$	$=N\sigma^2(\eta^2-R^2)$
Deviation within array.....	$N-a$	$S(y-\bar{y}_p)^2$	$=N\sigma^2(1-\eta^2)$
Total...............	$N-q$	$S(y-Y)^2$	$=N\sigma^2(1-R^2)$

where R is the correlation between y and Y. The transformation this time is

$$\frac{\eta^2-R^2}{1-\eta^2} = \frac{S_1}{S_2} = \frac{n_1}{n_2} e^{2z},$$

and if the sample is increased indefinitely

$$n_2 \frac{\eta^2-R^2}{1-\eta^2}$$

tends to χ^2 distribution for $(a-q)$ degrees of freedom.

This result is in striking contrast to a test which is in common use under the name of Blakeman's criterion, which is designed to test the linearity of regression lines, and which also uses the quantity η^2-r^2. Our formula shows that, for large samples, with linear regression,

$$(N-a)\frac{\eta^2-r^2}{1-\eta^2}$$

has a mean value $a-2$. The failure of Blakeman's criterion to give even a first approximation, lies in the fact that, following some of Pearson's work, the number of the arrays is ignored, whereas, in fact, the number of arrays governs the whole distribution.

SUMMARY

The four chief cases of the z distribution have the following applications

I. Normal Curve	II. χ^2	III. Student's	IV. z
Many statistics from large samples	Goodness of fit of frequencies Index of dispersion for Poisson and Binomial samples Variance of normal samples	Mean Regression coefficient Comparison of means and regressions	Intraclass correlations Multiple correlation Comparison of variances Correlation ratio Goodness of fit of regressions

COMPARISON OF DISTRIBUTION FORMULAE

	Constant factor	Frequency
χ^2	$\dfrac{1}{2^{\frac{n-2}{2}}\,\frac{n-2}{2}!}$	$\chi^{n-1}e^{-\frac{1}{2}\chi^2}d\chi = n^{\frac{1}{2}n}e^{nz-\frac{1}{2}ne^{2z}}dz$
Normal	$\dfrac{2}{\sqrt{2\pi}}$	$e^{-\frac{1}{2}x^2}dx = e^{z-\frac{1}{2}e^{2z}}dz$
Student's	$\dfrac{\frac{n-1}{2}!\;2n^{\frac{1}{2}n}}{\frac{n-2}{2}!\;\sqrt{\pi}}$	$(n+t^2)^{-\frac{n+1}{2}}dt = \dfrac{e^z\,dz}{(e^{2z}+n)^{\frac{1}{2}(n+1)}}$
z	$\dfrac{2\,.\,\frac{n_1+n_2-2}{2}!}{\frac{n_1-2}{2}!\;\frac{n_2-2}{2}!}\,n_1^{\frac{1}{2}n_1}n_2^{\frac{1}{2}n_2}$	$\dfrac{e^{n_1z}dz}{(n_1e^{2z}+n_2)^{\frac{1}{2}(n_1+n_2)}}$

Addendum Aug., 1927.

Since the International Mathematical Congress (Toronto, 1924) the practical applications of the developments summarized in this paper have been more fully illustrated in the author's book *Statistical Methods for Research Workers* (Oliver and Boyd, Edinburgh, 1925). The Toronto paper supplies in outline the mathematical framework around which the book has been built, for a formal statement of which some reviewers would seem to have appreciated the need.

SOME POINTS IN THE GENERAL THEORY OF GRADUATION

By Mr. Robert Henderson,
Vice-President, Equitable Life Assurance Society, New York, N.Y., U.S.A.

It is proposed in this paper to investigate the problem of graduation from the standpoint of the determination, in conjunction with a given set of observed facts, of a hypothesis such that the compound probability of the truth of the hypothesis and the occurrence of the observed facts shall be as large as possible. This compound probability is the product of the total probability that the observed facts will occur and the probability that if they occur the hypothesis will be true. Since from the standpoint of this problem the first probability is a constant it is required to determine the most probable hypothesis where the observed facts are given.

The compound probability may also be expressed as the product of the *a priori* probability of the truth of the hypothesis and the probability, on the assumption that the hypothesis is true, that the observed facts would occur. In the cases to which graduation is usually applied the hypothesis takes the form of a set of assumed values for a series of probabilities connected with a series of values of a variable. These values of the variable usually are separated by equal intervals. Let us denote the probability corresponding to the value x of the variable by q_x. Then if Q denotes the probability of the observed facts, E_x the number of instances observed at value x and θ_x the number of such instances in which the probability is realized we have

(1)
$$\log Q=\sum\left\{\theta_x\log q_x+(E_x-\theta_x)\log(1-q_x)\right\}$$

where Σ extends over all observed values of x. This assumes that the observed facts include the particular instances in which the probability is realized and not merely the number of such instances. If only the number is supposed to have been observed there would be an additional term in log Q which would, however, be independent of q_x and would not therefore affect the relative probabilities on different hypotheses as to the values of q_x.

If now we suppose that all possible sets of values of q_x are equally likely, then the set which is *a posteriori* most probable is that which makes log Q a maximum. Differentiating then with respect to q_x and equating to zero we have

$$\theta_x/q_x-(E_x-\theta_x)/(1-q_x)=0,$$

or

$$\theta_x(1-q_x)-(E_x-\theta_x)q_x=\theta_x-E_xq_x=0,$$

or

(2)
$$q_x=\theta_x/E_x.$$

This shows that on this supposition the crude or ungraduated values of q_x constitute the most probable set. Adjustment or graduation only enters when it is assumed that a smooth series of values of q_x is more probable *a priori* than an irregular series and in order that the problem should be capable of solution in precise mathematical form it is necessary to adopt some rule with respect to the relative *a priori* probabilities of different possible hypothetical series.

One method of dealing with this matter is to assume that if the successive terms of the series are the values, for successive values of a variable, of a function with a few arbitrary constants the series is a smooth one, that all such series are equally likely and that it is not necessary to consider any series not so determined. The problem is then comparatively simple theoretically and consists in so determining the values of the arbitrary constants in any proposed function as to make the probability of the observed facts as great as possible. If functions of different forms are being considered for the same series this would be done for each and the absolute maximum then selected. In general where $a_1, a_2, \ldots, a_s$ are the arbitrary constants in q_x we have a series of equations of the form

$$(3) \qquad \Sigma\{\theta_x/q_x-(E_x-\theta_x)/(1-q_x)\}\frac{dq_x}{da_r}=0, \quad (r=1, 2, \ldots, s).$$

This is the general solution of the problem on the assumptions indicated. It is necessary, however, to test whether the observed values are such as could reasonably be expected to arise from the smooth series of true values. Otherwise the hypothesis is not satisfactory. The theoretically correct basis of this test is the probability, calculated on the basis of the adjusted values, of the observed facts. We have seen that if Q be this probability then

$$\log Q=\Sigma\{\theta_x \log q_x+(E_x-\theta_x)\log(1-q_x)\}.$$

If we assume that $\theta_x=E_xq_x+h_x$ this equation takes the form

$$(4) \qquad \log Q=\Sigma\{(E_xq_x+h_x)\log q_x+(E_x-E_xq_x-h_x)\log(1-q_x)\}.$$

On the same assumption the maximum value of the probability of the observed facts, is

$$\log Q'=\sum\left\{(E_xq_x+h_x)\log\left(q_x+\frac{h_x}{E_x}\right)+(E_x-E_xq_x-h_x)\log\left(1-q_x-\frac{h_x}{E_x}\right)\right\}.$$

The logarithm of the probability according to the adjusted values is therefore less than the maximum value by

$$(5) \qquad \log Q'-\log Q=\sum\left[(E_xq_x+h_x)\log\left(1+\frac{h_x}{E_xq_x}\right)+\{E_x(1-q_x)-h_x\}\times\log\left\{1-\frac{h_x}{E_x(1-q_x)}\right\}\right]$$

$$=M\sum\left[\left\{h_x+\tfrac{1}{2}\frac{h_x^2}{E_xq_x}-\tfrac{1}{6}\frac{h_x^3}{E_x^2q_x^2}+\ldots\right\}+\left\{-h_x+\tfrac{1}{2}\frac{h_x^2}{E_x(1-q_x)}\right.\right.$$

$$\left.\left.+\tfrac{1}{6}\frac{h_x^3}{E_x^2(1-q_x)^2}+\ldots\right\}\right]$$

$$=M\sum\left\{\frac{h_x^2}{2E_xq_x(1-q_x)}-\frac{(1-2q_x)h_x^3}{6E_x^2q_x^2(1-q_x)^2}+\ldots\right\}.$$

The mean value of this is

(6) $$M\sum\left\{\tfrac{1}{2}-\frac{(1-2q_x)^2}{E_xq_x(1-q_x)}+\ldots\right\}.$$

When E_x is large the second and subsequent terms in the bracket become negligible compared with the first and the expected value of $\log Q'-\log Q$ is accordingly approximately $\frac{1}{2}Mn$ where n is the number of different values of x at which we have observations and M is $\log_{10}e$. The adjusted series should give a value of $\log Q'-\log Q$ which is not materially above the expected value thus arrived at.

If it is found impossible to determine any algebraic function which will satisfy this test it is necessary to make some assumption as to the relative *a priori* probabilities of different sets of values. The assumption which it is proposed to adopt may be expressed as follows:
Let

$$\lambda_x=\log\frac{q_x}{1-q_x},\ \text{and}\ y=\Sigma(\Delta^3\lambda_x)^2$$

and let P denote the *a priori* probability of the set of values.

Then we assume

(7) $$\log P=f(y)$$

where $f'(y)$ is negative.

The function λ_x is used instead of q_x in this statement because, while q_x is limited to values between 0 and 1, λ_x may have any value from $-\infty$ to $+\infty$. We may therefore assume any real variations whatever in the values of λ_x without introducing impossible values of q_x.

On these assumptions PQ is the compound probability that the observed facts arose as a chance variation from the assumed set of probabilities and we have

(8) $$\log PQ=\log P+\log Q=f(y)+\Sigma\{\theta_x\log q_x+(E_x-\theta_x)\log(1-q_x)\}.$$

The condition that the value of this should be stationary for variations of λ_x is

$$f'(y)\frac{dy}{d\lambda_x}+M\left\{\frac{\theta_x}{q_x}-\frac{E_x-\theta_x}{1-q_x}\right\}\frac{dq_x}{d\lambda_x}=0. \tag{9}$$

But

$$\frac{dy}{d\lambda_x}=2\Delta^3\lambda_{x-3}-6\Delta^3\lambda_{x-2}+6\Delta^3\lambda_{x-1}-2\Delta^3\lambda_x=-2\Delta^6\lambda_{x-3}$$

and

$$\frac{d\lambda_x}{dq_x}=M\left(\frac{1}{q_x}+\frac{1}{1-q_x}\right)=\frac{M}{q_x(1-q_x)},$$

therefore

$$\frac{dq_x}{d\lambda_x}=\frac{q_x(1-q_x)}{M}.$$

Hence

$$-2f'(y)\Delta^6\lambda_{x-3}+(\theta_x-E_xq_x)=-2f'(y)\Delta^6\lambda_{x-3}+h_x=0. \tag{10}$$

This must hold for all values of x. Therefore since $f'(y)$ is independent of x the values of $\Delta^6\lambda_{x-3}$ must be proportional to h_x or to the departure of the actual from the expected number of realizations of the probability. For those values of x for which there are no observed facts so that $\theta_x-E_xq_x=0$ we must have $\Delta^6\lambda_{x-3}=0$. It is usually the case that the observed facts correspond to a finite range of values of x. It will be noted that the necessary modification of the equations of condition near the limiting values of x is equivalent to assuming that the series of values of λ_x is extended in both directions in such fashion that all third differences, involving any value of λ_x outside of the range, vanish.

We have thus a set of equations of the form

$$\Delta^6\lambda_{x-3}+k(h_x)=0 \tag{11}$$

where k is some, as yet undetermined, quantity.

Also

$$\frac{d}{d\lambda_x}\left\{\Delta^6\lambda_{x-3}+k(\theta_x-E_xq_x)\right\}=-20-kE_x\frac{dq_x}{d\lambda_x}=-\left\{20+\frac{k}{M}E_xq_x(1-q_x)\right\}.$$

Considered therefore as an isolated adjustment the approximate correction to λ_x would be

$$\frac{\Delta^6\lambda_{x-3}+kh_x}{20+\frac{k}{M}E_xq_x(1-q_x)},$$

or

$$\frac{\Delta^6\lambda_{x-3}+kh_x}{20+kw_x}$$

if we write w_x for $\dfrac{E_xq_x(1-q_x)}{M}$.

The value of $\Delta^6\lambda_{x-3}$ will, however, be affected not only by the change in λ_x but also by the changes in the adjacent terms from λ_{x-3} to λ_{x+3} inclusive, and these effects may be cumulative or divergent. In fact if we write ρ_x for the correction to be applied to λ_x we have the equation

$$(12) \qquad \Delta^6\lambda_{x-3}+kh_x+\Delta^6\rho_{x-3}-kw_x\rho_x=0.$$

We cannot, however, determine the exact general solution of this difference equation, either in its exact form (11) or in the approximate form (12). Methods of successive approximation are therefore necessary. The following is suggested as fairly rapidly convergent and as possessing advantages in connection with the determination of the as yet unknown value of k.

An approximate value of ρ_x is supposed derived from the equation

$$(13) \qquad \rho_x=l(\Delta^6\lambda_{x-3}+kh_x).$$

A suitable value of l is determined by considering the particular cases where $k=0$ in which case the adjusted value of λ_x will be $\lambda_x+l\Delta^6\lambda_{x-3}$ and the adjusted value of $\Delta^3\lambda_x$ will be $\Delta^3\lambda_x+l\Delta^9\lambda_{x-3}$. If we consider that the individual values of λ_x are subject to independent fluctuations with the same mean square of error, the value of l for which the mean square of the error of the adjusted value of λ_x is a minimum is 5/231 and that for which the mean square of the error of the adjusted value of $\Delta^3\lambda_x$ is a minimum is 21/1105. For ease of calculation we may take the intermediate value 1/50 for l and the equation takes the form

$$(14) \qquad \rho_x=.02\ (\Delta^6\lambda_{x-3}+kh_x).$$

In order now to determine an appropriate value for k we note that

$$\frac{d\log Q}{d\lambda_x}=\theta_x-E_xq_x=h_x.$$

Hence the approximate change made in log Q by the above adjustment is

$$(15) \qquad \delta\log Q=.02(\Sigma h_x\Delta^6\lambda_{x-3}+k\Sigma h_x^2).$$

The value of log Q derived from the adjusted values of λ_x will accordingly be

$$\log Q+.02\{\Sigma h_x\Delta^6\lambda_{x-3}+k\Sigma h_x^2\}.$$

Our previous investigation has shown that the expected or mean value of log Q is approximately $\log Q'-\frac{1}{2}Mn$. If we, therefore, determine k so that the adjusted value of log Q shall have this value we have

$$(16) \qquad k\Sigma h_x^2=50\ (\log Q'-\log Q-\tfrac{1}{2}Mn)-\Sigma h_x\Delta^6\lambda_{x-3}.$$

From the nature of the argument, however, negative values of k are not admissible so that zero must be used for k where this formula would give a negative value.

This investigation presupposes a preliminary knowledge of approximate values for the graduated series. These approximate values may be derived by the usual graphic, interpolation or summation methods. Equations (14) and (16) are then applied to derive a set of corrections which applied to the first approximation gives a second which may be again corrected in the same way if desired.

It is interesting to note that, on account of the values of $\Delta^6\lambda_{x-3}$ at the ends of the series being derived on the assumption that third differences vanish if they include any term beyond the actual range of observations, the first, second and third sums of the values of $\Delta^6\lambda_{x-3}$ all vanish. If, therefore, a series could be arrived at which would exactly satisfy equation (11), the first, second and third sums of the values of h_x, the departure of the actual from the expected number of realization,s would all vanish. The total number and the first and second moments of the expected would therefore be the same as those of the actual.

Suppose now that we write equation (12) in the form

$$kw_x\rho_x - \Delta^6\rho_{x-3} = kh_x + \Delta^6\lambda_{x-3}$$

and add to each side $(m - kw_x)\rho_x$, where m is a convenient constant of the same order of magnitude as the average value of kw_x. Then

$$m\rho_x - \Delta^6\rho_{x-3} = (m - kw_x)\rho_x + kh_x + \Delta^6\lambda_{x-3}. \tag{17}$$

Since the factor $(m - kw_x)$ is small we may substitute on the right side the approximate value of ρ_x from equation (14) and we have

$$m\rho_x - \Delta^6\rho_{x-3} = \left(1 + \frac{m - kw_x}{50}\right)(kh_x + \Delta^6\lambda_{x-3}). \tag{18}$$

The practical solution of difference equations of this form is discussed by Professor E. T. Whittaker in his papers: *On a new method of graduation*, Proc. Edin. Math. Soc., XLI, p. 63 and *On the Theory of Graduation*, Proc. Royal Soc. Edin., XLIV, p. 77, and by the author of this paper in a paper entitled *A new method of graduation*, Trans. Act. Soc. of Am., XXV, p. 29.

It will be seen from the foregoing description that the method suggested is essentially one of approximate corrections applied to a preliminary adjusted table which is supposed to require improvement. It affords a method of determining the necesary corrections by a mathematical process. It may, therefore, be considered as a reduction to mathematical form of the graphic process while at the same time retaining its applicability to material too scanty for a summation graduation.

The final result is to approach by successive approximations to the series which will have the smallest value of $\sum(\Delta^3\lambda_x)^2$ consistent with the requirement that the value of $\log Q' - \log Q$ should not exceed $\frac{1}{2}Mn$.

Note.—Since presenting the above paper the author has, in a paper entitled *Further Remarks on Graduation*, Trans. Act. Soc. of Am., XXVI, p. 52, proposed a practical direct solution of difference equations of a form equivalent to (12). If the methods of that paper are used, the approximation described in equations (17) and (18) would be avoided.

INTERPOLATION WITH LEAST MEAN SQUARE OF ERROR

By Dr. W. F. Sheppard,
London, England.

1. What is the best interpolation-formula?

The answer to this question must depend on what we mean by "best," and also, of course, on the kind of formula we are considering. As regards this latter point, we are to confine our attention to polynomial interpolation from constant-difference data; *i.e.*, we are to suppose that the data are a series of values of z corresponding to a series of values of x which proceed by a constant difference h, and we are to assume that the z which we require to find can (to the degree of accuracy to which we are working) be represented by a polynomial function of x of degree m, whose $m+1$ coefficients can be determined by equating $m+1$ consecutive z's to the same polynomial functions of the corresponding x's. Our problem is then to determine what are the best z's to use for our purpose.

It will be observed that the problem has been stated as one of expressing the polynomial in terms of a series of z's. In actual practice it is more usual to seek a formula which involves one z and a series of differences taken by one or other of a number of different routes. We have, for instance, the advancing-difference formula, the receding-difference formula, the central-difference formula or formulae, and so on. But these are Newtonian forms, all of which can be reduced to the one Lagrangian form. Suppose, for instance, that the data are $\ldots, z_0, z_1, \ldots, z_m, \ldots$, corresponding to values $\ldots, x_0, x_1, \ldots, x_m, \ldots$ of x, and that we require the value z_θ which corresponds to the value $x_\theta \equiv x_0+\theta h$ of x; and let us, for simplicity, consider only the case of m odd, so that there will be only one central-difference formula. Then the advancing-difference formula

$$z_0+\theta\Delta z_0+\frac{\theta(\theta-1)}{2!}\Delta^2 z_0+\ldots+\frac{\theta(\theta-1)\ldots(\theta-m+1)}{m!}\Delta^m z_0$$

can be expressed in the Lagrangian form

$$L\{z_\theta;\, z_0,\ldots,z_m\} \equiv \frac{\theta(\theta-1)\ldots(\theta-m)}{m!}\left\{C_0\frac{z_m}{\theta-m}-C_1\frac{z_{m-1}}{\theta-m+1}+\ldots+(-)^m C_m\frac{z_0}{\theta}\right\},$$

where C_r denotes $m!/\{r!(m-r)!\}$; the receding-difference formula

$$z_1+(\theta-1)\Delta z_0+\frac{(\theta-1)\theta}{2!}\Delta^2 z_{-1}+\ldots+\frac{(\theta-1)\theta\ldots(\theta+m-2)}{m!}\Delta^m z_{-m+1}$$

is

$$L\{z_\theta;\ z_{-m+1},\ldots,z_1\} \equiv \frac{(\theta+m-1)(\theta+m-2)\ldots(\theta-1)}{m!}$$
$$\times\left\{C_0\frac{z_1}{\theta-1}-C_1\frac{z_0}{\theta}+\ldots+(-)^m C_m\frac{z_{-m+1}}{\theta+m-1}\right\};$$

and, replacing m by $2n+1$, the central-difference formula which involves differences up to $\delta^{2n+1}z_{\frac{1}{2}}\equiv\Delta^{2n+1}z_{-n}$ is $L\{z_\theta;\ z_{-n},\ \ldots,z_{n+1}\}$. It is therefore simpler to deal only with the Lagrangian form. Thus the formula which involves z_{p-m}, $z_{p-m+1},\ldots,z_p$, namely

$$(1.\mathrm{A})\qquad z_\theta=L\{z_\theta;\ z_{p-m},\ldots,z_p\}\equiv\frac{(\theta-p)(\theta-p+1)\ldots(\theta-p+m)}{m!}$$
$$\times\left\{C_0\frac{z_p}{\theta-p}-C_1\frac{z_{p-1}}{\theta-p+1}+\ldots+(-)^m C_m\frac{z_{p-m}}{\theta-p+m}\right\},$$

will be the advancing-difference formula for interpolation betweee z_{p-m} and z_{p-m+1}, the receding-difference formula for interpolation between z_{p-1} and z_p, and (for $m=2n+1$) the central-difference formula for interpolation between z_{p-n-1} and z_{p-n}.

When we are interpolating from a mathematical table, it is usually desirable that the result should be as accurate as possible; and, if other things—*e.g.*, ease of calculation and number of terms required—are equal, the best formula will be that which gives the smallest limit of error. Ordinarily, for this class of cases, the central-difference formula is to be preferred, on the ground that the calculations are easier, that we can sometimes do with fewer terms, and that, even if the number of terms to be used is the same, the limit of error is smaller. But there is another class of cases—the cases of observed quantities subject to error—in which the answer to the question as to the "best" formula involves different considerations. Each of the z's which we use is subject to an error which is practically unlimited, but small values of the error are more frequent than large ones; and what we have to take into account is not the limit of the error in the interpolated value but the relative frequencies of different values of this error. For these cases we may reasonably consider that, in ordinary circumstances, the best formula is that which gives the least mean square of error.

2. The problem may therefore be stated as follows. We require the best value of z_θ, where $0<\theta<1$. For z_θ we use the Lagrangian expression involving one of a number of possible sets of $m+1$ consecutive z's, such as

$$\left.\begin{array}{l}\vdots\\ \vdots\\ z_{-1},\ z_0,\ z_1,\ \ldots,\ z_{m-1}\\ z_0,\ \ z_1,\ z_2,\ \ldots,\ z_m\\ z_1,\ \ z_2,\ z_3,\ \ldots,\ z_{m+1}\\ \vdots\\ \vdots\end{array}\right\}.$$

If the first z in the set we use is denoted by z_{p-m}, the Lagrangian expression is as given in (1.A); and we have to find the value of p for which the m.s.e.—*i.e.*, the mean square of error—in the resulting value of z_θ is least.

If the error in z_r is a_r, the error in z_θ is to be found by replacing z by a throughout (1.A). The mean square of this error will involve the mean squares and mean products of the a's. For a general result, we can only consider the simple case in which the a's are statistically independent and all have the same mean square, which we will denote by k^2. If the m.s.e. of $L\{z_\theta; z_{p-m}, \ldots, z_p\}$ is denoted by $K^2\{z_\theta; z_{p-m}, \ldots, z_p\}$, it will be seen that, m being fixed, the ratio of this m.s.e. to k^2 is a function of $\theta-p$. We may write this in the form

$$K^2\{z_\theta;\ z_{p-m}, \ldots, z_p\} = F(\theta-p), \tag{2. 1}$$

where

$$F(\theta-p) \equiv \left\{\frac{(\theta-p)(\theta-p+1)\ldots(\theta-p+m)}{m!}\right\}^2 \times \left\{\frac{C_0^2}{(\theta-p)^2} + \frac{C_1^2}{(\theta-p+1)^2} + \ldots + \frac{C_m^2}{(\theta-p+m)^2}\right\} k^2. \tag{2. A}$$

We want to find the value of p for which $F(\theta-p)$ is least.

3. Let us begin by comparing the m.s.e. as given by (2. 1) with that due to using the formula next before it on the list (at beginning of § 2), namely the formula which involves $z_{p-m-1}, z_{p-m}, \ldots z_{p-1}$.

(*a*) To see which of these mean squares of error is the greater, we subtract the latter from the former, and see whether the difference is positive or negative. The difference is $F((\theta-p)-F(\theta-p+1)$. It will be seen that the last term of $F(\theta-p+1)$ is equal to the first term of $F(\theta-p)$, and that, if we write

$$Q \equiv \left\{\frac{(\theta-p+1)(\theta-p+2)\ldots(\theta-p+m)}{m!}\right\}^2 k^2, \tag{3. A}$$

so that Q is positive, then

$$F(\theta-p) - F(\theta-p+1) = Q . f(\theta-p), \tag{3. 1}$$

where

$$f(\theta-p) \equiv (\theta-p)^2 \left\{\frac{C_1^2}{(\theta-p+1)^2} + \frac{C_2^2}{(\theta-p+2)^2} + \ldots + \frac{C_m^2}{(\theta-p+m)^2}\right\}$$
$$-(\theta-p+m+1)^2 \left\{\frac{C_0^2}{(\theta-p+1)^2} + \frac{C_1^2}{(\theta-p+2)^2} + \ldots + \frac{C_{m-1}^2}{(\theta-p+m)^2}\right\}.$$

We have to investigate the sign of $f(\theta-p)$.

(*b*) We can simplify $f(\theta-p)$ by grouping the terms.

(i) If m is odd and $=2n+1$, there is a central term in each bracket, and these central terms taken together are equal to

$$-\frac{4(n+1)C_nC_{n+1}}{\theta-p+n+1} = -\frac{2(m+1)C_nC_{n+1}}{\theta-p+n+1}.$$

(ii) Taking together the rth term from the beginning and the rth term from the end in each bracket, it will be found that the four terms together are equal to

$$-2(m+1)C_{r-1}C_r\left(\frac{1}{\theta-p+r}+\frac{1}{\theta-p+m-r+1}\right).$$

(iii) Hence, on the whole, whether m is odd or even,

$$f(\theta-p)=-2(m+1)\left\{\frac{C_0C_1}{\theta-p+1}+\frac{C_1C_2}{\theta-p+2}+\ldots+\frac{C_{m-1}C_m}{\theta-p+m}\right\}. \tag{3.2}$$

(c) For convenience, we can divide the formulae which we are comparing into three overlapping groups: (i) the group which begins with the advancing-difference formula, the next being the formula which uses $z_1, \ldots, z_{m+1}$, and so on, (ii) the corresponding group on the other side, which ends with the receding-difference formula (z_{-m+1} to z_1), and (iii) the intermediate group, from the receding-difference to the advancing-difference (both included), each of the formulae in this group using both z_0 and z_1.

(i) For the first group, the last z's are successively $z_m, z_{m+1}, z_{m+2}, \ldots$, so that $p-1$ in (a) is never less than m. Since $0<\theta<1$, it follows that $\theta-p+m$, $\theta-p+m-1, \ldots, \theta-p+1$ are all negative; and therefore, by (3.2), $f(\theta-p)$ is always positive. In other words, the formulae, beginning with the advancing-difference formula, are successively worse and worse.

(ii) For the second group, taking the formulae in the reverse order, the last z's are successively $z_1, z_0, z_{-1}, \ldots$. Hence the greatest value of p is 1; and therefore $\theta-p+1$, $\theta-p+2, \ldots, \theta-p+m$ are all positive. It follows from (3.2) that $f(\theta-p)$ is always negative, so that the formulae, beginning with the receding-difference formula and working backwards, are successively worse and worse.

(iii) It remains to consider the intermediate group, in which there are m sets of z's to be compared, namely

$$\begin{array}{l} z_{-m+1}, z_{-m+2}, \ldots, z_0, \quad z_1 \text{ (receding-difference)},\\ \quad\vdots \\ \quad\vdots \\ z_{p-m-1}, z_{p-m}, \ldots, z_{p-2}, z_{p-1}\ (F),\\ z_{p-m}, z_{p-m+1}, \ldots, z_{p-1}, z_p\ (F'),\\ \quad\vdots \\ \quad\vdots \\ z_0, z_1, \ldots\ldots\ldots, z_{m-1}, z_m \text{(advancing-difference)}.\end{array}$$

If we write

$$\psi(\theta-p)\equiv\frac{C_0C_1}{\theta-p+1}+\frac{C_1C_2}{\theta-p+2}+\ldots+\frac{C_{m-1}C_m}{\theta-p+m}, \tag{3.B}$$

then (F') will give a better or a worse result than (F) according as $\psi(\theta-p)$ is positive or negative.

4. (i) Suppose that θ increases continuously from being just >0 to being just <1. Then $\psi(\theta-p)$ decreases continuously. Also, when $\theta\rightarrow 0$, $\psi(\theta-p)\rightarrow+\infty$; and, when $\theta\rightarrow 1$, $\psi(\theta-p)\rightarrow-\infty$. There is therefore a certain value of θ, and only one, for which $\psi(\theta-p)$ is $=0$. If we denote this value by θ_{m-p+1}, then (F') will give a better or a worse result than (F) according as θ is less or greater than θ_{m-p+1}.

(ii) Now let us in the same way compare (F') with the next formula (F''), which involves $z_{p-m+1}, z_{p-m+2}, \ldots, z_{p+1}$. There will similarly be a value θ_{m-p} of θ such that (F'') will give a better or a worse result than (F') according as θ is less or greater than θ_{m-p}; and this will be the value for which $\psi(\theta-p-1)=0$, where

$$\psi(\theta-p-1)\equiv\frac{C_0C_1}{\theta-p}+\frac{C_1C_2}{\theta-p+1}+\ldots+\frac{C_{m-1}C_m}{\theta-p+m-1}.$$

It may be shewn that θ_{m-p+1} is greater than θ_{m-p}. If θ is less than θ_{m-p} and therefore less than θ_{m-p+1}, (F'') will give a better result than (F'), and (F') a better result than (F); if θ is between θ_{m-p} and θ_{m-p+1}, (F') will give a better result than either (F) or (F''); if θ is greater than θ_{m-p+1} (F) will give a better result than (F'), and (F') a better result than (F'').

(iii) Dealing with all the sets of the group in this way, we see that there is a series of values $\theta_1, \theta_2, \ldots, \theta_{m-1}$ dividing the interval from 0 to 1 into m parts such that if θ falls in the first part the best formula is the advancing-difference formula, if it falls in the second part the best formula is the one next before the advancing-difference formula, and so on. Thus each of the m formulae involving the m sets of z's in § 3 (*c*) (iii) has a particular bit of the interval $\theta=0$ to 1 for which it is the best formula.

5. The values $\theta_1, \theta_2, \ldots, \theta_{m-1}$ have been found as the values of θ for which

$$\frac{C_0C_1}{\theta}+\frac{C_1C_2}{\theta-1}+\ldots.+\frac{C_{m-1}C_m}{\theta-m+1}=0,$$

$$\frac{C_0C_1}{\theta+1}+\frac{C_1C_2}{\theta}+\ldots.+\frac{C_{m-1}C_m}{\theta-m+2}=0,$$

$$\vdots$$

$$\frac{C_0C_1}{\theta+m-2}+\frac{C_1C_2}{\theta+m-3}+\ldots+\frac{C_{m-1}C_m}{\theta-1}=0;$$

but they can be expressed in a slightly different way.

Let us write

$$\chi(\Theta)\equiv\frac{(\Theta-1)(\Theta-2)\ldots(\Theta-m)}{m!}\left\{\frac{C_0C_1}{\Theta-1}+\frac{C_1C_2}{\Theta-2}+\ldots+\frac{C_{m-1}C_m}{\Theta-m}\right\},$$

so that $\chi(\Theta)$ is a polynomial of degree $m-1$ in Θ. It is clear that $\chi(\Theta)$ is positive if $\Theta>m$, and is of sign $(-)^{m-1}$ if $\Theta<1$. Also if we write $\Theta=m, m-1, m-2, \ldots, 1$, $\chi(\Theta)$ is alternately positive and negative. The $m-1$ roots of the equation $\chi(\Theta)=0$ therefore lie respectively in the $m-1$ intervals 1 to 2, 2 to 3, $\ldots, m-1$ to m. If we denote these roots by $\Theta_1, \Theta_2, \ldots, \Theta_{m-1}$, the relation between these roots and $\theta_1, \theta_2, \ldots, \theta_{m-1}$ is

$$\theta_1=\Theta_1-1,\ \theta_2=\Theta_2-2,\ \ldots,\ \theta_{m-1}=\Theta_{m-1}-(m-1).$$

The statement that $\theta_1, \theta_2, \ldots, \theta_{m-1}$ are in ascending order of magnitude is equivalent to the statement that $\Theta_1, \Theta_2, \ldots, \Theta_{m-1}$ are successively further advanced in their respective intervals.

6. The point of special interest is the way the suitable set of z's changes as θ changes. Suppose we arrange the z's horizontally at equal intervals, the positive direction being from left to right; and let the increase of θ be represented by the motion of a point P from left to right. Let the successive differences of the z's, up to the mth, be arranged above the z's in the usual relative position, so that Δz_r comes above the middle of the interval between z_r and z_{r+1}. Then any set of $m+1$ z's, and their successive differences from the m first differences to the single mth difference, will lie inside a triangle, which we may call the *tabular triangle*. For the m sets of z's which we have been considering, the tabular triangle has m positions, its vertex for the first set mentioned in § 3 (c) (iii) being above a point midway between z_{-m+1} and z_1, and for the last set being above a point midway between z_0 and z_m. Now if we only knew that, as stated at the end of § 4, each of the m formulae corresponds to a bit of the interval $\theta=0$ to 1, we might naturally suppose that the order of the m positions of the tabular triangle is the same as the order of the bits of the interval, so that as θ increases from 0 to 1 the triangle moves from the position of the receding-difference formula to that of the advancing-difference formula, and then, as θ passes from the interval $0-1$ to the interval $1-2$, comes back through $m-2$ intervals to the new receding-difference position. But the fact is exactly the opposite. When θ is very small, so that the point P is at the extreme left of the interval $0-1$, the tabular triangle is to the extreme right; and as P moves from left to right up to $\theta=1$ the triangle moves (by jumps) from right to left. As θ is approaching 1 the triangle is to the extreme left, covering $z_{-m+1}, \ldots, z_1$. When θ passes 1 the triangle is thrown over, through m instead of $m-2$ intervals, to the extreme right, so as to cover $z_1, \ldots, z_{m+1}$, and then works back again until it reaches $z_{-m+2}, \ldots, z_2$; and so on.

The effect of this is that the transition on passing from one interval to the next is as violent as possible. For the interval $0-1$ the z's which might be used are $z_{-m+1}, \ldots, z_m$. But z_{-m+1} is not used until the last bit of the interval; and when we pass from $0-1$ to $1-2$ the tabular triangle discards it and sweeps over at once to take in z_{m+1}. It speeds the parting, welcomes the coming guest.

The point is of interest in relation to the not uncommon practice, by the use of osculatory interpolation, of making the break, due to passing from one interval to the next, less violent than it is even when the same kind of formula—*e.g.*, the central-difference formula—is being used tbroughout. This practice succeeds in producing greater smoothness, but only at the expense of accuracy.

We can understand this by considering the simple case of interpolation by first difference. We can, as before, denote the errors in $z_0, z_1, z_2, \ldots$ by

α_0, α_1, α_2, ... Suppose that α_0, α_1, α_2 are all positive, and of approximately the same magnitude. Then the interpolated values of z between z_0 and z_2 will all have this same error. Now suppose that α_3 is of about the same absolute magnitude as the preceding errors, but is negative. Then the effect of using 1st-difference interpolation from z_2 to z_3 is that the negative error α_3 begins at once, to the best of its ability, to counteract the positive error α_2. Osculatory interpolation prevents this. With 1st-difference interpolation, the graph of z is an irregular line, approximately straight from z_0 to z_2 and then sloping suddenly towards z_3. Osculatory interpolation substitutes a curved line, allowing α_1 to continue its effect after its proper time. The resulting series of values of z is smoother, but the individual values are obviously less accurate. To put it briefly, increase of smoothness and increase of accuracy, when we retain the observed values. . ., z_0, z_1, z_2, . . . , are antagonistic.

7. We come now to the practical question: How far are these results to affect our choice of a formula for interpolation? We have found that different formulae are better for different parts of the interval; but is it really necessary that we should use these different formulae? We may conveniently divide the question under three heads:

(1) What are the respective magnitudes of the different parts of the interval?

(2) How much better is one formula than another, goodness being measured by smallness of m.s.e.?

(3) Would any serious inaccuracy be likely to result from our using one formula, such as the central-difference formula, throughout?

8. The following table (Table I) gives the division of the interval $0-1$ into bits, for values of m from 1 to 6, which is about as far as we need go.

TABLE I

Division of Interval $0-1$ According to Suitability of Formula

$m\equiv$degree of polynomial	$m+1$=number of z's involved in formula	No. of possible formulae	Determining values for division of the interval
1	2	1	0.0, 1.0
2	3	2	0.0, 0.5, 1.0
3	4	3	0.0, 0.2254, 0.7746, 1.0
4	5	4	0.0, 0.0985, 0.5, 0.9015, 1.0
5	6	5	0.0, 0.0412, 0.3301, 0.6699, 0.9588, 1.0
6	7	6	0.0, 0.0163, 0.1636, 0.5, 0.8364, 0.9837, 1.0

It will be seen that, when m is large, the portions of the interval in which we should in any case use a central-difference formula are considerable, while those suitable for advancing or receding differences are very small.

9. We have so far only compared different formulae by seeing whether one formula gives a greater or a less m.s.e. than another. In order to see how much

better or worse one formula is than another, we have to examine the magnitude of each m.s.e., *i.e.*, we have to examine the magnitude of

$$F(\theta-p)\equiv K^2\{z_\theta;\ z_{p-m},\ldots,z_p\}$$

throughout the interval $\theta=0$ to 1 for each of the relevant values of p. We could do this by taking each value of p separately. But as these separate variations of $F(\theta-p)$ are merely successive parts of the variation of the m.s.e. due to a single Lagrangian formula throughout the range bounded by its extreme z's, it is simpler to consider this variation first, as a whole, and then to compare the different parts of it. We therefore take the formula which includes z_0 and z_m, and see how the m.s.e. of z_θ varies as θ increases from 0 to m.

Let us denote $F(\theta-m)/k^2$ by R, so that $R=1$ for $\theta=0, 1, 2, \ldots, m$. The graph of R for $\theta=0$ to m is obviously symmetrical about its middle ordinate, and we need only consider the first half of the graph, corresponding to values of θ from 0 to $\frac{1}{2}m$.

For values of m from 2 to 5, the graph can be roughly traced from Table II, which gives (for the first half of each graph) the values of R and of θ when R is a minimum or a maximum, and also the values of θ for which $R=1$. The entries for the central ordinate are in heavier type. If the graph for any particular m is drawn, and the separate pieces for the $m+1$ intervals are then put together in one interval, their points of intersection will mark the points at which, as shown in Table I, we should pass from one formula to another in order always to have the least possibe m.s.e.

TABLE II

MATERIALS FOR GRAPH OF m.s.e.

		$R=1$	R min.	$R=1$	R max.	$R=1$	R min.	$R=1$	R max.	$R=1$	R min.	$R=1$
$m=2$	θ									0.0	0.2929	**1.0**
	R									1	0.6250	**1**
$m=3$	θ					0.0	0.1602	0.4276	0.7010	1.0	**1.5**	2.0
	R					1	0.7481	1	1.1779	1	**0.6406**	1
$m=4$	θ					0.0	0.0808	0.1873	0.5482	1.0	1.3494	**2.0**
	R					1	0.8488	1	1.6325	1	0.6963	**1**
$m=5$	θ	0.0	0.0378	0.0814	0.4584	1.0	1.2326	1.5962	1.7952	2.0	**2.5**	3.0
	R	1	0.9189	1	2.7526	1	0.7623	1	1.0746	1	**0.7060**	1

To see how the "best" formula compares with the central-difference formula, if the latter were used throughout the whole of each interval, it will be sufficient to take the values of θ for which the non-central formula produces its greatest reduction, *i.e.*, the values for which, as shown in the above table, R is a minimum. The comparison is made in Table III. Suppose, for instance, that $m=5$. Then the greatest reducing effect of the advancing-difference formula is when $\theta=.0378$, R being then 0.9189; if we used the central-difference formula the value of R, as found by taking $\theta=2.0378$ in Table II, would be 0.9734. Similarly the greatest effect of the next formula corresponds to $\theta=1.2326$ in Table II, R being then 0.7623; and the effect of using the central-difference formula at this point would be shown by taking $\theta = 2.2326$, which would give $R = 0.8189$.

TABLE III

COMPARISON OF CENTRAL-DIFFERENCE FORMULA WITH OTHER FORMULAE AT THEIR BEST

m	Portion of interval	Best formula		Corresponding values for central-difference formula		Ratio of values of R
		θ	R	θ	R	
3	Outer	0.1602	0.7481	1.1602	0.8349	0.8960
4	Outer	0.0808	0.8488	1.0808	0.8774	0.9674
5	Outer	0.0378	0.9189	2.0378	0.9734	0.9440
	Next	1.2326	0.7623	2.2326	0.8189	0.9310

It will be seen that, even at their best, the results given by the non-central formulae are not so very much better than those given by the central formula.

10. On the whole, therefore, the central-difference method comes out fairly well in the competition. Its advantages are that

(1) the formulae are more convenient, especially for sub-tabulation, *i.e.*, for interpolation by an interval which is a sub-multiple of the original interval;

(2) the coefficients being smaller, it is sometimes possible to do with a lower order of differences;

(3) the limit of error is (speaking generally) less;

(4) the m.s.e. is less, throughout the whole interval, than that of the tabulated values; and

(5) although this m.s.e. is, in some portions of the interval, greater than that which would be due to some other formula, it is not very much greater.

11. There is still one point which remains to be considered. How are we to meet the difficulty that, if m is even, say $=2n$, there is one central-difference formula for the first half of the interval and another for the second half? Are we to use the two formulae? And what formula are we to use for the mid-point of the interval?

It is obviously best that we should have a single formula for the whole interval. Three courses are open:

(1) We may use the central-difference formula for $m' = 2n+1$, so as to go up to differences of order $2n+1$.

(2) We may take the mean of the two central-difference formulae of order $2n$. The resulting formula is the same as if we were to express the formula in (1) in terms of $\mu z_{\frac{1}{2}}$, $\delta z_{\frac{1}{2}}$, $\mu\delta^2 z_{\frac{1}{2}}$, $\ldots$, $\delta^{2n+1} z_{\frac{1}{2}}$, and then omit the term in $\delta^{2n+1} z_{\frac{1}{2}}$; we can call it the *curtailed formula*, the formula of (1) being the *complete formula*.

(3) We may use an arbitrary formula obtained by adding to the curtailed formula some term in $\delta^{2n+1} z_{\frac{1}{2}}$ other than that which converts it into the complete formula.

It is not necessary to go into this question in detail. It will be sufficient to state the following results:

(i) The curtailed formula gives a smaller mean square of error than the complete formula; except at the mid-point, where the two formulae are identical.

(ii) The coefficient of $(\theta-\frac{1}{2})\delta^{2n+1}z_{\frac{1}{2}}$, in the term which converts the curtailed formula into the complete formula, is of sign $(-)^n$; and the addition, to the curtailed formula, of any other positive multiple of $(-)^n(\theta-\frac{1}{2})\delta^{2n+1}z_{\frac{1}{2}}$ will similarly increase the mean square of error.

(iii) In particular, the term which has to be added to the curtailed formula in order to produce a formula for osculatory interpolation is of this latter kind, and therefore, as we should expect from § 6, osculatory interpolation increases the mean square of error.

12. Our enquiry has been purely mathematical, and has been based on an arbitrary definition of "best". The object has been to see the results of different methods; the actual choice of method is a different matter. There is, indeed, some inconsistency between aiming at accuracy rather than smoothness, and retaining the observed values. If we really desire greater accuracy, the natural procedure is to alter the observed values so as to reduce their m.s.e., and then to interpolate in the usual way from the altered values.

QUADRATURE FORMULAE WHEN ORDINATES ARE NOT EQUIDISTANT

BY PROFESSOR JAMES W. GLOVER,
University of Michigan, Ann Arbor, Michigan, U.S.A.

The formula (2) in this paper, and several others of similar type were first derived by Hardy* and later appeared in King's Text-book†, where numerous applications to actuarial calculations are made. It is found in practice to yield surprisingly close approximations, even though a small number of ordinates are used, and on this account some further consideration of its theoretical basis and a method of generalization may prove acceptable. The general idea is to establish a formula of the type

$$\int_{-n}^{n} u_x dx = a(u_{-n}+u_n)+\sum_{i=1}^{s} a_i\,(u_{-n_i}+u_{n_i}) = a\bar{u}_n+\sum_{i=1}^{s} a_i\bar{u}_n \tag{1}$$

where

$$n_1<n_2<\ldots<n_s<n,$$

to express the value of the area bounded by the curve, $y=u_x$, the axis of x, and the two limiting ordinates u_{-n} and u_n. Then taking the origin at $-n$ the limits become 0, $2n$ (or $17n$), and the formula is applied to successive sections of the area, each standing over a base of length $2n$ (or $17n$), to obtain a general expression for the area in terms of the extreme and certain intermediate ordinates which are not equidistant. In some of these formulae a constant times the ordinate at the origin is employed. In the case $s=2$ the formula found by Hardy and given in the Text-book is:

$$\int_{0}^{17n} u_x dx = n[.60(u_0+u_{17n})+4.79(u_{6n}+u_{11n})+3.11(u_{2n}+u_{15n})]. \tag{2}$$

We proceed to derive this formula.

Assuming the expansion of the function by Maclaurin's theorem

$$u_x=u_0+\frac{n}{\underline{|1}}u_0^{(1)}+\frac{n^2}{\underline{|2}}u_0^{(2)}+\ldots+\frac{n^8}{\underline{|8}}u_0^{(8)}+\ldots$$

it follows that

$$\bar{u}_n=u_{-n}+u_n=2u_0+2\frac{n^2}{\underline{|2}}u_0^{(2)}+2\frac{n^4}{\underline{|4}}u_0^{(4)}+\ldots+2\frac{n^8}{\underline{|8}}u_0^{(8)}+\ldots$$

and similarly for $\bar{u}_{n_1}$ and $\bar{u}_{n_2}$.

*G. F. Hardy, *Formulas for Approximate Summation*, Journal Institute of Actuaries, Volume 24, page 95.

†George King, *Institute of Actuaries Text-book*, Part II, page 483 (Second edition).

Also the area

$$A=\int_{-n}^{n} u_x dx = 2nu_0 + 2\frac{n^3}{\lfloor 3}u_0^{(2)} + 2\frac{n^5}{\lfloor 5}u_0^{(4)} + \ldots + 2\frac{n^9}{\lfloor 9}u_0^{(8)} + \ldots$$

and comparing coefficients of like derivatives the following set of homogeneous equations is obtained:

$$\begin{aligned} a+\ a_1\ &+\ a_2\ -n\ =0\\ a+\alpha\, a_1\ &+\beta\, a_2\ -n/3=0\\ a+\alpha^2 a_1\ &+\beta^2 a_2\ -n/5=0 \qquad (3)\\ a+\alpha^3 a_1\ &+\beta^3 a_2\ -n/7=0\\ a+\alpha^4 a_1\ &+\beta^4 a_2\ -n/9=0 \end{aligned}$$

where $\alpha=n_1^2/n^2$ and $\beta=n_2^2/n^2$. For solutions other than the identical one (0, 0, 0, 0), we must have

$$D_1=\begin{vmatrix} 1 & 1 & 1 & 1\\ 1 & \alpha & \beta & 1/3\\ 1 & \alpha^2 & \beta^2 & 1/5\\ 1 & \alpha^3 & \beta^3 & 1/7 \end{vmatrix}=0 \qquad D_2=\begin{vmatrix} 1 & \alpha & \beta & 1/3\\ 1 & \alpha^2 & \beta^2 & 1/5\\ 1 & \alpha^3 & \beta^3 & 1/7\\ 1 & \alpha^4 & \beta^4 & 1/9 \end{vmatrix}=0$$

from which it follows that

$$(\alpha-1)\ (\beta-1)\ (\alpha-\beta)\left[\frac{\alpha\beta}{1.3}-\frac{\alpha+\beta}{3.5}+\frac{1}{5.7}\right]=0$$

$$\alpha\beta(\alpha-1)\ (\beta-1)\ (\alpha-\beta)\left[\frac{\alpha\beta}{3.5}-\frac{\alpha+\beta}{5.7}+\frac{1}{7.9}\right]=0.$$

All values of α and β which satisfy both determinants are trivial except those common to the two hyperbolas

$$\frac{xy}{1.3}-\frac{x+y}{3.5}+\frac{1}{5.7}=0 \text{ and } \frac{xy}{3.5}-\frac{x+y}{5.7}+\frac{1}{7.9}=0.$$

For example, when $\alpha=1$, the ordinate u_{n_1} coincides with u_n; when $\alpha=\beta$, u_{n_1} and u_{n_2} fall together.

Elimination of y leads to the quadratic

$$21x^2-14x+1=0, \qquad (4)$$

the solution of which gives the two symmetrically situated (with respect to the line $y=x$) points of intersection

$$[(7\pm 2\sqrt{7})/21,\ (7\mp 2\sqrt{7})/21]$$

of the hyperbolas or approximately (.5853096, .0813571) and (.0813571, .5853096). Hence $n_1=.28522n$ and $n_2=.76506n$.

Taking (α, β) at an intersection of the hyperbolas the determinants vanish, but

$$\begin{vmatrix} 1 & 1 & 1\\ 1 & \alpha & \beta\\ 1 & \alpha^2 & \beta^2 \end{vmatrix}=(\alpha-1)\ (\beta-1)\ (\alpha-\beta)\neq 0$$

and accordingly, we may take n at pleasure, and

$$a = \frac{15\alpha\beta-5\ (\alpha+\beta)+3}{15\ (\alpha-1)\ (\beta-1)}\ .\ n = \frac{1}{15}n = .06667n,$$

$$a_1 = \frac{2\ (5\beta-1)}{15\ (\alpha-1)\ (\alpha-\beta)}\ .\ n = \frac{1}{15}[7+\tfrac{1}{2}\sqrt{7}]n = .55486n,$$

$$a_2 = \frac{2\ (5\alpha-1)}{15\ (\beta-1)\ (\beta-\alpha)}\ .\ n = \frac{1}{15}[7-\tfrac{1}{2}\sqrt{7}]n = .37487n.$$

The area obtained is

$$(5)\quad A = \int_{-n}^{n} u_x dx = n[.06667(u_{-n}+u_n)+.55486(u_{-n_1}+u_{n_1})+.37487(u_{-n_2}+u_{n_2})],$$

or taking the origin at the ordinate u_{-n}

$$(6)A\quad = \int_{0}^{2n} u_x dx = n[.06667(u_0+u_{2n})+.55486(u_{n-n_1}+u_{n+n_1})+.37487(u_{n-n_2}+u_{n+n_2}).$$

The area is thus expressed in terms of the two bounding ordinates and two additional pairs of ordinates symmetrically placed between them. The bounding ordinates determine the value of n and this determines the position of the remaining four. If u_x is given in functional form these six ordinates may be computed and then the area calculated.

Hardy had chiefly in view the case where the ordinates are known for certain integral values of the arguments, and in this connection notice that when $n=8.5n$, the values of $n_1=2.42n$ and $n_2=6.50n$, nearly, and this led to taking them arbitrarily as $n_1=2.5$ and $n_2=6.6$. When these values are substituted back in the first three equations of (3) new values of a, a_1, a_2 are found and the area expressed in (6) reduces to Hardy's formula (2). In other words, it arises by a slight displacement of the values of a, a_1, a_2, n_1, n_2. It thus appears that (6) gives the area A based on and true to the first nine terms of the expansion of u_x by Maclaurin's theorem and Hardy's formula (2) gives a close approximation to this value.

This method of approach leads to an easy generalization which will be pointed out briefly for the case $s=3$, that is, when three pairs of symmetrically placed ordinates conforming to certain conditions are given between the extreme ordinates bounding the area. The latter would then take the form:

$$(7)\qquad A = \int_{-n}^{n} u_x dx = a(u_{-n}+u_n)+\sum_{i=1}^{3} a_i(u_{-n_i}+u_{n_i}),$$

and expanding by Maclaurin's theorem to thirteen terms, and comparing coefficients up to and including the twelfth derivative, we have the following set of homogeneous linear equations:

$$
(8)\qquad
\begin{aligned}
&a+\ \ a_1 \ +\ \ a_2 \ +\ \ a_3 \ -n \ \ \ \ =0,\\
&a+\alpha\, a_1 \ +\beta\, a_2 \ +\gamma\, a_3 \ -n/3 \ =0,\\
&a+\alpha^2a_1 \ +\beta^2a_2 \ +\gamma^2a_3 \ -n/5 \ =0,\\
&a+\alpha^3a_1 \ +\beta^3a_2 \ +\gamma^3a_3 \ -n/7 \ =0,\\
&a+\alpha^4a_1 \ +\beta^4a_2 \ +\gamma^4a_3 \ -n/9 \ =0,\\
&a+\alpha^5a_1 \ +\beta^5a_2 \ +\gamma^5a_3 \ -n/11=0,\\
&a+\alpha^6a_1 \ +\beta^6a_2 \ +\gamma^6a_3 \ -n/13=0,
\end{aligned}
$$

where

$$\alpha=n_1^2/n^2,\quad \beta=n_2^2/n^2,\quad \text{and } \gamma=n_3^2/n^2.$$

The determinants whose vanishing is a necessary and sufficient condition for solutions other than the identical one are taken from the matrix:

$$
\left\|\begin{matrix}
1 & 1 & 1 & 1 & 1 & 1 & 1\\
1 & \alpha & \alpha^2 & \alpha^3 & \alpha^4 & \alpha^5 & \alpha^6\\
1 & \beta & \beta^2 & \beta^3 & \beta^4 & \beta^5 & \beta^6\\
1 & \gamma & \gamma^2 & \gamma^3 & \gamma^4 & \gamma^5 & \gamma^6\\
1 & 1/3 & 1/5 & 1/7 & 1/9 & 1/11 & 1/13
\end{matrix}\right\| = 0
$$

and may be expanded as follows:

$$D_1=k\left[-\frac{\alpha\beta\gamma}{1.3}+\frac{\alpha\beta+\beta\gamma+\gamma\alpha}{3.5}-\frac{\alpha+\beta+\gamma}{5.7}+\frac{1}{7.9}\right]=0,$$

$$D_2=k\left[-\frac{\alpha\beta\gamma}{3.5}+\frac{\alpha\beta+\beta\gamma+\gamma\alpha}{5.7}-\frac{\alpha+\beta+\gamma}{7.9}+\frac{1}{9.11}\right]\alpha\beta\gamma=0,$$

$$D_3=k\left[-\frac{\alpha\beta\gamma}{5.7}+\frac{\alpha\beta+\beta\gamma+\gamma\alpha}{7.9}-\frac{\alpha+\beta+\gamma}{9.11}+\frac{1}{11.13}\right]\alpha^2\beta^2\gamma^2=0,$$

where, apart from a constant factor,

$$k=(\alpha-1)\ (\beta-1)\ (\gamma-1)\ (\alpha-\beta)\ (\beta-\gamma)\ (\gamma-\alpha).$$

All the solutions common to D_1, D_2, and D_3, are trivial (involving the coincidence of two or more of the ordinates), except the points of intersection of the three cubic surfaces

$$
(9)\qquad
\begin{aligned}
&\frac{xyz}{1.3}-\frac{xy+yz+zx}{3.5}+\frac{x+y+z}{5.7}=\frac{1}{7.9},\\
&\frac{xyz}{3.5}-\frac{xy+yz+zx}{5.7}+\frac{x+y+z}{7.9}=\frac{1}{9.11},\\
&\frac{xyz}{5.7}-\frac{xy+yz+zx}{7.9}+\frac{x+y+z}{9.11}=\frac{1}{11.13}.
\end{aligned}
$$

These points have abscissae satisfying the cubic equation

$$x^3+px^2+qx+r=0$$

where p, q, r, are determined by the equations

$$
(10)\qquad
\begin{aligned}
&\frac{r}{1.3}+\frac{q}{3.5}+\frac{p}{5.7}+\frac{1}{7.9}=0,\\
&\frac{r}{3.5}+\frac{q}{5.7}+\frac{p}{7.9}+\frac{1}{9.11}=0,\\
&\frac{r}{5.7}+\frac{q}{7.9}+\frac{p}{9.11}+\frac{1}{11.13}=0.
\end{aligned}
$$

The actual cubic in this case is

$$
(11)\qquad 429x^3-495x^2+135x-5=0
$$

and its roots are all real, distinct, and positive, and accordingly lead directly to the numerical coefficients in the expression for the area A. The roots are $\alpha=.0438062$, $\beta=.3501091$, $\gamma=.7599309$, and they furnish the following values of the quantities n_1, n_2, n_3, entering in the quadrature formula:

$$
\begin{aligned}
n_1&=n_1'n=.20930n,\\
n_2&=n_2'n=.59170n,\\
n_3&=n_3'n=.87174n,
\end{aligned}
$$

where primed letters are used to denote the corresponding numerical coefficients of n. It will be convenient to replace the arbitrary n by $\frac{1}{2}Mn$; the formula then becomes

$$
(12)\qquad A=\int_{-\frac{1}{2}Mn}^{\frac{1}{2}Mn}u_x dx=n\left[\frac{Ma'}{2}(u_{-n'n}+u_{n'n})+\sum_{i=1}^{3}\frac{Ma_i'}{2}(u_{-n'_in}+u_{n'_in})\right]
$$

or, taking the origin at the extreme left hand ordinate,

$$
(13)\qquad A=\int_{o}^{Mn}u_x dx=n\left[\frac{Ma'}{2}(u_o+u_{Mn})+\sum_{i=1}^{3}\frac{Ma_i'}{2}(u_{\frac{1}{2}(M-Mn'_i)n}+u_{\frac{1}{2}(M+Mn'_i)n})\right].
$$

This expression for the area agrees with that obtained from the series expansion to thirteen terms. If n is any integer and M is an integer so chosen as to make the Mn'_i integral (nearly) a quadrature formula is obtained in terms of ordinates at integral abscissae by a slight displacement of these values, that is, of the positions of the paired ordinates. This displacement in turn modifies slightly the values of a_1', a_2', and a_3', as determined from the first five equations in (8). The closeness of the approximation to the area depends upon the displacement of the paired ordinates determined by the cubic (11) and this displacement is in turn affected by the choice of the multiplier M. A great variety of formulae may thus be obtained for each value of s. When $n=1$ and M is small, say not more than 10, the formula may be applied to successive sections of the area and the result combined in a single quadrature formula, as in the case of Simpson's formula. On the other hand n may be taken sufficiently great to cover a long range Mn on the abscissa axis, and the area obtained with the use of only eight ordinates, or in general, $2(s+1)$ ordinates.

ON A CLASS OF QUADRATURE FORMULAE*

BY PROFESSOR J. F. STEFFENSEN,
University of Copenhagen, Copenhagen, Denmark.

It is an almost universal practice amongst those who have to deal with mathematics not for its own sake, but on account of its applications, to derive quadrature-formulae (and interpolation-formulae generally) on purely formal lines, that is: on the assumption that the function under consideration is a polynomial, usually of a low degree. Many useful formulae have been found in this way, but as long as their accuracy has only been tested by numerical examples, and their remainder-terms have not been given in a practical form, they cannot be said to rest on a satisfactory foundation.

I propose to-day to deal with a class of quadrature-formulae with which I have occupied myself before†, but without succeeding, in the first instance, in clearing up all the difficulties they present, namely the formulae of Cotes' type. While the remainder-term of the formulae of this class, containing an *odd* number of terms, was easily put into a practical form, the method by which this result was obtained proved inapplicable in the case of an *even* number of terms. We may here proceed as follows:

If we put

$$P(x) = (x^2 - \tfrac{1}{4})\ (x^2 - \tfrac{9}{4}) \dots \left(x^2 - \frac{(2r-1)^2}{4}\right), \tag{1}$$

$$P_v(x) = \frac{P(x)}{x - \dfrac{2v-1}{2}}, \tag{2}$$

where v is one of the numbers $0, \pm 1, \pm 2, \dots \pm(r-1), r$, we obtain by *Lagrange's* interpolation-formula

$$f(x) = \sum_{-r+1}^{r} \frac{P_v(x)}{P_v\left(\dfrac{2v-1}{2}\right)} f\left(\frac{2v-1}{2}\right) + P(x) f\left(x, \pm\tfrac{1}{2}, \dots, \pm\frac{2r-1}{2}\right). \tag{3}$$

In this formula, $f\left(x, \pm\tfrac{1}{2}, \dots, \pm\dfrac{2r-1}{2}\right)$ denotes the divided difference of order $2r$,

*See author's book: *Interpolation* (Baltimore, 1927), §16.

†Transactions of the Fifth Scandinavian Mathematical Congress (Helsingfors, 1922), p. 125. and papers quoted therein, published in Skandinavisk Aktuarietidskrift, 1921, p. 201 and 1922, p.20,

formed with the arguments $x,\ \pm\frac{1}{2},\ \ldots \pm \frac{2r-1}{2}$, and the formula is, so far, an identity.

We now put, denoting by m a positive integer or 0,

$$U_v = \int_{-m-\frac{1}{2}}^{m+\frac{1}{2}} \frac{P_v(x)}{P_v\left(\frac{2v-1}{2}\right)} dx. \tag{4}$$

Noticing that $P(-x) = P(x)$ and $P_v(-x) = -P_{-v+1}(x)$, it is easily proved that $U_v = U_{-v+1}$, and we therefore obtain by integrating (3)

$$\int_{-m-\frac{1}{2}}^{m+\frac{1}{2}} f(x)dx = \sum_1^r U_v \left[f\left(\frac{2v-1}{2}\right) + f\left(-\frac{2v-1}{2}\right)\right] + R \tag{5}$$

where

$$R = \int_{-m-\frac{1}{2}}^{m+\frac{1}{2}} P(x) f\left(x,\ \pm\tfrac{1}{2},\ \ldots \pm \frac{2r-1}{2}\right) dx. \tag{6}$$

Now, if $P(x)$ were a function that does not change its sign between the limits of integration, we might apply the Theorem of Mean Value to (6) and write, first, as is sometimes done‡,

$$R = f\left(\xi,\ \pm\tfrac{1}{2},\ \ldots,\ \pm\frac{2r-1}{2}\right) \int_{-m-\frac{1}{2}}^{m+\frac{1}{2}} P(x)dx, \tag{7}$$

and thereafter

$$R = \frac{f^{(2r)}(\eta)}{(2r)!} \int_{-m-\frac{1}{2}}^{m+\frac{1}{2}} P(x)dx, \tag{8}$$

denoting by ξ and η certain mean values of the arguments employed; provided only, that $f^{(2r)}(x)$ is continuous over the range considered. But $P(x)$ vanishes at the points $\pm\frac{1}{2}, \pm\frac{3}{2}, \ldots, \pm\frac{2r-1}{2}$, and (7) can, therefore, except in the case $m=0$, certainly not be true for every $f(x)$ that satisfies the condition of continuity. It is the more remarkable, that (8) can be proved, by a different method, to hold for every $f(x)$, possessing a continuous $f^{(2r)}(x)$.

In order to prove this, we divide the integral into two, integrating first from $-m-\frac{1}{2}$ to $m-\frac{1}{2}$, and thereafter from $m-\frac{1}{2}$ to $m+\frac{1}{2}$. In the second part of the interval of integration $P(x)$ does not change its sign, and the Theorem of Mean Value is, therefore, applicable here. We thus obtain

$$R = \int_{-m-\frac{1}{2}}^{m-\frac{1}{2}} P(x) f\left(x,\ \pm\tfrac{1}{2},\ \ldots,\ \pm\frac{2r-1}{2}\right) dx + \frac{f^{(2r)}(\xi_1)}{(2r)!} \int_{m-\frac{1}{2}}^{m+\frac{1}{2}} P(x)dx. \tag{9}$$

‡O. Biermann: *Vorlesungen über mathematische Näherungsmethoden, Braunschweig* 1905, p. 176; E. Lefrancq: *De l'interpolation,* Bulletin du Comité permanent des Congrès Internationaux d'Actuaires, 1922, pp. 49-98. The latter author, while admitting that he has not proved formula (8), uses it as an "indice de la précision".

Now we have, by the definition of divided differences,

$$f\left(x, \pm\tfrac{1}{2}, \ldots, \pm\frac{2r-1}{2}\right)=\frac{f\left(x, \pm\tfrac{1}{2}, \ldots, \pm\frac{2r-3}{2}, -\frac{2r-1}{2}\right)-f\left(\pm\tfrac{1}{2}, \ldots, \pm\frac{2r-1}{2}\right)}{x-\frac{2r-1}{2}};$$

inserting this in (9), and expressing $P(x)$ in the usual notation of factorials as $P(x)=(x+r-\frac{1}{2})^{(2r)}$, we obtain, for the first integral in (9),

$$\int_{-m-\frac{1}{2}}^{m-\frac{1}{2}}(x+r-\tfrac{1}{2})^{(2r-1)}f\left(x, \pm\tfrac{1}{2}, \ldots, \pm\frac{2r-3}{2}, -\frac{2r-1}{2}\right)dx$$

$$-f\left(\pm\tfrac{1}{2}, \ldots, \pm\frac{2r-1}{2}\right)\int_{-m-\frac{1}{2}}^{m-\frac{1}{2}}(x+r-\tfrac{1}{2})^{(2r-1)}dx.$$

If, in both of these integrals, we write $x-\frac{1}{2}$ for x, it is at once seen, that the second integral vanishes, while the first may be written

$$\int_{-m}^{m} x(x^2-1)\ldots(x^2-\overline{r-1}^2)\,f\left(x-\tfrac{1}{2}, \pm\tfrac{1}{2}, \ldots, \pm\frac{2r-3}{2}, -\frac{2r-1}{2}\right)dx. \tag{10}$$

It will here be convenient to introduce my usual notation of central factorials. I write

$$x^{[n]}=x\left(x+\frac{n}{2}-1\right)^{(n-1)}, \tag{11}$$

so that, denoting by δ the central difference, $\delta x^{[n]}=nx^{[n-1]}$. It follows from (11) that

$$\begin{cases} x^{[2v]} = x^2(x^2-1)\ (x^2-4)\ldots(x^2-\overline{v-1}^2), \\ x^{[2v+1]}=x(x^2-\frac{1}{4})\ (x^2-\frac{9}{4})\ldots\left(x^2-\frac{(2v-1)^2}{4}\right), \end{cases} \tag{12}$$

and it will be convenient to write $x^{[n]-1}=\dfrac{x^{[n]}}{x}=\left(x+\dfrac{n}{2}-1\right)^{(n-1)}$, so that $\delta x^{[n+1]-1}=nx^{[n]-1}$, and

$$\begin{cases} x^{[2v]-1}=x(x^2-1)\ (x^2-4)\ldots(x^2-\overline{v-1}^2), \\ x^{[2v+1]-1}=(x^2-\frac{1}{4})\ (x^2-\frac{9}{4})\ldots\left(x^2-\frac{(2v-1)^2}{4}\right). \end{cases} \tag{13}$$

It is evident, that (10) has the same form as the remainder-term in Cotes' formula for an *odd* number of terms. Making use of the results obtained in my above-quoted papers we may, therefore, at once write (10) in the form

$$\frac{2f^{(2r)}(\xi_2)}{(2r)!}\int_0^m x^{[2r]}dx,$$

so that (9) becomes

$$R = \frac{2f^{(2r)}(\xi_2)}{(2r)!}\int_0^m x^{[2r]}dx + \frac{f^{(2r)}(\xi_1)}{(2r)!}\int_{m-\frac{1}{2}}^{m+\frac{1}{2}} x^{[2r+1]-1}dx. \tag{14}$$

We proceed to show that the two terms of which R consists may be contracted into one. We begin by observing, that

$$\int_0^t x^{[2r]}dx = \frac{1}{2r+1}\int_{t-\frac{1}{2}}^{t+\frac{1}{2}} x^{[2r+1]}dx; \tag{15}$$

for both sides vanish for $t=0$, and their differential coefficients are $t^{[2r]}$ and $\frac{1}{2r+1}\,\delta t^{[2r+1]} = t^{[2r]}$respectively. We may, therefore, write (14)

$$R = \frac{2f^{(2r)}(\xi_2)}{(2r+1)!}\int_{m-\frac{1}{2}}^{m+\frac{1}{2}} x^{[2r+1]}dx + \frac{f^{(2r)}(\xi_1)}{(2r)!}\int_{m-\frac{1}{2}}^{m+\frac{1}{2}} x^{[2r+1]-1}dx. \tag{16}$$

But it is now obvious, that the coefficients of $f^{(2r)}(\xi_2)$ and $f^{(2r)}(\xi_1)$ have *the same sign*, so that we may put

$$R = \frac{2f^{(2r)}(\xi)}{(2r+1)!}\int_{m-\frac{1}{2}}^{m+\frac{1}{2}} [x^{[2r+1]} + (r+\tfrac{1}{2})x^{[2r+1]-1}]dx = \frac{2f^{(2r)}(\xi)}{(2r+1)!}\int_{m-\frac{1}{2}}^{m+\frac{1}{2}} (x+r+\tfrac{1}{2})^{(2r+1)}dx,$$

or finally, writing $x-\frac{1}{2}$ for x,

$$R = \frac{2f^{(2r)}(\xi)}{(2r+1)!}\int_m^{m+1} x^{[2r+2]-1}dx. \tag{17}$$

This formula can easily be proved to be identical with (8). We have, in fact,

$$\int_t^{t+1} x^{[2r+2]-1}dx = (2r+1)\int_0^{t+\frac{1}{2}} x^{[2r+1]-1}dx; \tag{18}$$

for both sides vanish for $t=-\frac{1}{2}$, and their differential coefficients are $(t+1)^{[2r+2]-1} - t^{[2r+2]-1} = \delta(t+\frac{1}{2})^{[2r+2]-1}$ and $(2r+1)\ (t+\frac{1}{2})^{[2r+1]-1}$ respectively, and these expressions are identical. Instead of (17) we may, therefore, write

$$R = \frac{2f^{(2r)}(\xi)}{(2r)!}\int_0^{m+\frac{1}{2}} x^{[2r+1]-1}dx \tag{19}$$

which differs only in the notation from (8).

In (5), the interval of integration was taken as $2m+1$. It is customary to choose the interval of integration as unity in which case we must put $f(x) = F\left(\frac{x}{2m+1}\right)$. If, for abbreviation, we write

(20) $$V_v=\frac{1}{2m+1}U_v,\ F_{\pm v}=F\left(\frac{v-\frac{1}{2}}{2m+1}\right)+F\left(-\frac{v-\frac{1}{2}}{2m+1}\right),$$

(21) $$O_{2m+1}^{[2r+1]-1}=\frac{2}{(2r)!\ (2m+1)^{2r+1}}\int_0^{m+\frac{1}{2}} x^{[2r+1]-1}dx,$$

the definite formula becomes

(22) $$\int_{-\frac{1}{2}}^{\frac{1}{2}} F(x)dx=\sum_1^r V_vF_{\pm v}+R, \qquad R=O_{2m+1}^{[2r+1]-1}F^{(2r)}(\xi).$$

As in the case of an odd number of terms, we consider also here two main types of quadrature-formulae for which I propose the names the *closed* and the *open* type. These names are intended to suggest that in the former case use is made of the end-points of the interval of integration, but not in the latter. The open type is, therefore, as I have shown elsewhere, suited for the numerical integration of differential equations. The closed type is obtained for $r=m+1$, the open one for $r=m$. The corresponding coefficients in the remainder-term are $O_{2r-1}^{[2r+1]-1}$ and $O_{2r+1}^{[2r+1]-1}$ respectively. I have found, for the first few values of $-O_{2r-1}^{[2r+1]-1}$, commencing with $r=1$,

$$\frac{1}{12},\ \frac{1}{6480},\ \frac{11}{37800000},\ \frac{167}{426924691200},\ \frac{173}{458209960750080},$$

and for $O_{2r+1}^{[2r+1]-1}$,

$$\frac{1}{36},\ \frac{19}{90000},\ \frac{751}{1016487360},\ \frac{2857}{1928493100800},\ \frac{434293}{225892143341061120}.$$

The resulting formulae with 2, 4, 6, 8 and 10 terms may readily be written down from the appended tables where the coefficient of the remainder-term, except the first, is approximate, but on the safe side, *e.g.*, $.0^316=.00016>\frac{1}{6480}$. Thus, the four-terms formula of the *closed* type, or the *three-eighths* rule, is

(23) $$\int_{-\frac{1}{2}}^{\frac{1}{2}} F(x)dx=\frac{3F_{\pm 1}+F_{\pm 2}}{8}-.0^316\ F^{(4)}(\xi),$$

while the four-terms formula of the *open* type is

QUADRATURE-FORMULAE OF THE CLOSED TYPE

Number of terms	$F_{\pm 1}$	$F_{\pm 2}$	$F_{\pm 3}$	$F_{\pm 4}$	$F_{\pm 5}$	Common Divisor	Remainder-Term
2	1					2	$-\frac{1}{12}F^{(2)}(\xi)$
4	3	1				8	$-.0^3\ 16F^{(4)}(\xi)$
6	50	75	19			288	$-.0^6\ 30F^{(6)}(\xi)$
8	2989	1323	3577	751		17280	$-.0^9\ 40F^{(8)}(\xi)$
10	5778	19344	1080	15741	2857	89600	$-.0^{12}\ 38F^{(10)}(\xi)$

QUADRATURE-FORMULAE OF THE OPEN TYPE

Number of terms	$F_{\pm 1}$	$F_{\pm 2}$	$F_{\pm 3}$	$F_{\pm 4}$	$F_{\pm 5}$	Common Divisor	Remainder-Term
2	1					2	$\frac{1}{36}F^{(2)}(\xi)$
4	1	11				24	$.0^3\,22F^{(4)}(\xi)$
6	562	−453	611			1440	$.0^6\,74F^{(6)}(\xi)$
8	−1711	4967	−2803	1787		4480	$.0^8\,15F^{(8)}(\xi)$
10	8891258	−17085616	15673880	−6603199	2752477	7257600	$.0^{11}\,20F^{(10)}(\xi)$

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} F(x)dx = \frac{F_{\pm 1}+11F_{\pm 2}}{24} + .0^3 22F^{(4)}(\xi). \tag{24}$$

In (23), $F_{\pm 2}$ is calculated at the end-points of the interval $\pm\frac{1}{2}$, but not in (24). The corresponding six-term formulae are of the closed type

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} F(x)dx = \frac{50F_{\pm 1}+75F_{\pm 2}+19F_{\pm 3}}{288} - .0^6 30F^{(6)}(\xi) \tag{25}$$

and in the case of the open type

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} F(x)dx = \frac{562F_{\pm 1}-453F_{\pm 2}+611F_{\pm 3}}{1440} + .0^6 74F^{(6)}(\xi). \tag{26}$$

A comparison with the remainder-terms of the formulae with an odd number of terms (see my above-quoted papers) shows that comparatively little is gained by adding a term to the formula, if the result is a formula with an even number of terms. The latter kind of formula should, therefore, as a rule not be employed, if the values of $F(x)$ required for the formulae with an odd number of terms are readily accessible.

It is well known that from a given quadrature-formula we may derive a new one by adding $K\delta^m f(0)$ to the formula, in which case the remainder-term must evidently be diminished by $K\delta^m f(0) = Kf^{(m)}(\eta)$. Most of the formulae obtained in this way present no particular interest, but occasionally a result of practical value may be obtained, as the resulting formula may either have simpler coefficients or be more accurate than Cotes' formula with the same number of terms.

In particular, the constant K may be chosen in such a way, that a term in Cotes' formula vanishes. As an example, we will consider the seven-term formula of the closed type (see my above-quoted paper) which, if the interval of integration is chosen as 6, may be written

$$\int_{-3}^{3} f(x)dx = \frac{1}{140}\,(272f_0+27f_{\pm 1}+216f_{\pm 2}+41f_{\pm 3}) - \frac{9}{1400}f^{(8)}(\xi). \tag{27}$$

If, from this formula, we deduct

$$\frac{9}{700}\delta^6 f(0)=\frac{9}{700}(-20f_0+15f_{\pm 1}-6f_{\pm 2}+f_{\pm 3})$$

and add $\frac{9}{700}f^{(6)}(\eta)$ to the remainder-term, we obtain *Hardy's* formula

(28) $$\int_{-3}^{3} f(x)dx=2.2f_0+1.62f_{\pm 2}+.28f_{\pm 3}+R$$

where

(29) $$R=\frac{9}{700}\,[f^{(6)}(\eta)-\tfrac{1}{2}f^{(8)}(\xi)].$$

It is seen, that Hardy's formula only contains 5 terms, as $f_{\pm 1}$ has disappeared, and it must therefore be compared with Cotes' five-term formula of the closed type or

(30) $$\int_{-\frac{1}{2}}^{\frac{1}{2}} F(x)dx=\frac{12F_0+32F_{\pm 1}+7F_{\pm 2}}{90}-\frac{F^{(6)}(\xi)}{1935360}.$$

Transforming the interval in (28), so that it becomes ± 1, we obtain

(31) $$\int_{-\frac{1}{2}}^{\frac{1}{2}} F(x)dx=\frac{1.1F_0+.81F_{\pm 2}+.14F_{\pm 3}}{3}+\frac{F^{(6)}(\eta)-\frac{1}{72}F^{(8)}(\xi)}{21772800}.$$

Comparing (31) with (30), it appears that in the frequently occurring cases where $\frac{1}{72}F^{(8)}(\xi)$ is small in comparison with $F^{(6)}(\eta)$, a decimal is gained by employing Hardy's formula instead of Cotes'. On the other hand, the accurate examination of the remainder-term is a little more troublesome in the former case than in the latter.

Another popular formula is obtained by adding $\frac{1}{140}\,\delta^6 f(0)$ to (27) and deducting $\frac{1}{140}\,f^{(6)}(\eta)$ from the remainder-term. The result is *Weddle's* formula

(32) $$\int_{-3}^{3} f(x)dx=\frac{3}{10}(6f_0+f_{\pm 1}+5f_{\pm 2}+f_{\pm 3})+R$$

where

(33) $$R=-\frac{1}{140}\,[f^{(6)}(\eta)+.9f^{(8)}(\xi)].$$

In this case no term vanishes, and the formula should, therefore, be compared with the seven-term formula (27). It is seen, that the only advantage of Weddle's formula is the simpler coefficients, and that (27) should be preferred if $f^{(8)}(\xi)$ is numerically smaller than $f^{(6)}(\eta)$.

It may be added, in conclusion, that quadrature-formulae may occasionally produce results which are better than may be inferred from the particular form we have given the remainder-term. This is not strange; and it is more than likely, that it is possible to indicate more precise limits for the error, if we know more about the function than is assumed in this paper, *i.e.*, that it possesses a continuous differential coefficient of a certain order. Much remains to be done in the way of establishing special remainder-terms for functions with more special properties; but until it is done, it would be unwise to disregard the warning given by the remainder-forms we do possess and, to quote an example, apply the three-eighths rule over a range including a point of discontinuity of $F^{(4)}(x)$.

A FORMULA FOR THE SOLUTION OF SOME PROBLEMS IN SAMPLING

BY MR. EDWARD C. MOLINA,
American Telephone and Telegraph Company, New York, N.Y., U.S.A.

The class of problems contemplated is that which may be simulated by drawings from an urn containing a known number of balls of which an unknown number are white.

Let

N = total balls in urn,
x = unknown number of white balls in urn,
n = number of balls drawn,
c = number of white balls drawn,
Σ = symbol for finite integration,
S = symbol for summation,
$P(X_1, X_2)$ = a posteriori probability that $X_1 \leq x \leq X_2$,
$W(X)$ = a priori probability that $x = X$.

By Bayes' Theorem we have, assuming that the drawn balls are not returned to the urn,

$$(1) \qquad P(X_1, X_2) = \frac{\overset{X_2}{\underset{X_1}{S}} W(x) \binom{x}{c}\binom{N-x}{n-c}}{\overset{N-n+c}{\underset{c}{S}} W(x) \binom{x}{c}\binom{N-x}{n-c}}.$$

The practical application of this formula presents two serious difficulties. In the first place, a decision has to be made regarding the function $W(x)$ the form of which is seldom known. The usual procedure is to assume $W(x)$ a constant and write

$$(2) \qquad P(X_1, X_2) = \overset{X_2}{\underset{X_1}{S}} \binom{x}{c}\binom{N-x}{n-c} \Big/ \overset{N-n+c}{\underset{c}{S}} \binom{x}{c}\binom{N-x}{n-c}.$$

As you know this procedure has lent itself to endless misunderstanding and disputes. However, it will be adopted here and in justification thereof, let me recall your attention in particular to the views of Pearson and Whittaker.

In the second place how is (2) to be computed when the numbers involved are large? The standard text-books replace (2) by the Bernoulli-Laplace integral

$$P=\frac{2}{\sqrt{\pi}}\int_0^K e^{-t^2}\,dt, \tag{3}$$

taking $K=L\sqrt{nN/2pq(N-n)}$ instead of $L\sqrt{n/2pq}$ and

$$(X_2/N)=(c/n)+L,$$
$$(X_1/N)=(c/n)-L,$$

a restriction on X_1, X_2 which is not assumed in (2). When L is small compared with (c/n) and (c/n) is in the neighbourhood of 1/2 the Bernoulli-Laplace integral gives satisfactory results. But for the class of problems here contemplated $(c/n)\not>1/10$ and may be less than 1/1000. A transformation of (2) other than (3) is therefore needed.

In the Department of Development and Research of the American Telephone and Telegraph Company extensive tables are being computed from the following formula which is convenient when c is small compared with (X_2-X_1);

$$P(X_1, X_2)=\frac{\underset{t=0}{\overset{c}{S}}\left[\binom{X_1}{t}\binom{N-X_1+1}{n+1-t}-\binom{X_2+1}{t}\binom{N-X_2}{n+1-t}\right]}{\binom{N+1}{n+1}}. \tag{4}$$

This formula is an exact transformation of (2) obtained by passing from summation to finite integration and then applying the well known theorem

$$\Sigma u_x v_x = u_x\Sigma v_x - \Delta u_x \Sigma^2 v_{x+1} + \Delta^2 u_x \Sigma^3 v_{x+2} + \ldots\ldots$$

For $X_1=c$, $X_2=X$ the formula reduces to

$$P(c, X)=1-\frac{\underset{t=0}{\overset{c}{S}}\binom{X+1}{t}\binom{\overline{N+1}-\overline{X+1}}{n+1-\ t}}{\binom{N+1}{n+1}}. \tag{5}$$

Using (5) we can compute (2) from the equation

$$P(X_1, X_2)=P(c, X_2)-P(c, X_1-1).$$

A direct interpretation of the expression for $P(c, X)$ gives us the following interesting theorem:

The *a posteriori probability* that an urn contains not more than X white balls when c whites have resulted from n drawings from a total of N balls is equal to the *a priori probability* of drawing at least $(c+1)$ whites in $(n+1)$ drawings from an urn containing $(N+1)$ balls of which $(X+1)$ are white. This theorem assumes that all values of X are a priori equally likely.

LES MESURES D'APRÈS ÉCHANTILLONS

PAR M. LUCIEN MARCH,
Directeur honoraire de la Statistique Générale de la France, Paris, France.

Quand on peut observer tous les individus dont on se propose d'étudier numériquement un caractère mesurable quelconque, on détermine aisément quelques termes de comparaison conventionnels: moyenne, écart moyen, écart type, déviation, etc., à l'aide desquels on apprécie, soit la grandeur du caractère observé, pour l'ensemble des individus considérés en masse, soit la variabilité de ce caractère parmi les individus de l'ensemble.

Cependant l'observation complète n'est pas toujours possible. Souvent on y renonce parce qu'il y faudrait consacrer trop de temps, ou trop d'argent, ou bien parce que certaines parties de l'ensemble sont inaccessibles.

Par exemple Laplace a montré comment on pouvait évaluer le nombre des habitants d'un pays en calculant un coefficient que l'on applique ensuite au total des naissances enregistrées dans le pays. Moyennant certaines hypothèses, il a calculé l'erreur probable du résultat.

Parmi les hypothèses, la plus importante consiste à supposer que la population du pays, d'une part, et le groupe observé dans cette population, d'autre part, constituent deux échantillons pris au hasard, et indépendamment l'un de l'autre, parmi un nombre infiniment grand d'individus.

Ces hypothèses étant admises, et moyennant une bonne définition de l'expression «pris au hasard», la théorie des probabilitiés permet de déterminer l'erreur probable à laquelle on s'expose quand on adopte, au lieu du nombre exact inconnu, le nombre calculé d'après l'échantillon.

Plus particulièrement, un échantillon de n éléments étant pris au hasard dans un ensemble extrêmement nombreux de N éléments dont pN possèdent un certain caractère, tandis que $(1-p)N$ ne le possèdent pas, si l'on considère ν échantillons comprenant le même nombre d'éléments n, on admet que la distribution de ces échantillons, suivant le nombre des éléments qui possèdent le caractère étudié, est donnée par les termes du développement de l'expression

$$\nu[p+(1-p)]^n.$$

L'erreur probable de la moyenne, calculée d'après l'échantillon, par rapport à la moyenne calculée pour l'ensemble, est égale à $\frac{2}{3}\sqrt{\frac{p(1-p)}{n}}$ quand n est assez grand. D'après cette formule, pour comparer les tailles moyennes de deux groupes importants, l'un de 10,000 conscrits, l'autre de 100,000, il suffit de comparer les tailles moyennes de deux échantillons *d'égale importance*, de 100 conscrits par exemple, dans le groupe le plus nombreux, comme dans le groupe le moins nombreux.

Cette conclusion, bien que légitime, dans les cas où les hypothèses signalées plus haut seraient rigoureusement fondées, est difficile à admettre dans la réalité.

La remarque en a été faite par Pearson qui a traité le problème sous une autre forme, en considérant comme indépendants, non pas l'ensemble des N éléments et le groupe de n pris dans son sein, mais les deux groupes de n et de $N-n$ éléments. On ne saurait en effet regarder comme entièrement indépendants deux groupes dont l'un est une partie de l'autre. Mais l'hypothèse suivant laquelle les deux groupes de n et de $N-n$ éléments sont puisés dans une masse infinie subsiste néanmoins.

Or cette hypothèse n'est acceptable que si le caractère étudié est un caractère dit «spécifique». Mais pour définir un caractère spécifique, il faut pouvoir s'assurer que tous les individus de la masse le possèdent, soit au même degré, soit à des degrés qui ne varient que suivant la loi du hasard. Autant dire qu'aucune vérification n'est possible, en dehors de ce qu'apprend l'échantillon.

Quand on traite la question en considérant les deux groupes de n et de $N-n$ éléments qui composent l'ensemble de N éléments, on obtient comme l'erreur probable une expression* qui dépend du nombre N, qui devient nulle lorsque n devient égal à N, et devient égale à l'écart type des N éléments lorsque n devient égal à 1, résultats qui s'imposent évidemment.

Cependant, dans de nombreuses applications de la méthode statistique, les caractères étudiés sont manifestement tout autres que des caractères spécifiques. C'est donc un procédé dangereux pour la formation de l'esprit que d'appuyer le jugement sur des hypothèses manifestement en opposition avec la réalité.

Par exemple, on ne saurait admettre l'existence d'une natalité spécifique, et surtout que la natalité réelle a quelque rapport avec une natalité spécifique idéale possible, alors que l'on connait l'influence des volontés particulières et la direction de ces volontés.

De même quand on procède à des mesures anthropologiques pour des races que l'on sait particulièrement constituées de mélanges.

De même encore dans une étude des salaires dans un ensemble déterminé d'ouvriers.

Or quand on se propose de déterminer de simples caractéristiques d'un ensemble limité d'éléments, un traitement arithmétique simple suffit. Nous signalerons par exemple les propositions suivantes qui permettent—d'après la méthode appliquée par Bienaymé à d'autres problèmes—d'apprécier l'écart existant entre la moyenne observée sur un échantillon et la moyenne inconnue applicable à l'ensemble, suivant la formule signalée plus haut, mais sans faire intervenir la notion d'infinité, en restant au contraire sur le terrain de l'observation.

*Indiquée plus loin.

Soit un ensemble de N éléments différents dont les grandeurs sont $x_1, x_2, \ldots, x_N$ et qui ont pour somme S_1, la valeur moyenne étant $M = \frac{S_1}{N}$, la fluctuation* $\frac{\Sigma(x-M)^2}{N}$. On forme des groupes de ces N éléments, de façon qu'un élément particulier n'entre pas plus d'une fois dans chaque groupe. Soit ν le nombre de ces groupes.

Les propositions suivantes permettent de déterminer la fluctuation des ν moyennes calculées pour les ν groupes.

I. Moyenne générale des ν moyennes

Les moyennes des ν groupes ont pour moyenne la moyenne des N éléments de l'ensemble.

Car le nombre des éléments contenus dans la totalité des ν groupes est $n\nu$. Mais N seulement sont différents. Par suite, un élément particulier entre $\frac{n\nu}{N}$ fois dans cette totalité. La somme des éléments distincts étant S_1, le total des éléments des groupes est égal à $\frac{n\nu}{N} \times S_1$; comme ces éléments sont en nombre égal à $n\nu$, leur valeur moyenne est $\frac{S_1}{N}$.

II. Fluctuation des moyennes des groupes

La fluctuation des moyennes des groupes est égale à la fluctuation des N éléments de l'ensemble diminuée de la moyenne des fluctuations des éléments de chaque groupe.

Soit un élément x_h d'un groupe désigné par la lettre i et M_i la moyenne des éléments de ce groupe. En vertu de la propriété bien connue de la moyenne

$$\sum_{h=1}^{n} (x_h - M_i)^2 = \sum_{h=1}^{n} (x_h - M)^2 - n(M_i - M)^2.$$

Additionnons les égalités semblables que l'on peut former pour les ν groupes. Le résultat peut s'écrire:

$$\sum_{i=1}^{\nu} \sum_{h=1}^{n} (x_h - M_i)^2 = \sum_{h=1}^{n\nu} (x_h - M)^2 - n \sum_{i=1}^{\nu} (M_i - M)^2.$$

Dans le premier membre de cette égalité, ainsi que dans le premier terme du second membre, l'élément particulier x_h entre $\frac{n\nu}{N}$ fois, pour la raison indiquée au paragraphe précédent; on peut donc mettre ce second terme sous la forme $\frac{n\nu}{N}\sum_{1}^{N}(x_h - M)^2$, puisque M est la même quantité dans toutes les parenthèses.

*Pour simplifier le langage, nous adoptons la terminologie du professeur Edgeworth en appelant fluctuation: le carré de l'écart quadratique moyen.

Le second terme du second membre comprend ν carrés. Dès lors l'égalité précédente peut s'écrire sous la forme suivante:

$$n \sum_{i=1}^{\nu} \left\{ \sum_{h=1}^{n} \frac{(x_h - M_i)^2}{n} \right\} = \frac{\nu n}{N} \sum_{h=1}^{N} (x_h - M)^2 - n \sum_{i=1}^{\nu} (M_i - M)^2.$$

Cette fois, les éléments x_h qui entrent dans le premier terme du second membre sont tous différents. Désignons par μ^2 la moyenne

$$\frac{1}{\nu} \sum_{i=1}^{\nu} (M_i - M)^2$$

des écarts des moyennes des groupes par rapport à la moyenne de l'ensemble; par μ'^2 la moyenne

$$\frac{1}{N} \sum_{h=1}^{N} (x_h - M)^2$$

relative aux éléments de l'ensemble primitif;
par μ''^2 la moyenne

$$\frac{1}{n} \sum_{h=1}^{n} (x_h - M_i)^2$$

relative aux éléments d'un groupe.

Avec ces notations, l'égalité précédente prend la forme:

$$n \sum_{1}^{\nu} \mu''^2 = n\nu\mu'^2 - n\nu\mu^2,$$

d'où en divisant par $n\nu$:

$$\frac{1}{\nu} \sum_{1}^{\nu} \mu''^2 = \mu'^2 - \mu^2.$$

Si l'on désigne par μ'''^2 la moyenne des fluctuations μ''^2, on peut écrire finalement:

$$\mu'''^2 = \mu'^2 - \mu^2 \text{ ou } \mu^2 = \mu'^2 - \mu'''^2.$$

Les propositions suivantes permettent de calculer μ'''^2.

III. Somme des carrés des éléments des groupes

La somme des carrés des éléments des groupes est égale à la somme S_2 des carrés des éléments de l'ensemble primitif multipliée par $\frac{\nu n}{N}$.

En effet, dans le total des carrés des éléments des groupes, chaque élément particulier se présente $\frac{\nu n}{N}$ fois, d'après la proposition I.

Comme la somme des carrés des éléments particuliers, qui se présentent chacun une fois dans l'ensemble primitif, est égale à S_2, il en résulte que la somme des carrés des éléments des groupes est égale à:

$$\frac{\nu n}{N} \times S_2.$$

IV. Produits deux a deux des éléments de l'ensemble primitif

Dans la somme P' de ces produits, laquelle est égale à $\frac{1}{2}(S_1^2 - S_2)$, chaque élément particulier entre $N-1$ fois.

En effet, chaque élément donne avec les autres $N-1$ produits.

V. Produits deux a deux des éléments des groupes

La somme P de ces produits est égale à la somme P' des produits deux à deux des éléments de l'ensemble primitif multipliée par le facteur $\frac{(n-1)n\nu}{N-1}$.

En effet, chaque groupe contient n éléments dont chacun peut former, avec les $(n-1)$ autres, $n-1$ produits.

D'après la proposition I, parmi les $n\nu$ éléments de tous les groupes, $\frac{n\nu}{N}$ seulement sont différents. Chacun d'eux donnant naissance, dans le groupe auquel il appartient, à $(n-1)$ produits, il y a $(n-1)\frac{n\nu}{N}$ produits dans lesquels entre un élément déterminé.

D'autre part, d'après la proposition IV, dans la somme P' un élément particulier se trouve compris dans $N-1$ produits. Si donç P' est répété $\frac{(n-1)n\nu}{N}$ fois et si P l'est $N-1$ fois, on obtiendra deux sommes égales, puisque chaque élément particulier entrera le même nombre de fois dans l'une ou dans l'autre des sommes et qu'à chaque fois il reste associé avec les $N-1$ autres éléments.

On peut donc écrire:

$\frac{(n-1)n\nu}{N} P' = (N-1)P$, d'où en remplaçant P' par sa valeur $\frac{1}{2}(S_1^2 - S_2)$

$$P = \frac{(n-1)n\nu}{2(N-1)N}(S_1^2 - S_2).$$

VI. Fluctuation moyenne des éléments d'un groupe

La fluctuation moyenne des éléments d'un groupe est égale à la fluctuation des N éléments de l'ensemble multipliée par le facteur $\frac{(n-1)N}{n(N-1)}$.

Considérons un groupe quelconque de rang i dont les éléments ont pour somme nM_i. Le carré de la somme des éléments de ce groupe a pour valeur $n^2M_i^2$. Il comprend d'abord la somme des carrés de ces éléments, puis le double de la somme des produits deux à deux des mêmes éléments.

Additionnons tous les carrés formés pour les ν groupes. Le total est égal à:

$$\sum_{i=1}^{\nu} n^2 M_i^2 = n^2 \sum_{i=1}^{\nu} M_i^2.$$

Ce total comprend d'abord la somme des carrés de tous les éléments qui figurent dans les groupes, chacun de ceux-ci se présente $\frac{n\nu}{N}$ fois (prop. I), tandis que dans l'ensemble primitif il ne figure qu'une fois.

La somme des carrés des éléments qui figurent dans les ν groupes est donc égale à $S_2 \times \frac{\nu n}{N}$.

En outre, le total considéré comprend les produits deux à deux des éléments de chaque groupe. La somme de ces produits, d'après la proposition V, est égale à:

$$\frac{(n-1)n\nu}{2(N-1)N}(S_1^2 - S_2).$$

Et comme ce total, d'ailleurs égal à $n^2 \sum_{i=1}^{\nu} M_i^2$ contient le double des produits dont on vient de calculer la somme, on peut finalement écrire:

$$n^2 \sum_{i=1}^{\nu} M_i^2 = \frac{\nu n}{N} S_2 + \frac{(n-1)}{(N-1)} \frac{\nu n}{N} (S_1^2 - S_2)$$

d'où, en divisant par n et simplifiant:

$$n \sum_{i=1}^{\nu} M_i^2 = \frac{\nu}{n}\left(1 - \frac{n-1}{N-1}\right) S_2 + \frac{n-1}{N-1} \frac{\nu}{N} S_1^2.$$

D'autre part, si l'on désigne par x un élément quelconque pris parmi les $n\nu$ que l'on trouve dans les ν groupes, on a (proposition III):

$$\sum_1^{n\nu} x^2 = \frac{\nu n}{N} S_2.$$

Retranchons membre à membre les deux égalités précédentes, il vient:

$$\sum_1^{n\nu} x^2 - n \sum_1^{\nu} M_i^2 = \frac{\nu n}{N} S_2 - \frac{\nu}{N} \frac{N-n}{N-1} S_2 - \frac{n-1}{N-1} \frac{\nu}{N} S_1^2.$$

Or on sait que

$$\sum_1^{n} (x - M_i)^2 = \sum_1^{n} x^2 - n M_i^2.$$

Le premier membre de l'égalité précédente peut donc être remplacé par la somme:

$$\sum_1^{n\nu} (x - M_i)^2,$$

formée pour les ν groupes. Dès lors l'égalité précédente peut s'écrire:

$$\sum_1^{n\nu} (x - M_i^2) = \frac{n\nu}{N} \left\{ S_2 \left[1 - \frac{N-n}{n(N-1)} \right] - \frac{n-1}{n(N-1)} S_1^2 \right\}.$$

Et de là, en divisant par νn les deux membres,

$$\frac{1}{\nu}\sum_1^\nu \frac{1}{n}\sum_1^n (x-M_i)^2 = \frac{1}{N}\left\{S_2\frac{N(n-1)}{n(N-1)} - \frac{n-1}{n(N-1)}S_1^2\right\} = \frac{n-1}{n(N-1)N}(NS_2-S_1^2).$$

En employant les notations indiquées au paragraphe II, et reprenant l'égalité $\sum_1^N (x-M)^2 = S_2 - NM^2$, on remarquera que $\sum_1^N (x-M)^2 = N\mu'^2$ et $M^2 = \frac{S_1^2}{N^2}$. Donc on peut écrire:

$$N^2\mu'^2 = NS_2 - S_1^2.$$

L'égalité précédente peut alors se mettre sous la forme:

$$\frac{1}{\nu}\sum_1^\nu \mu''^2 = \frac{n-1}{N-1}\frac{1}{nN}N^2\mu'^2 = \frac{(n-1)N}{n(N-1)}\mu'^2$$

ou

$$\mu'''^2 = \frac{n-1}{N-1}\frac{N}{n}\mu'^2,$$

en désignant par μ'''^2 la moyenne des fluctuations des n éléments de chacun des groupes possibles.

Remarque.—Lorsque n et N sont des nombres un peu grands, le facteur $\frac{n-1}{N-1}\frac{N}{n}$ est presque égal à l'unité. Alors la moyenne des fluctuations des groupes est à peu près égale à la fluctuation de l'ensemble primitif.

VII. Fluctuation des moyennes des groupes

La fluctuation des moyennes des groupes est égale à la moyenne des fluctuations des groupes multipliée par le rapport $\frac{N-n}{N(n-1)}$.

En effet, d'après la proposition VI,

$$\mu'''^2 = \mu'^2 \times \frac{n-1}{N-1}\frac{N}{n}.$$

D'après la proposition II:

$$\mu^2 = \mu'^2 - \mu'''^2.$$

En conséquence:

$$\mu^2 = \mu'^2\left(1 - \frac{N(n-1)}{n(N-1)}\right) = \frac{N-n}{n(N-1)}\mu'^2 = \frac{N-n}{N(n-1)}\mu'''^2.$$

Remarques.—1. Lorsque n est assez grand et que N est un très grand nombre, l'égalité précédente se réduit à peu près à:

$$\mu^2 = \frac{\mu'''^2}{n} = \frac{\mu'^2}{n}$$

formule convenant en particulier au cas d'un ensemble infini.

2. Si $N=n$, c'est-à-dire si l'ensemble ne comprend qu'un seul groupe identique à lui-même, $\mu'''^2=\mu'^2$ et $\mu^2=0$, comme on pouvait le prévoir.

3. Si $n=1$, c'est-à-dire s'il n'y avait point de groupes formés, chaque groupe se réduisant à l'un des éléments de l'ensemble, les formules précédentes donnent:

$$\mu^2=\mu'^2, \ \mu''^2=\mu'''^2=0,$$

comme on pouvait également le prévoir.

4. Si, au lieu des groupes de n éléments, on considère des groupes de $N-n$ éléments, l'expression

$$\mu^2=\frac{N-n}{n(N-1)}\mu'^2$$

se transforme en:

$$\pi^2=\frac{n}{(N-n)(N-1)}\mu'^2.$$

Lorsque $n=1$, cette expression donne

$$\pi^2=\frac{\mu'^2}{(N-1)^2},$$

ainsi qu'il convient.

D'autre part, quand $n=N$, on a $\pi=0$. Des deux expressions précédentes, on déduit:

$$\frac{\mu}{\pi}=\frac{N-n}{n}.$$

En appelant *groupe complémentaire* celui qui est formé par les éléments restants dans l'ensemble primitif quand on a enlevé ceux d'un groupe, et en appelant écart type d'un groupe la racine carrée de la fluctuation, on peut dire que l'écart type d'un groupe et celui de son complément présentent un rapport inverse de celui du nombre des éléments composants*.

5. Lorsque N croit indéfiniment, l'expression de μ^2 tend vers $\frac{\mu'^2}{n}$ ou μ vers $\frac{\mu'}{\sqrt{n}}$; $\frac{\mu}{\pi}$ tend vers 1; les deux écarts types tendent à s'égaliser.

Application.—Vers 1893, l'Office du Travail en France a procédé à une enquête atteignant 550,000 ouvriers de la grande et de la moyenne industrie sur environ 2,600,000 qu'un recensement de ces industries eût pu atteindre. D'après cette enquête, le salaire moyen journalier des ouvriers atteints par l'enquête a été évalué à peu près à 4 francs. D'autre part, on a pu déterminer pour un groupe de 15,000 ouvriers la fluctuation des salaires de ces ouvriers; on l'a trouvé égale à 2,25.

*Cette remarque a été faite par M. Isserlis dans le Journal of the Royal Statistical Society, numéro de janvier 1918, page 78. M. Isserlis a signalé les formules précédentes démontrées d'après la théorie des courbes de fréquence.

En admettant cette valeur pour μ'''^2 on déduit des données précédentes:

$$\mu^2 = \frac{N-n}{N(n-1)}\,\mu'''^2 = \frac{2{,}6-0{,}65}{2{,}6\times 0{,}65}\,2{,}25 = 2{,}38$$

d'où

$$\mu = 1{,}54.$$

Par comparaison avec la loi de distribution normale, le salaire moyen de l'ensemble des ouvriers peut être caractérisé sous la forme $4 \pm \frac{2}{3} \times 1.54$ ou 4 ± 1. On doit admettre qu'il est compris entre 3 et 5 francs quand l'échantillon considéré donne comme moyenne 4 francs.

En résumé, de simples propositions d'arithmétique suffisent pour justifier les procédés de comparaison que comporte l'application élémentaire de la méthode statistique*. En utilisant ces propositons, on réduit au minimum les hypothèses nécessaires.

*Voir l'essai publié dans le Journal de la Société de Statistique de Paris, Décembre, 1910.

SUR UNE FORMULE GÉNÉRALE POUR LE CALCUL DES PRIMES PURES D'ASSURANCES SUR LA VIE

PAR M. MAURICE FRÉCHET,
Professeur à l'Université de Strasbourg, Strasbourg, France.

INTRODUCTION

Dans son ouvrage *Introduction à la science actuarielle,* M. Du Pasquier obtient quelques formules générales qu'il n'avait trouvées «dans aucun traité et qui renferment comme cas particuliers les formules classiques pour le calcul de la prime unique aussi bien que de la prime échelonnée».

Les formules nouvelles auxquelles il fait ainsi allusion sont des formules dans lesquelles le second membre variant avec le genre d'assurance envisagée ne se distingue pas de celui des formules classiques et représente la valeur actuelle des contrats. La nouveauté consiste dans le premier membre qui a l'expression générale suivante

$$PD_x + a(N_x - N_{x+t})$$

et représente la valeur actuelle des primes (soient une prime unique P, et une prime annuelle a constante, anticipée, temporaire pendant t années) pour l_x assurés, à l'époque de leur naissance.

Il nous a paru que l'utilité d'une formule générale était beaucoup plus manifeste pour l'expression du second membre, étant donnée la grande variété des engagements pris par les assureurs. Tandis que le paiement des primes ne s'écarte guère des trois types: prime unique, prime constante vie entière; prime constante temporaire.

On trouvera dans ce qui suit une formule générale qui s'applique, quelle que soit la forme du contrat (supposé sur une seule tête).

L'un des buts poursuivis est celui même de M. Du Pasquier: obtenir une formule générale qu'il suffit de particulariser dans chaque cas concret sans avoir besoin d'imaginer à chaque fois un raisonnement nouveau. La généralité obtenue—concernant aussi bien les différentes formes d'engagements de l'assureur que les différents modes de paiement des primes—est considérablement plus étendue que celle qu'a atteint M. Du Pasquier.

Toutes les fois qu'on sera embarrassé par une combinaison nouvelle, on pourra, au lieu de rechercher le prix de revient par une subtile méthode indirecte, remplacer purement et simplement les quantités qui figurent dans la formule générale (7) par celles qui correspondent à la combinaison considérée.

Nous ne recommandons pas pourtant particulièrement cette manière de faire. C'est pour d'autres raisons que nous signalons cette formule. D'une part, en effet, elle montre que toute prime peut s'exprimer en fonction des nombres de commutation, vérité d'expérience qui se trouve ici établie de façon tout à fait générale. Et, d'autre part, elle permet d'écarter, comme nous le verrons plus loin, une objection justifiée portant sur le mode d'établissement des formules particulières.

Données et notations.—Nous supposerons d'abord, dans ce qui suit, que l'unité de temps envisagée est l'année.

Soit alors x l'âge en années de la tête assurée, au moment de la signature du contrat.

On ne sait pas d'avance quelles seront les sommes payées soit par l'assureur, soit par le payeur de primes au cours de la n^e année du contrat. Mais on sait d'avance quelles seront ces sommes en cas de vie, et quelles seront ces sommes en cas de décès. C'est cela que doit spécifier le contrat.

On sait donc quelle somme V_{x+n} sera payée par l'assureur au momen où l'assuré entre dans l'âge $x+n$, s'il atteint cet âge et l'on sait quelle somme Δ_{x+n} paiera l'assureur l'année du décès de l'assuré, si ce décès a lieu à l'âge $x+n$. Cette dernière somme pouvant être payée à des époques variables suivant le mois du décès, nous supposerons par exemple qu'elle est *en moyenne* payable en fin d'année de décès.

D'autre part, les primes n'étant jamais payables que pendant la vie du payeur de primes, nous savons qu'en entrant dans l'âge $x+n$ celui-ci devra payer une prime p_{x+n} et que l'assureur ne recevra plus rien du chef de ce contrat après le décès du payeur de primes.

Bien entendu, dans les cas particuliers, certaines des sommes V_{x+n}, Δ_{x+n}, p_{x+n} pourront être nulles.

Calcul du prix de revient du contrat.—Ceci étant, appelons U le prix de revient du contrat, c'est-à-dire la prime pure unique équivalente à l'ensemble des primes convenues.

Puisque pour les primes pures, nous faisons abstraction des charges, des bénéfices et des pertes, il est indifférent pour le payeur de primes de s'adresser à un seul assureur qui lui garantira l'ensemble des engagements inscrits dans le contrat, ou de s'adresser à autant d'assureurs qu'il y a de sommes V_{x+n}, Δ_{x+n} à assurer.

Ainsi U est égal à la somme des primes pures relatives aux paiements de chacune de ces sommes prises isolément.

Or V_{x+n} est une somme à payer au moment où l'assuré entre dans l'âge $x+n$ et seulement s'il atteint cet âge. La prime afférente à V_{x+n} seul est donc la prime pure unique à payer à l'âge x pour un capital V_{x+n} différé de n années en cas de vie. Cette prime est manifestement proportionnelle à V_{x+n}. Elle est égale à $V_{x+n}.\ ({}_nE_x)$ en désignant par la notation $({}_nE_x)$ la prime correspondant à un capital assuré de 1.

De même, la prime afférente à Δ_{x+n} est égale à $\Delta_{x+n}.\ ({}_nF_x)$ en désignant par $({}_nF_x)$ la prime pure unique à payer à l'âge x pour une assurance en cas de décès temporaire pendant une année, différée de n années et pour un capital de 1. Finalement

$$U = V_x + V_{x+1} \cdot ({}_1E_x) + \ldots + V_{x+n} \cdot ({}_nE_x) + \ldots$$
$$+ \Delta_x \cdot ({}_0F_x) + \Delta_{x+1} \cdot ({}_1F_x) + \ldots + \Delta_{x+n} \cdot ({}_nF_x) + \ldots$$

Mais U est équivalent à l'ensemble des primes p_{x+n}. C'est-à-dire qu'au lieu de payer ces primes aux époques convenues, l'assuré pourrait en payant U charger un nouvel assureur de lui payer les sommes p_{x+n}. Autrement dit, U est le prix de revient du contrat consistant à substituer aux sommes V_{x+n} les sommes p_{x+n} et à supprimer les paiements des sommes Δ_{x+n}. Par suite, la formule précédente conviendra à ce nouveau contrat moyennant les modifications indiquées. D'où

$$U = p_x + p_{x+1} \cdot ({}_1E_x) + p_{x+2} \cdot ({}_2E_x) + \ldots + p_{x+n} \cdot ({}_nE_x) + \ldots$$

On a donc finalement une première formule générale

$$\begin{aligned} & p_x + p_{x+1} \cdot ({}_1E_x) + p_{x+2} \cdot ({}_2E_x) + \ldots + p_{x+n} \cdot ({}_nE_x) + \ldots = \\ (1) \qquad & V_x + V_{x+1} \cdot ({}_1E_x) + V_{x+2} \cdot ({}_2E_x) + \ldots + V_{x+n} \cdot ({}_nE_x) + \ldots \\ & + \Delta_x \cdot ({}_0F_x) + \Delta_{x+1} \cdot ({}_1F_x) + \Delta_{x+2} \cdot ({}_2F_x) + \ldots + \Delta_{x+n} \cdot ({}_nF_x) + \ldots \end{aligned}$$

On peut simplifier cette formule en calculant les $({}_nF_x)$ en fonction des $({}_nE_x)$, suivant un raisonnement connu.

Si un assuré d'âge x paie une prime unique $({}_nF_x) + ({}_{n+1}E_x)$ l'assureur devra payer 1 à la fin de la $(n+1)^{\text{e}}$ année du contrat, soit au bénéficiaire si l'assuré décède au cours de cette année, soit à l'assuré si celui-ci survit à la fin de la même année. L'assureur ne paie rien dans les autres cas. On voit qu'en somme, l'assureur s'engage à payer 1 à la fin de la $(n+1)^{\text{ième}}$ année du contrat, si l'assuré est en vie en entrant dans l'âge $x+n$. Il revient au même pour l'assureur d'avancer ce paiement de 1 d'une année en escomptant ce paiement, c'est-à-dire en payant $1 \times \frac{1}{1+i}$ ou v francs si on pose $v = \frac{1}{1+i}$, i désignant le taux d'intérêt annuel pour un franc. Finalement $({}_nF_x) + ({}_{n+1}E_x)$ est la prime pure unique payable à l'âge x pour un capital v différé de n années en cas de vie. D'où:

$$(2) \qquad ({}_nF_x) = v \cdot ({}_nE_x) - ({}_{n+1}E_x).$$

La formule (1) devient alors:

$$\begin{aligned} & p_x + p_{x+1} \cdot ({}_1E_x) + p_{x+2} \cdot ({}_2E_x) + \ldots + p_{x+n} \cdot ({}_nE_x) + \ldots = \\ (3) \qquad & V_x + V_{x+1} \cdot ({}_1E_x) + V_{x+2} \cdot ({}_2E_x) + \ldots + V_{x+n} \cdot ({}_nE_x) + \ldots \\ & + v \cdot \Delta_x + (v\Delta_{x+1} - \Delta_x)({}_1E_x) + (v\Delta_{x+2} - \Delta_{x+1})\ ({}_2E_x) + \ldots \\ & + (v\Delta_{x+n} - \Delta_{x+n-1})({}_nE_x) + \ldots \end{aligned}$$

Cette formule, s'appliquant à toutes les combinaisons d'assurances sur une seule tête, *est indépendante de la théorie des probabilités*. Elle suppose seulement que, pour chaque durée n et pour chaque âge x, le prix d'un contrat d'assurances de capital différé de n années en cas de vie a une valeur déterminée proportionelle au montant du capital assuré. Le prix $({}_nE_x)$ de ce contrat pour un capital assuré de 1 franc *pourrait, par exemple, être déterminé expérimentalement* au moment où les opérations de l'assureur atteindraient la stabilité. Il est d'avance évident que ce prix doit être inférieur à la valeur actuelle v^n d'un paiement certain de 1 franc différé de n années. On aurait donc à déterminer expérimentalement le coefficient k tel que

$$({}_nE_x) = k \,.\, v^n.$$

Introduction de la probabilité de survie.—Pour calculer ${}_nE_x$ le raisonnement ordinairement usité peut se présenter sous la forme suivante:

Supposons qu'au même moment, un certain nombre L_x d'assurés d'âge x souscrivent chacun une assurance pour un capital de 1 franc différé de n années en cas de vie. A moins que certains assurés présentent des risques spéciaux tarifés à part, l'assureur est obligé de leur demander à chacun la même prime (pure unique) ${}_nE_x$. Il reçoit donc au total la somme $L_x \,.\, ({}_nE_x)$. Au bout de n années, cette somme placée à intérêts composés est devenue

$$L_x \,.\, ({}_nE_x) \,.\, (1+i)^n$$

et elle devrait compenser le versement que l'assureur doit faire aux survivants. Si donc le nombre de ceux-ci est désigné par L_{x+n} on voit qu'il n'y aurait ni bénéfice, ni perte si l'on avait

$$L_x \,.\, ({}_nE_x) \,.\, (1+i)^n = L_{x+n}$$

ou en posant encore $v = \dfrac{1}{1+i}$:

$${}_nE_x = \frac{L_{x+n}}{L_x} \,.\, v^n.$$

Mais l'assureur ne connait pas d'avance la valeur de L_{x+n}; il sait seulement que *si le nombre L_x des contractants est assez grand*, le rapport $\dfrac{L_{x+n}}{L_x}$ est, pratiquement, voisin d'un nombre donné par les tables de survie sous la forme $\dfrac{l_{x+n}}{l_x}$, et qu'on appelle probabilité de survie de l'âge x à l'âge $x+n$. Dans l'ignorance de la valeur précise de $\dfrac{L_{x+n}}{L_x}$ pour ceux de ses assurés qui ont contracté à l'âge x une assurance d'un capital différé de n années, l'assureur adopte la solution équitable consistant à remplacer ce rapport inconnu par sa valeur moyenne connue $\dfrac{l_{x+n}}{l_x}$ et il prend

$${}_nE_x = \frac{l_{x+n}}{l_x} \,.\, v^n.$$

Objection.—Le raisonnement précédent repose sur l'hypothèse que L_x est grand. Or, même pour une grosse entreprise d'assurance, le nombre de ceux qui contractent une assurance de capital différé, précisément à l'âge x et précisément pour une durée de n années n'est pas très considérable.

Réfutation.—La formule (2) permet de réfuter l'objection, ou, plutôt d'en diminuer la portée en étendant celle du raisonnement. En effet, la formule (3) montre qu'on peut considérer toute combinaison d'assurance comme se ramenant à une combinaison d'assurances de capitaux différés de diverses durées. De sorte que tout assuré d'âge x peut être considéré comme ayant contracté une assurance d'un capital différé de n années pourvu que l'un des montants correspondants p_{x+n}, V_{x+n} ou $v\Delta_{x+n}-\Delta_{x+n-1}$ ne soit pas nul. Il suffira pour cela qu'il ait à payer une prime à l'âge $x+n$ ou qu'il soit couvert par une assurance soit en cas de vie, soit en cas de décès jusqu'à l'âge $x+n$.

On voit alors que le groupe sur lequel portait le raisonnement précédent est beaucoup plus étendu que celui considéré tout d'abord et par suite qu'il n'est plus excessif d'admettre que le rapport observé $\frac{L_{x+n}}{L_x}$ est voisin du rapport prévu $\frac{l_{x+n}}{l_x}$.

Introduction des nombres de commutation.—On peut maintenant remplacer dans la formule (2) les nombres ${}_nE_x$ par leurs valeurs. On a vu que

$$(4)\qquad {}_nE_x=v^n\,\frac{l_{x+n}}{l_x}\,.$$

On évitera l'emploi simultané de deux tables, une de survie, une d'intérêts, en calculant d'avance les quantités $D_x=l_xv^x$ et en remarquant qu'alors

$$(5)\qquad {}_nE_x=\frac{D_{x+n}}{D_x}\,.$$

Au lieu de remplacer directement $({}_nE_x)$ par cette valeur dans la formule (3) il vaut mieux se servir de la formule (1) après avoir calculé ${}_nF_x$ au moyen des formules (2) et (4):

$${}_nF_x=v\,.\left(v^n\frac{l_{x+n}}{l_x}\right)-v^{n+1}\,\frac{l_{x+n+1}}{l_x}\,,$$

ou

$${}_nF_x=v^{n+1}\,\frac{d_{x+n}}{l_x}\,,$$

en appelant d_{x+n} le nombre de décès $l_{x+n}-l_{x+n+1}$ à l'âge $x+n$ donné par la table de survie. En posant $C_x=d_x\,.\,v^x$ on aura

$$(6)\qquad {}_nF_x=v\,\frac{C_{x+n}}{D_x}.$$

Il ne reste plus qu'à remplacer dans (1) les quantités ${}_nE_x$ et ${}_nF_x$ au moyen des formules (5) et (6). On obtient ainsi la formule finale

$$(7)\qquad \begin{aligned} &p_x\,.\,D_x+p_{x+1}\,.\,D_{x+1}+p_{x+2}\,.\,D_{x+2}+\,\ldots\,+p_{x+n}\,.\,D_{x+n}+\,\ldots\,=\\ &V_x\,.\,D_x+V_{x+1}\,.\,D_{x+1}+V_{x+2}\,.\,D_{x+2}+\,\ldots\,+V_{x+n}\,.\,D_{x+n}+\,\ldots\\ &+v[\Delta_x\,.\,C_x+\Delta_{x+1}\,.\,C_{x+1}+\Delta_{x+2}\,.\,C_{x+2}+\,\ldots\,+\Delta_{x+n}\,.\,C_{x+n}+\,\ldots\,].\end{aligned}$$

Telle est la formule générale qui permet le calcul des primes p_{x+n} payables en entrant dans l'âge $x+n$ pour une assurance contractée à l'âge x et qui garantit le paiement de sommes telles que V_{x+n} payables au n^e anniversaire du contrat si l'assuré est en vie à cette époque, et de sommes telles que Δ_{x+n} payables en moyenne en fin d'année de décès si l'assuré meurt à l'âge $x+n$.

Méthode dite eulérienne.—Une fois légitimée la formule (7), on peut la retrouver d'une façon plus rapide en écrivant que si l_x assurés du même âge contractent la même assurance, la valeur actuelle, à une époque quelconque des engagements des l_x assurés est égale à la valeur actuelle, à la même époque, des engagements de l'assureur en supposant que les décès des contractants suivent exactement la loi de mortalité exprimée par la table dressée d'avance des nombres l_x. En prenant pour époque de base leur naissance, la valeur actuelle des primes p_{x+n} payées par les l_{x+n} survivants à l'âge $x+n$ sera égale à la somme $p_{x+n} \cdot l_{x+n}$ escomptée de $x+n$ années, soit

$$p_{x+n} \cdot l_{x+n} \cdot v^{x+n} = p_{x+n} \cdot D_{x+n}.$$

De même, la valeur actuelle des capitaux V_{x+n} sera $V_{x+n} \cdot D_{x+n}$. Enfin, la valeur actuelle des capitaux Δ_{x+n} payés à d_{x+n} bénéficiaires sera, $x+n+1$ années auparavant:

$$\Delta_{x+n} \cdot d_{x+n} \cdot v^{x+n+1} = v\Delta_{x+n} \cdot C_{x+n}.$$

D'où la formule (7).

Seulement, cette façon de procéder constitue plutôt un moyen mnémotechnique qu'une méthode de démonstration. Les nombres réels des survivants et des décès sont *très différents* des nombres l_x, d_x. Ce sont les *rapports* de ces nombres qui sont voisins des rapports correspondants. Ce qui fait que l'équation obtenue par ce second moyen est cependant valable, c'est que cette équation est homogène par rapport aux l_x.

Dans le cas particulier envisagé par M. Du Pasquier on a:

$$p_x = P+a,\ p_{x+1} = a,\ \ldots,\ p_{x+t-1} = a,\ p_{x+t} = p_{x+t+1} = \ldots = 0;$$

le premier membre de (7) devient

$$(P+a)D_x + \ldots + aD_{x+t-1} = PD_x + a(D_x + \ldots + D_{x+t-1})$$

ou

$$PD_x + a(N_x - N_{x+t})$$

en posant

(8) $$N_x = D_x + D_{x+1} + \ldots$$

On retrouve ainsi l'expression de M. Du Pasquier.

Remarques

I. D'après ce qui précède, nous avons appelé p_r, V_r les sommes à payer respectivement par l'assuré et par l'assureur, si l'assuré est en vie en entrant dans l'âge r. Et nous avons supposé ces sommes payables au $r^{ième}$ anniversaire, ou comptées pour leurs valeurs à cette époque. Si l'existence de l'assuré, en entrant dans l'âge r déterminait le paiement de sommes payables à d'autres époques, nous aurions donc à prendre pour p_r, V_r les valeurs actuelles de ces sommes (respectivement dues par l'assuré et par l'assureur) capitalisées ou escomptées au $r^{ième}$ anniversaire. De même, Δ_r dans notre notation désigne la valeur capitalisée ou escomptée à la fin de l'année du décès lorsque ce décès a lieu à l'âge r.

Si donc, par exemple, on admet que les paiements en cas de décès ont lieu immédiatement après le décès et que comme les décès se répartissent assez régulièrement tout le long de l'année, les paiements ont lieu en moyenne au milieu de l'année du décès, alors si l'on appelle δ_r la valeur effective du capital assuré en cas de décès à l'âge r on devra prendre pour Δ_r

$$\Delta_r = (1+i)^{\frac{1}{2}}\delta_r = \frac{\delta_r}{\sqrt{v}}$$

et comme les termes en Δ dans la formule (7) sont multipliés par v, il faudra remplacer dans cette formule Δ_r par δ_r et v par $\sqrt{v}$.

II. Le raisonnement et les formules obtenues ne seraient pas altérées si l'unité de temps était, non pas l'année, mais le semestre, le trimestre, le mois, . . . Seulement, les notations x, i, v . . . se rapporteraient à la nouvelle unité de temps. Comme les tables de mortalité se rapportent, en général, au cas où l'unité de temps est l'année, il faudrait interpoler selon les méthodes développées dans les traités d'assurances pour obtenir les valeurs des l_x, D_x, . . . relatives aux semestres, trimestres, . . .

Applications

I. Traitons, d'abord, à titre d'exemple, une des combinaisons usuelles d'assurances: *rente viagère à jouissance différée et à capital réservé.* Soit à calculer la prime annuelle P à payer à partir de l'âge x pour assurer le paiement d'une rente viagère de 1, différée de n années, étant entendu qu'au décès de la personne assurée, les ayants-droit recevront la totalité des primes versées, sans intérêt. Le premier terme de la rente sera payé à l'entrée de l'assuré dans l'âge $x+n$ et suivra d'une année le versement de la dernière prime.

On a évidemment, dans ce cas:

$$p_x = p_{x+1} = \ldots = p_{x+n-1} = P;\ p_{x+n} = p_{x+n+1} = \ldots = 0,$$
$$V_x = V_{x+1} = \ldots = V_{x+n-1} = 0;\ V_{x+n} = V_{x+n+1} = \ldots = 1,$$
$$\Delta_x = P;\ \Delta_{x+1} = 2P;\ \ldots \Delta_{x+n-1} = nP = \Delta_{x+n} = \Delta_{x+n+1} = \ldots$$

En portant dans la formule (7) les valeurs des p_{x+i}, V_{x+j}, Δ_{x+k}, indiquées ci-dessus, celle-ci devient:

$$P(D_x + D_{x+1} + \ldots + D_{x+n-1}) = D_{x+n} + D_{x+n+1} + \ldots$$
$$+vP(C_x + 2C_{x+1} + \ldots + nC_{x+n-1} + nC_{x+n} + nC_{x+n+1} + \ldots)$$

ou

$$P(N_x - N_{x+n}) = N_{x+n} + vP(R_x - R_{x+n}),$$

en posant conformément à l'usage

$$M_x = C_x + C_{x+1} + \ldots ; \; R_x = M_x + M_{x+1} + \ldots$$

D'où enfin

$$P = \frac{N_{x+n}}{N_x - N_{x+n} - v(R_x - R_{x+n})}.$$

II. Mais il est surtout intéressant d'appliquer la formule (7) à des combinaisons peu courantes, pour lesquelles les Traités d'assurances n'offrent pas une formule toute faite et qui cependant se présentent à l'occasion aux actuaires. Nous conseillons, pour éviter toute erreur, sur l'interprétation des indices des quantités p_r, V_r, Δ_r et de leurs dates de paiement, de rétablir directement dans chaque cas la formule (7) en employant le raisonnement abrégé indiqué plus haut.

A titre d'exemple, nous traiterons un *problème posé aux examens de l'Institut des Actuaires français.*

En novembre 1905, on demandait aux candidats de calculer la prime pure P à verser viagèrement pendant n' années pour garantir le paiement par l'assureur en cas de décès de l'assuré des annuités restant dues pour amortir un emprunt A que l'assuré a contracté à l'âge x en même temps que sa police d'assurance au taux t pour 1 avec remboursement en n années. Les annuités sont payables en fin d'année.

Escomptons les engagements de l'assureur et de l_x assurés à l'époque de la naissance, en supposant que la mortalité des assurés se conforme à une table de survie donnant d'avance les valeurs des l_r survivants à l'âge r, et désignons par i le taux technique (pour 1) de l'assurance. La valeur actuelle des primes est, en posant $v = \dfrac{1}{1+i}$.

$$(11) \quad Pl_x v^x + Pl_{x+1} v^{x+1} + \ldots + Pl_{x+n'-1} v^{x+n'-1} = P(D_x + \ldots + D_{x+n'-1}) = P(N_x - N_{x+n'}).$$

Le seul engagement de l'assureur consiste à payer en cas de décès à l'âge r et en fin d'année de décès une certaine somme Δ_r. De sorte que la valeur actuelle des engagements de l'assureur est:

$$(12) \qquad \Delta_x d_x v^{x+1} + \Delta_{x+1} d_{x+1} v^{x+2} + \ldots = v[\Delta_x C_x + \Delta_{x+1} C_{x+1} + \ldots].$$

Calculons les Δ; Δ_{x+r} est la somme des valeurs escomptées en fin d'âge $x+r$ des annuités (que nous appellerons a) restant à payer. Soit, en posant $w = \dfrac{1}{1+t}$:

$$\Delta_{x+r} = a + aw + aw^2 + \ldots + aw^{n-1-r} = \frac{a}{tw}(1 - w^{n-r}).$$

La valeur de l'expression (12) est donc

$$\frac{av}{tw}\{(1-w^n)C_x + \ldots + (1-w^{n-r})C_{x+r} + \ldots + (1-w)C_{x+n-1}\}$$

$$= \frac{av}{tw}\{M_x - M_{x+n} - (w^n C_x + \ldots + w^{n-r} C_{x+r} + \ldots + wC_{x+n-1})\}.$$

On a donc l'expression de P:

$$P=\frac{av}{tw(N_x-N_{x+n'})}\left\{M_x-M_{x+n}-(w^nC_x+\ldots+w^{n-r}C_{x+r}+\ldots+wC_{x+n-1})\right\}.$$

Si on donne, non pas a, mais le capital A à amortir, on remplacera a par $\frac{At}{1-w^n}$. De sorte que

$$P=\frac{Av}{w(1-w^n)\,(N_x-N_{x+n'})}\left\{M_x-M_{x+n}-(w^nC_x+\ldots+w^{n-r}C_{x+r}+\ldots+wC_{x+n-1})\right\}.$$

Dans le cas particulier où $i=t$, la formule se simplifie:

$$w^nC_x+\ldots+w^{n-r}C_{x+r}+\ldots+wC_{x+n-1}=$$

$$v^nd_xv^x+\ldots+v^{n-r}d_{x+r}v^{x+r}+\ldots+vd_{x+n-1}v^{x+n-1}=v^{x+n}(l_x-l_{x+n})=v^nD_x-D_{x+n}$$

et

$$P=\frac{A}{1-v^n}\left\{\frac{M_x-M_{x+n}+D_{x+n}-v^nD_x}{N_x-N_{x+n'}}\right\}.$$

MATHEMATICAL LAW OF MORTALITY—A SUGGESTION

BY W. PALIN ELDERTON,
Actuary and Manager, Equitable Life Assurance Society, London, England.

1. The suggestion which I put forward in this preliminary paper may be set out as follows:

If the numbers exposed to risk in a mortality experience are expressed as a frequency curve then it is well worth while to see whether the survivors at the end of a year out of those exposed to risk will also be a frequency curve and to study the law resulting from such an assumption.

2. The exposed to risk represents a number of persons living at various ages and the survivors calculated from this exposed to risk by multiplying it at each age by the probability of surviving a year at that age will give another exposed to risk. Thus if E_x be written for the exposed at age x and p_x for the probability of surviving a year then $E_x p_x$ is an exposed to risk although it would not be the exposed found by continuing the investigation which led to E_x for a further year unless

(*a*) the rates of mortality remained unchanged,

(*b*) there are no additions by way of new business (or immigrations) and no withdrawals (or emigrations).

The theory to which our suggestion leads may, therefore, be expressed as follows: if E_x be represented by, or be assumed to be, a frequency curve $f(x)$, then $p_x E_x$ will be a frequency curve.

3. Before we turn to such statistical evidence as can be given conveniently at the present time it will be well to consider the algebraic form of p_x implied by our theory. In order to do this we must describe briefly the forms usually taken by the exposed to risk. In the case of experiences drawn from life assurance statistics, etc., the exposed to risk is small at the early ages, rises, reaches a maximum and then falls and finally becomes insignificant: it conforms in general appearance to the more ordinary frequency curves reached in general with the Pearson type curves or with the Gram-Charlier-Edgeworth curves which latter we shall call the Charlier curves as they are most conveniently given by that writer and nominated by him types A and B. The exposed would generally be markedly skew but it might take a symmetrical form and the simplest curve of either system would therefore be the normal curve of error.

When experiences are drawn from censuses the shape of the curve representing the exposed would usually start at a maximum at birth and decrease throughout life in varying decrements. The form might be like a distorted J shaped frequency curve which can be reached with either system of curves. One of the simplest J shaped curve is the ordinary geometrical progression.

4. The first point to be made clear is that the theory does not demand that every exposed to risk, whether of the assurance or census type, can be graduated accurately by a frequency curve; it only demands that if the exposed be assumed to take such a form certain consequences will follow. Naturally, however, it is of some help in visualizing the point of view to see that the systems of frequency curves generally adopted can give roughly the appearance of the exposed to risk.

5. Now let us take first some elementary considerations:

(*a*) It seems natural to assume that the curve for E_x will not have values prior to birth, or at any rate prior to conception, *i.e.*, prior to $x = -.75$.

(*b*) If the exposed curve starts at age r then the curve $E_x p_x$ cannot strictly start before r and it can only start at a later age than r if $p_x = 0$ which it does not.

(*c*) If the exposed curve ends at age s then the curve $E_x p_x$ cannot extend beyond s and may end at any age between $s-1$ and s.

(*d*) If the exposed curve be assumed in old age to become asymptotic and to continue to infinity then we should expect $E_x p_x$ to take a similar form.

6. Although these considerations seem fairly obvious it should be borne in mind that in practice it may be necessary to relax some of them and the defence for doing so would be that, in curve fitting, range is subject to large probable error and that even if the suggested theory be sound it might only hold if we had a perfectly suitable system of frequency curves with which to work. The Pearson and Charlier systems are both of wide application but this does not mean that no one will ever be able to produce a still better system. The tables published in Biometrika (see vol. XI, pp. 328 to 513) afford evidence that we might regard the Pearson system of frequency curves as a close approximation to and not an exact reproduction of the particular kind of frequency distribution discussed in that investigation which dealt with the distribution of the correlation coefficient in small samples. If, therefore, the curves we use are approximations a relaxation of the primary considerations might bring us nearer to the statistical law. It will, however, be convenient to simplify our problem in the first place in the way we have described leaving other aspects of the question to be examined when we have reached the stage of detailed statistical evidence.

7. In our preliminary work bringing in the considerations indicated in paragraph 5 the Pearson system of frequency curves will be used. They show

the range of the curve clearly and seem suitable for our purpose. Now, bearing in mind that we wish to generalize if possible for exposed to risk whether reached from censuses or assurances, it is natural to use as our theoretical exposed to risk a curve capable of assuming either the ordinary or the J shape. The Pearson system of curves is made up of three main types, Nos. I, IV and VI, and a number of transition types of which the normal curve of error and Pearson's type III are best known and more frequently useful. The system was evolved from the hypergeometrical series and Type III can be reached in a similar manner from the binomial series as a special case of the more general solution. Type V might possibly be suitable for our present purpose. Type IV does not assume the J shape and we shall not use it.

Let us start with the Type III curve and write it as $kx^m a^x = E_x$ and assume that when $E_x p_x$ has been formed a similar form of curve will result; then bearing in mind the considerations of § 5 we reach the conclusion that

(1) $$p_x = Hx^M A^x$$

or

(2) $$\log p_x = \alpha + \beta x + \gamma \log x.$$

From (1) it appears that if E_x and $E_x p_x$ can both be represented by curves of type III and start at the same value of x then p_x takes the form of a part of another type III curve. Now we know that p_x has as its value .9 approximately at age 0, rises to a maximum at age 10 when it is nearly unity (.998), and then falls very slowly at first and finally with rapidity. A statistical examination sufficed to show that this solution of the problem by type III curves is unsatisfactory for if the rapid rise in p_x at the early ages be reproduced we cannot find a curve of the particular type which will give the high values that prevail during middle age and then decrease to produce the low values of old age. If for instance we reproduce the rates from ages 0 to 15 and the rate at age 80 we shall have values of p_x which are far too low round age 50. We must, therefore, proceed to examine the more complicated cases.

9. We may next try two type I curves and we reach

$$p_x = Hx^{n_1}(b-x)^{n_2}(b-r-x)^{n_3}$$

where r by considerations of § 5 must be not greater than unity because the range is limited by b. The simplest form is when $r=0$ and we reach $Hx^{n_1}(b-x)^N$ another type I curve but, though better than in the case of type III, it is not a satisfactory solution. Now if the exposed to risk be assumed to end definitely at age b (where age may be reckoned from $-.75$) we may be held to assume that b is the last age at which any life can be living. This means that in the equation for p_x at the beginning of this paragraph r must be equal to 1. To simplify our solution we may put $b=110.75$, say, or a rather lower value, and it is not hard to see from general considerations that n_2 and n_3 will not differ greatly in value but will have opposite signs and n_1 will be small.

Taking logarithms we have

$$\log p_n = \log H + n_1 \log x + n_2 \log (b-x) + n_3 \log (b-1-x)$$

A few trials indicated that this might give much improved results.

10. Pearson's type VI curve may be written $y = y_0 x^{q_2}(x+a)^{-q_1}$ with a range from 0 to ∞. Hence if we assume that this type of curve represents E_x and $E_x p_x$ where x is reckoned from an age of $-.75$ we have $Hx^{Q_1}(x+b)^{-Q_2}(x+a)^{Q_3}$. This is not a form which is easy to test except by fixing one curve multiplying by p_x and fitting the resulting distribution; it may therefore be left for consideration at a later date.

11. There is no necessity to assume that if a curve of a particular type be used for E_x that a curve of the same type will be given by $E_x p_x$. It would probably simplify matters if this could be proved to be a rule but it is interesting to see what might happen if the type be varied. A priori consideration and some arithmetical work show that the curves for E_x and $E_x p_x$ will be of like type unless one of the curves falls just within its type.

On the basis of the theoretical considerations of § 5 if we assume a type III or VI curve for E_x then $E_x p_x$ must be a curve of a form having a limited range in one direction only and similarly if type I be assumed for E_x we require for $E_x p_x$ a limited range in both directions so that theoretically we are restricted to type I. This may be modified in practical work for reasons already indicated keeping, however, strictly to the theoretical line of argument we may see what would happen if type III be assumed for E_x and type VI for $E_x p_x$. Then we get

$$p_x = hx^{q_2}(x+a)^{-q_1}x^{-m}c^{-x} = hx^t(x+a)^{-q_1}c^{-x}.$$

This is a type III curve divided by $(x+a)^{q_1}$ or a type VI curve divided by a geometrical progression, and the latter means that we could in order to find a graduated series of rates of mortality from census statistics assume a geometrical progression for the population, multiply it by p_x, and graduate the result $E_x p_x$ by a type VI curve—theoretically a comparatively simple procedure. It must be added, however, that such arithmetical experiment as I have been able to make at present suggests rather that a type III curve assumed for E_x would be more probably associated with a type I curve for $E_x p_x$ than with a type VI curve.

12. A rather interesting point may be indicated as showing another possible development. If the exposed to risk in a strictly stationary population takes the form of, say, a part of a type I frequency curve, or in other words if l_x can be expressed as $h(a+x)^{m_1}(b-a-x)^{m_2}$ then the force of mortality takes the form $(a_0+a_1x)/(b_0+b_1x+b_2x^2)$. Remembering that the force of mortality is equal to $-\frac{d \log l_x}{dx}$ we can by examining the differences of the logarithms of the ordinates

of frequency curves see if there is any impossibility in this result. With type I curves these differences taken negatively are large, then small, then large again and although the point calls for further study there seems therefore no impossibility in this form for the force of mortality.

13. So far as the Charlier curves are concerned it is difficult to apply the considerations set out in § 5 to reach any simple form for the values of the probabilities of life at the various ages. The explanation of this difficulty is that these curves are applied in the form of a series and the division of one of them by another does not lead to any simple expression. So far, however, as the general theory is concerned there should be no difficulty in applying these curves as a mere process of graduation but it would be advantageous to be able to use tables prepared for a large number of ordinates and to a sufficient number of decimal places.

14. We may now turn to another aspect of our problem and consider it as a method of graduation. If we find arbitrarily, or from the data, a curve to represent an actual exposed to risk and multiply the ordinates by the unadjusted probabilities of living a year we should, on the basis of the theory enunciated, reach graduated values of the probabilities by fitting a curve to the product we have obtained and dividing the ordinates by those formerly assumed for the exposed to risk. This method has been tried but it presents some practical difficulties because the two curves (for E_x and E_xp_x) are so much alike in general form that extreme delicacy in the arithmetical work is necessary to avoid incongruities. It will suffice at the present time to give a numerical example on the 0^M table in which the work was simplified by using only every fifth value of p_x from Spencer's graduation. The curves found were as follows, the unit for x being 5 years of age:

Exposed $(\text{antilog } \overline{11}.0)x^{4.243992}(23.5893-x)^{9.41792}$

with the origin at age 7.

Exposed×probability

$(\text{antilog } \overline{10}.657122)x^{4.08443}(22.18714-x)^{8.43208}$

with the origin at 7.3462.

The arithmetical results for the E_xp_x graduated and ungraduated are shown in the following table. The constants and origin show that as compared with the considerations of § 5 the graduation is distorted a little at the "tails" and especially at the youngest ages and the probabilities which are also shown bring this out and indicate the difficulties of the method as a process of graduation. It is of interest to notice that the probabilities follow the form of p_x closely and they even can give the decrease in p_x at the start which shows that the method might in an improved arithmetical form produce the whole range of probability throughout life. In the present case the rise in p_x at the beginning should, however, be either absent or much smaller.

Age	*"Ungraduated" from hypothetical curve*	*Graduated*	*Colog p_x*
12	56	53	.027680
17	695	694	.002158
22	2482	2489	.000580
27	5261	5267	.001717
32	8269	8267	.002916
37	10631	10625	.003835
42	11761	11754	.004641
47	11510	11508	.005642
52	10119	10124	.007242
57	8056	8061	.009962
62	5819	5820	.014487
67	3796	3797	.021754
72	2222	2221	.033092
77	1149	1148	.050485
82	512	513	.077010
87	193	192	.117693
92	59	57	.181301
97	12	12	.284411

15. These notes are merely tentative. Although I have made many arithmetical trials it is clear that a very much larger volume of tedious arithmetical work is required before any definite conclusions can be reached.

Possibly, however, the subject is worth developing in the hope of reaching a general method of graduation and a close approach to a mathematical representation of mortality.

SOME LIFE-TABLE APPROXIMATIONS

BY MR. G. UDNY YULE,

University Lecturer in Statistics, and Fellow of St. John's College, Cambridge, England.

The approximate relation between p_x, the probability of surviving for a year from exact age x, and m_x the central death-rate between ages x and $x+1$, which is usually employed is

(1) $$p_x = \frac{2-m_x}{2+m_x}.$$

This equation is obtained on the assumption that l_x is a linear function of the age x, or the deaths uniformly distributed over the year.

Suppose we assume instead that the force of mortality μ_x may be taken as approximately constant over the year, *i.e.*, that l_x follows a logarithmic law of decrease,

(2) $$l_{x+h} = l_x e^{-bh}.$$

Then

(3) $$m_x = \mu_x = b$$

and

(4) $$p_x = e^{-m_x}.$$

This relation is so simple and obvious that I cannot but think it must have previously been pointed out, but it was new to me and apparently to others*, and I cannot cite any other writer by whom it has been given†. The expression is useful as showing at once the effect on p of changes in m. If we double m, for example, we have (to a first approximation) to square p; if we halve m, we have to take the square root of p. Thus in English Life Table No. 7, Males, the probability of surviving from age 20 to age 30 is 0.9557; what would be the probability of surviving from age 20 to age 30 if the death-rate at each age between these limits were halved? The answer is approximately the square root of 0.9557, or 0.9776.

The approximation of (1) and (4) to each other is extremely close so long as m is small. We have:

*I mentioned the result in the course of a discussion on life-tables at the Royal Statistical Society of London in 1922: Journal Royal Stat. Soc., vol. LXXXV, p. 545.

†But not long after my return from Canada I referred to Professor Westergaard's *Die Lehre von der Mortalität und Morbilität* (it is the old edition of 1882 on my shelves) where (4) is not only given but treated as the fundamental form. My apologies are due, and are hereby tendered, for not having looked up so well known a work before.

(5)
$$(1-m/2)/(1+m/2)=(1-m/2)\left(1-\frac{m}{2}+\frac{m^2}{4}-\frac{m^3}{8}+\dots\right)$$
$$=1-m+\frac{m^2}{2}-\frac{m^3}{4}+\frac{m^4}{8}-\dots,$$

while

(6)
$$e^{-m}=1-m+\frac{m^2}{2}-\frac{m^3}{6}+\frac{m^4}{24}-\dots$$

Thus the excess of the second approximation over the first is only of the order $m^3/12$. For $m=0.01$ the difference is less than unity in the seventh place of decimals, for $m=0.1$ less than unity in the fourth place.

A good deal of attention has recently been devoted to short methods of constructing "abridged" life-tables, *i.e.*, tables in which l_x and other life-table functions are given, not for every year of age, but for every fifth year or every tenth year only, or for five-year intervals over part of the range and ten-year intervals over the remainder. In particular, Dr. E. C. Snow, in Part II of the Supplement to the 75th Annual Report of the Registrar General for England and Wales (Cmd. 1010, 1920), developed a very brief method of construction by means of empirical equations, founded on existing life-tables, relating ${}_hp_x$ the chance of surviving h years from age x, to ${}_hr_x$ the corresponding *observed* death-rate in the same age-group, which must of course be distinguished from the life-table death-rate ${}_hm_x$. Omitting the years of infancy and early childhood, there is one equation for the quinquennial period 5-10; three equations linear and quadratic, for successive intervals of the range of r, for the quinquennial age-groups from 10 to 25; and six quadratic equations, for successive intervals of the range of r, for the 10-year age-groups from 25 onwards. "A single equation might be obtained to represent the whole of the curve shown in each diagram" (the curve relating p to r) says Dr. Snow (*loc. cit.*, p. ix): "Such an equation, however, could only be made to give the degree of accuracy required by the expenditure of an amount of labour which appeared quite prohibitive, while its practical use would be doubtful".

It occurred to me, however, to ask myself the question whether an equation of the type of (1) or (4) might not hold approximately for intervals of h years instead of a single year only. On the same assumptions we would then have

(7)
$${}_hp_x=\frac{2-h\,.\,{}_hm_x}{2+h\,.\,{}_hm_x}$$

or

(8)
$${}_hp_x=e^{-h\,.\,{}_hm_x}$$

as approximate relations. If now we substitute the observed death-rate ${}_hr_x$ for the life-table death-rate ${}_hm_x$ the practical question arises whether (7) or (8) may not afford a reasonable approximation for ${}_hp_x$ or may not lead to such an approximation by some slight adjustment.

For 5-year intervals I tried the data for the life-tables (No. 8, England and Wales, 1910-12) on which Dr. Snow based his formulae. Table I shows in col. 2 the observed death-rate, in col. 3 the value of ${}_hp_x$ for the life-table (cited from Dr. Snow's table, *loc. cit.*, p. xxx), and in col. 4 the value of ${}_hp_x$ calculated from (7) or (8)—both formulae give precisely the same results for these small values of r. It will be seen that the agreement is unexpectedly close. For the age-groups 5-10 the formulae seem to give values that are slightly too low, by 7 and 9 units respectively in the fourth place of decimals. For the age-groups over 10 the error never exceeds 4 in the last place and averages 0.00015. This is not quite so close an agreement as is given by Dr. Snow's equations, but it seems to me to be all that is necessary for work of the kind. For 5-year age-groups between 10 and 40 (7) or (8) can be used with confidence, and fair certainty that the error will not exceed more than a few units in the fourth place.

Turning to the ten-year groups, I have not thought it necessary to set out in full the figures for all the groups brought into use by Dr. Snow, but give in Table II the results for the consecutive age-groups 25-35, 35-45, etc. Column (4) shows the values of p given by equation (7), to be compared with the values from the life-table in col. (3). It will be seen that in the age-group 25-35 there is good agreement, but thereafter the value given by (7) tends to be slightly too high up to the age-group 65-75, where the differences reach 0.006 and 0.005. There is then an abrupt change in the sign of the error, and for the following age-group 75-85 the value given is markedly too small. The deficiency is still more marked in the age-group 80-90, which I have not shown. It is in fact evident that when r exceeds 0.2, (7) will give a negative value, and hence it is not likely to give a good approximation for high values of r.

Column (5) gives the approximations by equation (8). Up to age 45 these are identical with those of column (4), but after that age exceed the latter more and more, so that in the group 65-75 the difference from the life-table value of p lies in the second place of decimals. But the sign of the error is here the same throughout, and this suggests that equation (8) will be the more easily modified so as to give a fair approximation over the whole range.

As the error continuously increases with r the modified equation suggested itself,

(9) $${}_hp_x = e^{-hr(1+ar)} .$$

From this we have

(10) $$a = -\frac{\Sigma\,(\log p) + h\ \log e\,\Sigma(r)}{h\ \log e\,\Sigma(r^2)} .$$

Using only the data in columns (2) and (3) of Table II, I find for the value of a 0.864, so that approximately for these data

(11) $${}_{10}p_x = e^{-10r(1+0.864r)}$$

A comparison between the life-table values of p, the values given by the above equation, and the values obtained from Dr. Snow's tables is given in Table III.

TABLE I

Life Tables for England and Wales (No. 8), 1910-12

Five-Year Periods

Age Group	Observed death-rate, ${}_5r_x$	Chance of living from beginning to end of period, ${}_5p_x$ Life-table	Equation (7) or (8)	Error
(1)	(2)	(3)	(4)	
10-15 M	.00192	.9904	.9904	0
10-15 F	.00201	.9901	.9900	+1
15-20 F	.00269	.9868	.9866	+2
15-20 M	.00288	.9861	.9857	+4
20-25 F	.00314	.9845	.9844	+1
5-10 F	.00318	.9835	.9842	−7
5-10 M	.00320	.9832	.9841	−9
25-30 F	.00367	.9819	.9818	1
20-25 M	.00372	.9814	.9816	−2
25-30 M	.00428	.9788	.9788	0
30-35 F	.00450	.9776	.9778	−2
30-35 M	.00534	.9737	.9737	0
35-40 F	.00585	.9716	.9712	4

TABLE II

Life Tables for England and Wales (No. 8), 1910-12

Ten-year Periods

Age Group	Observed death-rate	Chance of living from beginning to end of period Life-table	Equation (7)	Equation (8)
(1)	(2)	(3)	(4)	(5)
25-35 F	.00408	.9598	.9600	.9600
25-35 M	.00480	.9531	.9531	.9531
35-45 F	.00652	.9365	.9369	.9369
35-45 M	.00799	.9224	.9232	.9232
45-55 F	.01124	.8920	.8936	.8937
45-55 M	.01465	.8620	.8635	.8637
55-65 F	.02273	.7939	.7959	.7967
55-65 M	.02969	.7375	.7415	.7431
65-75 F	.0508	.5878	.5949	.6017
65-75 M	.0631	.5152	.5203	.5321
75-85 F	.1161	.277	.2654	.313
75-85 M	.1356	.221	.1919	.258

TABLE III

Age interval	Chance of living from beginning to end of period			Differences	
	Life-table	Dr. Snow	Equation (11)	(3)-(2)	(4)-(2)
(1)	(2)	(3)	(4)	(5)	(6)
25-35	.9598	.9599	.9599	+ 1	+ 1
25-35	.9531	.9530	.9529	− 1	− 2
35-45	.9365	.9365	.9365	0	0
35-45	.9224	.9223	.9227	− 1	+ 3
45-55	.8920	.8923	.8927	+ 3	+ 7
45-55	.8620	.8612	.8621	− 8	+ 1
55-65	.7939	.7929	.7931	−10	− 8
55-65	.7375	.7366	.7375	− 9	0
65-75	.5878	.5882	.5884	+ 4	+ 6
65-75	.5152	.5139	.5141	−13	−11
75-85	.277	.281	.279	+40	+20
75-85	.221	.220	.220	−10	−10
Total	..	..	..	100	69

It will be seen that the fit given by Equation (11) is exceedingly good, at least to age 75, and is in fact considerably better than the fit given by Dr. Snow's formulae. It appears to me that the latter has put himself under a disadvantage by including data for age-groups 20-30, 30-40, etc., in determining his constants, instead of limiting himself to the groups shown, which are those used in the English returns. There appear to be some almost periodic differences between the figures for groups 25-35, etc., and groups 20-30, etc., possibly due to errors in the Census numbers at single years of age which are not entirely eliminated by the process of graduation from quinquennial groups used by Mr. King. It is clear that if Dr. Snow's figures can be held to be of sufficient accuracy for the construction of such short life-tables as are desired by the public health official—and I quite concur with his view on this point—*a fortiori* the results of equation (11) are also of sufficient accuracy. A single formula has at least this advantage, that it can be easily remembered or transmitted, and is more readily available for a worker who has not got a copy of the valuable Report containing Dr. Snow's equations and tables. An equation of the form (11) can also be determined on the basis of life-tables other than those used by Dr. Snow with only a fraction of the work that is necessary by his method to obtain a whole series of equations.

The question next arises whether any similar approximations can be given to enable the expectation of life to be calculated. Dr. Snow used as the basis of his work under this head equations relating ${}_hp_x$ to a quantity he denoted by ${}_hk_x$, viz., the ratio to l_x of $l_{x+1}+\ldots+l_{x+h-1}$.

I preferred to use a slightly different quantity, namely the mean number of years lived in the interval from x to $x+h$. On any assumption as to the form of l_x as a function of the age, we have for the numbers living between x and $x+h$,

$$(12)\qquad {}_hL_x=l_x\frac{{}_hq_x}{{}_hm_x}.$$

Hence the mean number of years lived in the interval is,

(13) $$_ht_x = {}_hq_x / {}_hm_x.$$

We saw that equation (8) (reading r for m) gave without any correcting terms an adequate approximation to the value of p so far as the five-year age-groups were concerned. We might therefore hope that for these groups the equation

(14) $$_ht_x = (1 - e^{-hr})/r$$

would give a fair approximation to t. Table IV shows that this is in fact the

TABLE IV

FIVE-YEAR GROUPS

Age-group	Mean number of years lived in the interval: Life-table	Equation (14)	Difference
(1)	(2)	(3)	(4)
10-15 M	4.977	4.976	−.001
15-20 M	4.967	4.964	−.003
20-25 M	4.955	4.954	−.001
25-30 M	4.948	4.947	−.001
30-35 M	4.937	4.934	−.003

TABLE V

TEN-YEAR GROUPS

Age-group	Mean number of years lived in the interval: Life-table	q/r	Equation (15)	Difference (4)-(2)
(1)	(2)	(3)	(4)	(5)
25-35 F	9.812	9.833	9.803	−.009
25-35 M	9.781	9.803	9.773	−.008
35-45 F	9.704	9.734	9.703	−.001
35-45 M	9.642	9.673	9.641	−.001
45-55 F	9.506	9.545	9.511	+.005
45-55 M	9.368	9.411	9.376	+.008
55-65 F	9.054	9.101	9.062	+.008
55-65 M	8.781	8.842	8.799	+.018
65-75 F	8.072	8.102	8.050	−.022
65-75 M	7.644	7.701	7.644	0
75-85 F	6.137	6.212	6.136	−.001
75-85 M	5.668	5.753	5.671	+.003

case; the differences lie only in the third place of decimals. It is true that they are all of the same sign, but they are not regular, and the only correction that can be suggested to the simple result given by (14) is to add 0.002 to the result. I have only made comparison with the values for t given by the life-table for males, following Dr. Snow, but the test seems sufficient.

For the ten-year age-groups equation (8) no longer gave an adequately close approximation to p, and we had to use the modified form (11). Taking the values of q so determined, and dividing each by the corresponding observed death-rate r, the figures are obtained that are given in col. (3) of Table V, with the corresponding values for the male and female life-tables in col. (2) for comparison. It will be seen that the values of column (3) are always too high, the life-table death-rate being higher than the observed death-rate for the same age-group owing to the larger proportion of the old, and the error increases towards the bottom of the table. If the figures of col. (3) are plotted to those of col. (2) it will be seen that they are nearly a linear function of the latter. Determining the equation by the method of least squares and denoting the first approximation of col. (3) by t', I find

$$t = -.1557 + 1.0128t'. \tag{15}$$

The corrected values given by this equation are shown in col. (4) and the errors in col. (5). The agreement is, I think, very fairly satisfactory. The worst disagreements are in the lines for 55-65 M and 65-75 F. The differences between the figures of col. (3) and col. (2) here run so erratically as to make one suspect a blunder in arithmetic, but I have checked the work and failed to find any mistake. To obtain some idea of the cumulative effect of the errors, I took out the values of l_x in the original life-table (England and Wales, No. 8, Males) at 25, 35, . . . 75, multiplied each by the corresponding value of t in col. (4) and summed. The total showed 30,975,000 years of life lived between 25 and 85, against 30,965,000 by the life-table, so that the error is just less than 1 part in 3000. The life-tables for females gave a precisely similar result: 34,057,000 by the values of t from col. (4) of Table V, and 34,068,000 by the life-table. The accuracy would be ample for work of the kind considered by Dr. Snow.

It may be mentioned that if equation (14) is applied to the ten-year groups—I have not thought it worth while to give the figures—the results are consistently too *low*, the error increasing up to the line 65-75 F and then decreasing. Apart from the fact that the use of (14) would imply more calculation, this seemed to render it less likely to afford a good basis for a further approximation.

CORRELATIONS BETWEEN CLIMATIC FACTORS AND DEATH RATES*

By Professor Lowell J. Reed,
Johns Hopkins University, Baltimore, Maryland, U.S.A.

From time immemorial man has associated certain of his ailments with the disturbances in climate, this association probably arising originally from the fact that marked changes in climate interfered with his comfort. More recently he has made comparison of the seasonal waves of such climatic factors as temperature, rainfall, humidity, etc., and the corresponding waves of death rates, and from the examination of these curves he finds certain relationships seeming to indicate that climatic factors do, directly or indirectly, influence the death rate, for example, the rise in the infant death rate that occurs with the increase in temperature that comes with the summer season. If, however, these climatic factors do influence mortality rates in any way, it is reasonable to suppose that changes in the climate above and below the normal values for the season in general should be reflected in the death rates. It is the purpose of the present paper to examine this question.

For such a study a long time series is necessary, and since such a series is available for London, the data were taken from the weekly reports of the Registrar General of England and Wales for the years 1865 to 1914 inclusive. This gives a series of 50 years or 2,608 weeks. The climatic factors investigated were mean temperature, temperature range, rainfall, humidity, movement of air, and the death rates considered were total death rate, infant death rate, and the death rate from respiratory diseases. The last two rates were considered because they seem to involve people that are perhaps more susceptible to climatic changes than the population in general.

To take care of seasonal variations the material was treated by weeks, and since there has been no trend of the climatic factor, the "normal" value for any particular week of the year was taken as the arithmetic mean of the values for that week over the fifty years.

To establish the "normal" death rate for any particular week of the year it was necessary to take account of the trend of the rates, since all three of the rates considered have been declining during the period under discussion. This was done by fitting straight lines by the method of least squares to the total death rates and the respiratory death rates and second order parabolas to the

*Papers from the Department of Biometry and Vital Statistics, School of Hygiene and Public Health, Johns Hopkins University, No. 131.

infant death rates. Fig. 1 is a typical example of the normal line for the total death rate, the points being observed values of the death rate and the line representing their trend. The values given by these curves were then taken as the "normal" values, and a particular week was considered as high or low according as it fell above or below the curve.

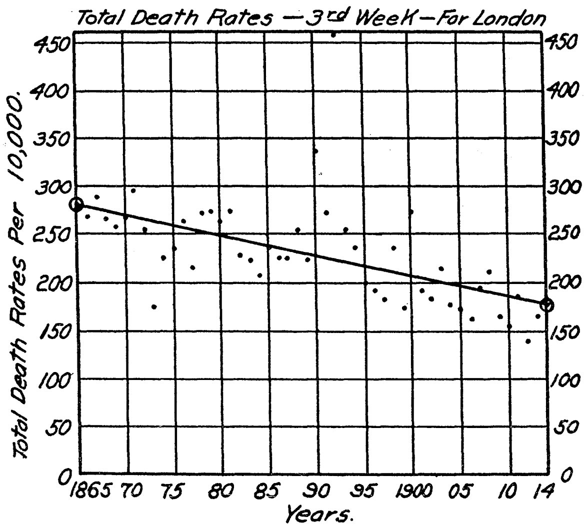


Fig. 1

The normals having been obtained deviations from them were found for all the weeks of the entire period, and the correlation question was then one of the amount of association between these deviations.

In dealing with the correlation between these values the question of lag arises, and to examine this correlation tables were formed between the climatic deviations of a particular week and the death rate of the same week, the week following, and the second week following. In general, the greatest correlation was found when the lag was one week and the present paper will be limited to that case.

Correlating the material for the entire period we find that the coefficients, although significant with respect to their probable errors, are all low. Table I shows these correlations for total and infant death rates. Probable errors for all values of this table are less than .01.

TABLE I

Correlations Between Death Rates and Climatic Factors

	Total Death Rate	Infant Death Rate
Temp. Range	+.03	+.18
Mean Temperature	−.29	+.08
Humidity	−.17	−.28
Movement of Air	−.22	−.12
Rainfall	−.17	−.12

One immediately surmises that the reason for these low correlations is that there is a seasonal change in the correlation itself, and to test this the weeks were grouped by tens and the correlations between mean temperature and the death rates were then formed. The results of these correlations are given in Table II and are shown graphically in Fig. 2.

TABLE II

CORRELATIONS BETWEEN DEATH RATES AND MEAN TEMPERATURE

Weeks	Total Death Rate	Respiratory Death Rate	Infant Death Rate
0–9	–.52	–.45	–.36
10–19	–.39	–.43	–.27
20–29	+.05	–.16	+.28
30–39	+.31	–.23	+.47
40–49	–.49	–.47	–.20
50–53	–.63	–.64	–.39

All the probable errors in this table are less than .04.

The interpretation of the curves is obvious. For the total and infant death rates deviations of the mean temperature above its normal are favourable during the winter and spring but unfavourable in the summer and then favourable again in the fall. The opposite holds for deviations of the mean temperature below its normal.

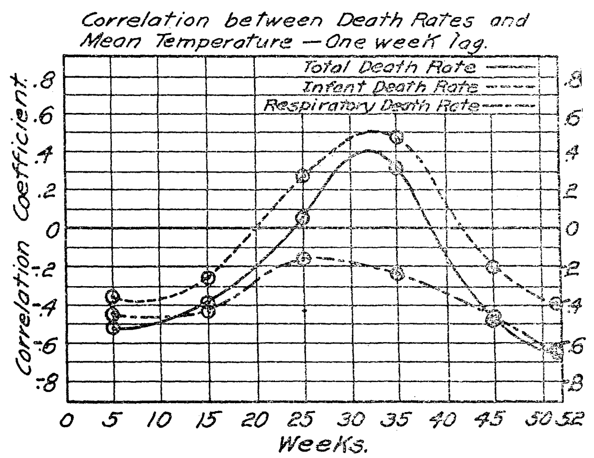


Fig. 2

For the respiratory death rate deviations of the mean temperature above normal are favourable and below normal unfavourable for all seasons of the year.

The surprising thing brought out by these correlations is the high values of the winter coefficients. One would not, I think, have expected that variation in the death rate would be as closely associated with variations of mean temperature as is indicated by coefficients ranging from –.4 to –.6.

ACTUARIAL SCIENCE IN THE FIELD OF WORKMEN'S COMPENSATION INSURANCE

BY MR. ALBERT W. WHITNEY,
Associate General Manager, National Bureau of Casualty and Surety Underwriters, New York, N.Y., U.S.A.

Workmen's Compensation Insurance in the United States is only about twelve or thirteen years old. During that time, however, a rating system has been developed which is probably more complex intrinsically than that of any other form of insurance. It has many of the elements of other rating systems and in addition some that are peculiar to itself. It involves not only the usual contingencies of life insurance but such contingencies as re-marriage and degree of dependency, it has many of the features of fire insurance such as a basic manual and a schedule for measuring physical conditions, but in addition it has features peculiar to itself, such as the experience rating plan. Apparently no other country has attempted to carry the analysis of the problem so far as the United States; although just why this should be so is not altogether clear.

The payment of Workmen's Compensation depends upon three primary elements, first the nature and severity of the injury; second, the degree of dependency in the case of the worker and third, the proportion of the wages that are allowed for the particular condition in question. This latter is something that depends upon the law and this varies in detail from state to state. The probability of physical injury, both as to frequency and severity, is on the other hand something that varies from industry to industry. The rate for Workmen's Compensation Insurance will therefore vary jointly both with the state and with the industry. For Workmen's Compensation Insurance purposes the industries are divided into about 800 different groups, and therefore the rate-making for, say, 30 states will involve the determination of 24,000 basic rates. The fixing of these basic rates is the most fundamental and most important problem of Workmen's Compensation Insurance. They are established on the basis of actual experience. In the last rate-revision the accident experience arising out of an exposure of 30 billion dollars of payroll paid was made use of. Even this enormous mass of experience would be, however, insufficient to give a dependable rate for every classification in every state.

The problem is therefore simplified by making an assumption that is justifiable as a basis for a first approximation, namely, that the relativity of hazard between industries will be independent of the state. This is not strictly true but it is not difficult to take care of the exceptions by special treatment. This assumption makes it possible to separate the problem of relativity of hazard as a function of industry from the problem of rate-level as a function of the particular state, so that these problems can be studied independently. It is possible by making use of our knowledge of rate-level by states to reduce all our accident experience to the basis of a single state so that our whole 30 billion dollars of experience can be thrown into the solution of a single problem, namely the basic relativity of hazard among these 800 classifications. For some classifications even this is not sufficient and still further expedients have to be resorted to.

For the problem of rate-level by state there are two lines of attack which can be used as corroborative and supplementary, namely the actuarial and the statistical. For the actuarial solution it is necessary to have an experience table which gives the relative frequency of accidents of varying degrees of severity. This corresponds to the mortality table in life insurance. The use of such a table involves the tacit assumption that the standard distribution of industries upon which it was based can be counted upon in all states, an assumption which is not good for more than a first approximation. Upon the basis of such a table the cost of say 100,000 accidents can be figured out under the benefits of each state. After a state has had several years' experience under a law the statistical method of determining rate-level can be applied by using the actual experience of the state. The only method available for computing rates, however, at the time that the law goes into effect is the *a priori* actuarial method and this method and the frequency table were developed to take care of this particular situation.

The problem of relativity of hazard is now well in hand and also the problem of rate-level as a function of the state, but the problem of determining rate-level as a function of the time is far from solved. In fact, this is undoubtedly the outstanding problem of Workmen's Compensation rate-making to-day.

Loss-ratios can be had that are based upon actual loss experience and that are therefore entirely dependable, but such experience to be valuable must be fully matured and this means that it must be at least two years old. Here is the difficulty: the fact that there is a two year lag in our knowledge of the proper rate-level. We must either be content to have our rates always at least two years behind the times, which is an intolerable situation both as a matter of equity and for competitive reasons, or we must devise methods of bringing this information up to date. The solution must be looked for in the establishment of a dependable correlation between rate-level and the indices that express economic and industrial conditions, for variation in general rate-level is unquestionably a function of such conditions. The relationship must be established on an empirical basis.

It is not sufficient, however, to determine the hazard for each of eight hundred classifications for each of all the states having Workmen's Compensation laws, with however great a degree of accuracy this may be accomplished; for there are still enormous differences between different individual risks in the same industry. Of two machine-shops, for instance, one may have its machinery thoroughly guarded, wide aisles, good ventilation, abundant light, first aid and hospital facilities, efficient and sympathetic supervision and a good morale among the employees; in the other many or all of these qualities may be lacking. The rate for those two shops should manifestly not be the same, and a rating system which fails to take these differences into account is seriously lacking both from the standpoint of equity and from the standpoint of encouraging accident prevention.

This situation is met in part by means of the schedule. The schedule is an instrument by which an analysis both qualitative and quantitative is made of the hazard as a function of the contributory causes, physical in nature, such as some of those referred to above. A trained inspector employed either by the insurance company or by the rating bureau makes a detailed inspection of the plant in question, noting the presence and absence of both the good and bad features referred to in the schedule, and, by an application of the schedule upon the basis of such a report, the rate for the individual risk is figured, by applying debits and credits to the basic rate for the particular classification and state as given in the manual.

The mathematical problem involved in the establishment of the schedule is an interesting one. The problem is fundamentally one of structure, and the analysis based upon the theory of probabilities develops an expression for the hazard as a function of certain parameters representing the effect of the contributory causes. The value of these parameters was established upon the basis of an analysis by cause of some 350 thousand accidents.

The schedule by itself, however, is not capable of making the fineness of analysis that is necessary. It can be used only to take account of such physical features of the risk as are revealed by inspection. But two plants may look alike and yet be quite different in morale and in their human attitude toward safety. The only way to get at such differences is through their actual accident record. It has been necessary therefore to superimpose upon the schedule another rating system called experience rating.

The mathematics of the experience rating plan consists in finding the most probable value of the hazard for the risk in question in the light of conflicting testimony. The risk belongs to a class and the experience of the class indicates a certain hazard. The experience of the risk itself is however available and in general indicates something different. Obviously if the risk is very large, so that its own experience is ample, we shall place great credence upon this experience and be inclined to neglect the experience of the class on the ground that this risk is clearly not typical of its class. If on the other hand the risk is small we shall ascribe the disparity between the indications of the risk and the indications of the class as due to chance and we shall be inclined to evaluate the hazard almost entirely upon the evidence furnished by the class experience. For risks intermediate in size the experience rating plan gives a consistent system for making an estimate of the degree to which the indications of the individual risk are on the one hand trustworthy and on the other hand due to chance. The mathematical structure of the experience rating plan appears to be correct but it has not up to the present time been found possible to determine values of the parameters statistically, so that the plan from a quantitative point of view is not as perfect as could be desired.

It should be noted in passing that the effect of refinements in Workmen's Compensation rating, and particularly schedule rating and experience rating, is not merely to produce greater equity and less competitive abuse but to offer inducements for accident prevention, for any plant that is put into good physical condition and that has a good accident record can get direct advantage from its efforts in a lowered rate and on the other hand the bad risk is in a similar way penalized. This ever-acting pecuniary inducement toward safety is a powerful influence in the saving of life and limb.

A word should be said about the administration of this rating system. The insurance carriers, with the exception of the monopolistic State Funds have united in the creation and support of an organization called the National Council on Compensation Insurance. This makes the rates for all the States except those having monopolistic State Funds. These rates are wholly advisory and it rests with the state authorities or the companies themselves as the case may be in the state in question to accept them and apply them; they are, however, generally applied. The National Council on Compensation Insurance was voluntarily placed under the supervision of the National Convention of Insurance Commissioners and a representative of that body has his permanent office in the National Council and occupies a strategic position on the more important committees.

PREMIÈRES RECHERCHES SUR LA FÉCONDABILITÉ DE LA FEMME

PAR M. CORRADO GINI,
Professeur à l'Université de Rome, Rome, Italie.

J'appelle *fécondabilité de la femme* la probabilité que la femme mariée soit fécondée dans le mois, abstraction faite de toute pratique malthusienne ou néo-malthusienne destinée à limiter la procréation.

Il est inutile d'insister sur l'importance théorique et pratique qu'aurait la mesure de la fécondabilité. Nous aurions en particulier le moyen de décider quelle partie des différences que l'on observe entre les coefficients de natalité de différents pays, de différentes classes sociales, de différentes époques provient de causes physiologiques et quelle partie provient de causes volontaires.

Il est inutile aussi de s'arrêter à démontrer que l'on ne peut pas mesurer directement la fécondabilité. On comprend aisément en effet que l'on n'est pas en état de mesurer les conséquences des pratiques malthusiennes et néo-malthusiennes sur la natalité. Nous ne connaissons pas d'ailleurs la fréquence des avortements ovulaires qui ont lieu dans les premiers mois de la grossesse.

On peut, au contraire, songer à une mesure indirecte de la fécondabilité, basée sur les considérations qui suivent.

Supposons que les femmes mariées, disponibles pour la procréation, soient, pendant toute la période considérée, au nombre de n et aient toutes la même fécondabilité p. Dans ces hypothèses, le nombre des femmes qui, après le mariage, seraient fécondées pour la première fois dans le premier mois, sera pn; le nombre des femmes qui, après le mariage, seraient fécondées pour la première fois dans le deuxième mois, sera $(1-p).pn$; le nombre des femmes qui, après le mariage, seraient fécondées pour la première fois dans le troisième mois, sera $(1-p)^2.pn$, et ainsi de suite. Or, les nombres que l'on obtient de la sorte constituent une progression géométrique, dont la raison est $1-p$.

En supposant que le pourcentage des avortements (ou des avortements et des morts-nés) et de même le pourcentage des grossesses plus longues ou moins longues que 9 mois soient les mêmes pour les produits des conceptions qui se sont effectuées dans les mois successifs du mariage, on peut substituer, au rapport entre les premières conceptions qui se sont effectuées dans le mois $x+1$ du mariage et les premières conceptions qui se sont effectuées dans le mois précédent x, le rapport entre les premiers nés (ou respectivement les premiers nés vivants) venus au monde dans le mois $x+10$ du mariage, et les premiers nés (ou respectivement les premiers nés vivants) venus au monde dans le mois $x+9$.

Il est évident que la valeur de p, à laquelle on parvient par cette voie, est indépendante du nombre plus ou moins grand des femmes mariées qui ne sont pas disponibles pour la procréation, soit à cause de la stérilité du mariage, soit à cause des pratiques malthusiennes ou néo-malthusiennes.

Il est évident aussi qu'elle est indépendante de la fréquence des avortements et des morts-nés.

La valeur de p ainsi déterminée mesurerait donc, dans les hypothèses sus-mentionnées, la fécondabilité des femmes fécondables.

Parmi ces hypothèses, il y en a pourtant une qui, évidemment, ne répond pas à la réalité: c'est l'hypothèse que toutes les femmes mariées fécondables aient la même fécondabilité p. Cette circonstance, dont on ne peut pas faire abstraction, n'empêche pas de parvenir à la mesure de la fécondabilité; la voie à suivre est pourtant un peu plus longue.

Soit $s \leq n$ le nombre des groupes dans lesquels les n femmes mariées peuvent être classées d'après leur fécondabilité, p_i $(i=1, 2, \ldots, s)$ étant la fécondabilité des femmes du groupe i et n_{ix} le nombre des femmes du groupe i qui, dans le mois x après le mariage restent encore à féconder.

La valeur probable du nombre des femmes qui, après le mariage, sont fécondées pour la première fois dans le mois x, sera $\sum_1^s {}_i p_i n_{ix}$; et la valeur probable du nombre des femmes qui, après le mariage, sont fécondées pour la première fois dans le mois suivant $x+1$ sera $\sum_1^s {}_i (1-p_i) p_i n_{ix}$. La valeur probable du rapport entre le second et le premier nombre (si l'on convient de donner à chacune des valeurs possibles du rapport un poids proportionnel à la probabilité que cette valeur se réalise et au nombre qui en constitue le dénominateur) sera:

$$\frac{\sum_1^s {}_i (1-p_i) p_i n_{ix}}{\sum_1^s {}_i p_i n_{ix}} = 1 - \frac{\sum_1^s {}_i p_i^2 n_{ix}}{\sum_1^s {}_i p_i n_{ix}}.$$

Mais $\frac{\sum_1^s {}_i p_i^2 n_{ix}}{\sum_1^s {}_i p_i n_{ix}}$ est précisément la valeur probable de la fécondabilité moyenne des femmes qui, après le mariage, sont fécondées pour la première fois dans le mois x.

Nous pourrons donc déduire la fécondabilité moyenne des femmes qui, après le mariage, sont fécondées pour la première fois dans le mois x, du complément à l'unité du rapport entre les femmes qui, après le mariage, sont fécondées pour la première fois dans le mois $x+1$ et les femmes qui, après le mariage, sont fécondées pour la première fois dans le mois x.

Ce calcul suppose justement les hypothèses suivantes, que l'on peut admettre sans difficulté:

(*a*) le nombre des femmes qui, dans le mois x étaient fécondables pour la première fois, après le mariage, et n'ont pas été fécondées dans ce mois, est égal au nombre des femmes qui, dans le mois $x+1$ sont fécondables pour la première fois après le mariage. Cela revient à dire que l'on peut négliger les effets, d'un mois au mois suivant, de la mortalité, des migrations, des passages de la catégorie des femmes mariées non fécondables à la catégorie des femmes fécondables pour la première fois après le mariage (par exemple par cessation des pratiques limitatrices de la procréation, ou par élimination de la cause de la stérilité), ainsi que des passages inverses de la deuxième catégorie à la première (par exemple par effet de stérilité survenue ou d'adoption de pratiques limitatrices);

(*b*) dans chacun des s groupes des femmes fécondables pour la première fois après le mariage, la fécondabilité n'a pas varié du mois x au mois $x+1$.

En supposant en outre (hypothèse *c*) que la probabilité des avortements (ou des avortements et des morts-nés) ainsi que la probabilité d'une durée de la gestation supérieure ou inférieure à 9 mois, soient les mêmes pour les produits conçus dans le mois x et pour ceux conçus dans le mois $x+1$, on pourra substituer, au rapport entre les femmes qui, après le mariage, sont fécondées pour la première fois dans le mois $x+1$ et les femmes qui, après le mariage, sont fécondées pour la première fois dans le mois x, le rapport entre les premiers nés (ou respectivement les premiers nés vivants) du mariage qui sont venus au monde dans le mois $x+10$ et les premiers nés (ou respectivement les premiers nés vivants) du mariage qui sont venus au monde dans le mois $x+9$.

On pourra par conséquent obtenir la mesure de la fécondabilité moyenne des primipares qui, après le mariage, ont conçu pour la première fois dans le mois x. Cette fécondabilité, à son tour, pourra être regardée comme sensiblement égale à la fécondabilité des primipares qui ont eu leur enfant dans le mois $x+9$ du mariage.

Le complément à l'unité du rapport entre les premiers nés (ou les premiers nés vivants) du mariage venus au monde dans le mois $x+10$ et les premiers nés (ou respectivement les premiers nés vivants) du mariage venus au monde dans le mois $x+9$, nous donnera, par conséquent, dans les hypothèses (*a*), (*b*), (*c*) une mesure approximative de la fécondabilité moyenne des primipares qui ont eu leur enfant dans le mois $x+9$ du mariage.

Si on calcule la moyenne des rapports entre les premiers nés des mois du mariage $x+10$ et $x+9$, $x+11$ et $x+10$ ...., $x+y+10$, et $x+y+9$ en donnant à chaque rapport un poids proportionnel au dénominateur, et si l'on prend le complément de cette moyenne à l'unité, on obtiendra, dans les hypothèses (*a*), (*b*), (*c*), une mesure approximative de la fécondabilité moyenne des primipares qui ont eu leur enfant dans les mois du mariage de $x+9$ à $x+y+9$.

Voici les résultats obtenus pour quelques États. Ils se basent sur des données publiées par les statistiques officielles, sauf pour l'Italie, pour laquelle les données sont le fruit de recherches spéciales établies sur les registres d'état-civil de 24 communes:

Fécondabilité moyenne des primipares qui ont eu leur enfant pendant les mois suivants du mariage:

Pays	Années	10-17	11-17	10-23	11-23
Berlin	1894-1895	26.1	24.4	23.0	20.5
Confédération australienne	1917-1921	19.4	20.6	18.5	19.2
Australie occidentale	1895-1915	23.9	22.7	21.9	20.2
Nouv. Galles du Sud	1893-1905 et 1916-1921	20.6	22.2	18.9	19.7
Victoria Tasmanie	1898-1900 1905-1906	21.0	24.4	....	....
Italie (24 communes)	1900-1921	23.8	21.0	21.4	18.6

L'uniformité des résultats obtenus est remarquable, surtout lorsqu'on exclut du calcul les premiers nés venus au monde pendant le 10^{e} mois du mariage, correspondant aux conceptions du premier mois du mariage. Pour le rapport entre les conceptions du 1er et celles du 2^{e} mois du mariage, on ne peut pas en effet admettre la validité de l'hypothèse (a), la fécondabilité étant, dans le premier mois, moindre que dans le 2^{e}, à cause de la fréquente virginité de l'épouse, ni celle de l'hypothèse (c), les avortements ovulaires étant spécialement fréquents pour les produits conçus dans le premier mois à cause des voyages de noces et peut-être d'autres circonstances.

L'obstacle à la fécondation représenté par la virginité augmente naturellement d'importance avec l'âge de l'épouse, mais, le premier mois passé, il ne paraît pas que la fécondabilité pour les épouses âgées soit plus faible que pour les plus jeunes. C'est là un résultat remarquable qui est mis en lumière par les données de la Confédération australienne pour la période 1907-1914, ainsi que par celles des Nouvelles-Galles du Sud pour la période 1893-1898.

Un autre résultat important est que la diminution de la natalité qui s'est vérifiée de 1901/2 à 1911/12 dans le Royaume de Saxe, et qui a atteint la mesure de 50%, ne parait pas avoir été accompagnée par une diminution de la fécondabilité des primipares.

Les méthodes et les résultats résumés dans cette note sont exposés avec des détails complémentaires dans deux mémoires présentés à l'«Istituto Veneto di Scienze, Lettere ed Arti» le 1er juillet 1923 et le 13 juillet 1924, et dans l'article *Decline in the Birth-rate and 'Fecundibility' of Women*, in The Eugenics Review, January, 1926. Des recherches ultérieures sont en cours qui pourront peut-être permettre de recueillir d'autres résultats intéressants dans le nouveau domaine que la méthode proposée paraît ouvrir à la recherche statistique.

JOTTINGS FROM THE CANADIAN CENSUS.

By MR. R. H. COATS, *Dominion Statistician*, and MR. M. C. MACLEAN, *Dominion Bureau of Statistics, Ottawa, Canada.*

The Census is the largest of statistical jobs: it is even the largest of ordinary administrative jobs in the average country under average conditions*. It deals primarily with the human factor, the most elusive of all, but at the same time the most important, as forming the background of all other factors. The technic of the Census is correspondingly difficult and tremendous in its consequences. These "jottings"—desultory as their name implies—are not offered primarily for their intrinsic importance, but as lending "local colour" to the present occasion, of which the Dominion Bureau of Statistics desires to express its meed of appreciation.

HISTORICAL

To Canada, we are fond of telling the world, belongs the credit of taking the first recorded Census of modern times—a veritable "nominal" enumeration of the population of New France, on the *de jure* system, covering sex, age, occupation, and conjugal and family condition†. The year was 1666. It was not a very large affair—only 3,215 souls. This was repeated no less than fifteen times in the next hundred years. After the British Conquest in 1763, the frequency of the Census at first diminished, but it rose sharply half a century later, and for eighteen years we had actually an annual census in Upper Canada, with fairly recurrent enumerations in the other parts as well. It was not, however, until 1848 that a regular Census was instituted by legislation in the Canadas. We can reckon ourselves today as having a past fairly replete with Censuses, though we have had only nine consecutive decennial censuses of the whole of Canada on a regular plan.

THE CENSUS AND PARLIAMENTARY REPRESENTATION

The legal *raison d'être* of the Canadian Census is to determine the representation of the several provinces in Parliament. Under the British North

*"Of all the peace-time activities of the Federal Government taking and compiling the Census is the largest."—Report of Advisory Committee on the U.S. Census, 1921.

†Unless the credit belongs to a Census of Castille mentioned by Edgeworth in the Dictionary of Political Economy (article *Census*), as "tolerably complete" and as having been taken in 1593. The Swedish Census commonly mentioned as the first was not taken until 1748.

America Act, the Province of Quebec is given a fixed number of members in the House of Commons, namely, 65*, the other provinces being assigned representation *pro rata* according to the Census,—with certain important checks and balances. One or two interesting points arise in this connection:

(1) The British North America Act provides (Section 51, subsection 3) that "in the computation of the number of members for a Province a fractional part not exceeding one half of the whole number requisite to entitle the province to a member shall be disregarded, but a fractional part exceeding one half of that number shall be equivalent to the whole number". This, with the provision already quoted, frees us at once from the problem of fractional representation that has so vexed the statisticians of the United States. But it does not free us from the responsibilities always attaching to the use of a standard. Errors, as we know, should not be permitted to occur in statistics. Perhaps therefore it will be helpful to point out, as a further spur to effort on the part of the Census, (and, may we add, as a further plea to the public for their interest and cooperation, often so nonchalantly given, though so absolutely essential to success in this exceedingly difficult task), how necessary it is to achieve the extreme of accuracy, especially at the governing points along the line. The practical issue may be stated as follows, the formulae in each case being of the most obvious kind, yet of value as putting in definite form what is frequently overlooked:

(*a*) What would be the effect on representation of an error, either of under-statement or of over-statement, in the Census of Quebec, the standard province?

Let s represent the true population of the standard province, p that of any other province or all provinces, e the population included in the supposed error, and r the number of representative units resulting from that error.

Then

$$\frac{65p}{s-e} - \frac{65p}{s} = r.$$

Simplifying and transposing:

$$e = \frac{rs^2}{65p+rs}.$$

Applying this to the actual population figures of Quebec and the other provinces as at the last census, the error which, if deducted from or added to the standard province, would give one representative unit too many or too few to the other provinces, will be found to be 13,270.

Now as one-half a representative unit may mean one representative, it follows that an error of 6,635 in taking the Census of Quebec might under present population conditions in Canada result in giving a member too many or too few to the rest of Canada.

*The number of seats held by Lower Canada in the Canadian Legislature prior to Confederation.

(*b*) The effect on representation of an error in the enumeration of another province than Quebec is, of course, the fraction that the error is of the unit of representation $\left(\frac{65e}{s} = r, \text{ or } e = \frac{sr}{65}\right)$. The unit of representation by the recent Census was 36,283. Half of this, or 18,142, would on the average affect representation by one member.

Thus under present population conditions in Canada an error in the Census of Quebec would be attended with nearly three times as serious consequences to representation in Parliament as a similar error elsewhere.

(*c*) As a refinement we might ask, would an error in the standard province, Quebec, be equal in its incidence upon each of the other Provinces? Obviously the answer is in the negative: the effect will be greater in the provinces that are smaller than Quebec than in Ontario, (the one province which is larger), for the reason that the error will be a greater proportion of the former than of the latter $\left(e_1 \times \frac{p}{s - e_1} = e\right)$. Thus if the Census of Quebec should contain an error in defect (which would have the result of increasing representation outside of Quebec), the increase in representation would tend to be greater in Ontario than in the smaller provinces. On the other hand, if the Census of Quebec should be in error through excess (which would decrease representation in the rest of Canada), the smaller provinces would tend to lose in point of representation less than would Ontario.

(2) As illustrating how the statistician must "watch his step" in working out results under such heavy responsibilities, we may adduce the following: One of the subsections of the British North America Act (subsection 4 of section 51), in addition to that already cited, provides that on the readjustment of representation "the number of members for a Province shall not be reduced unless the proportion which the number of the population of the Province bore to the number of the aggregate population of Canada at the then last preceding re-adjustment of the number of members for the Province is ascertained at the then latest Census to be diminished by one-twentieth part or upwards". The computation arising out of this provision must be so close that even an exceedingly small difference, such as that arising from the use of a decimal to five points instead of a vulgar fraction, must be guarded against, as it might easily result in the displacement of a representative. Ontario, for example, according to the last Census, fell off in the proportion which its population bore to the total population of Canada by a ratio of decrease very close to the one-twentieth prescribed in the Act, namely, .0481. The calculations by decimals to five places and by vulgar fractions show ratios of proportional decrease amounting to .048162 and .048138 respectively, a difference of .000024. Suppose the exact ratio had been .049976: an error of .000024 would have brought this up to .05 thus causing the loss of a representative. By carrying the decimal computation to more places the error would, of course, be reduced, but it would persist.

THE ANALYSIS OF A CENSUS—SAMPLE TESTS AND INTERPRETATIONS

The general subject of Census analysis is too vast to be approached here, save to mention that our Bureau contemplates a series of studies on special features which though far from comprehensive will represent considerably more than has been attempted hitherto in Canada. Two or three random exemplifications, however, of the importance, nay the necessity, of the most careful examination of Census data may be given, choosing for opening text the item on illiteracy, with special reference to the question of the reliability of the Census.

(1) *The Accuracy of the Census of Illiteracy.*

The Census returns on illiteracy have been under fire in many countries, notably in the United States, and undoubtedly the difficulties connected with the item are many. In the first place, the question "Can you read?" is indefinite: *e.g.*, some might answer "Yes" of a few simple words of print, others "No" having in mind an abstruse work or poorly written script. There is the further possibility of unwillingness to acknowledge illiteracy. Again, the enumerator may rely too much upon impression, a non-English-speaking person of poor appearance being prejudged as illiterate, and an English-speaking person of good appearance as literate. The U.S. Census Bureau concludes that while the illiteracy figures may be fairly accurate for the country as a whole and for States, (by the intercancellation of errors), they do not hold for smaller districts. The U.S. Census Advisory Committee suggests, for the resolution of the doubt, a follow-up survey of certain districts, say, in the neighborhood of a University by students of statistics. We have arrived at a somewhat different conclusion for Canada.

The returns of illiteracy in Canada in 1921 have been compiled for Canadian-, British-, and foreign-born over ten years of age, for the rural and urban parts separately of each of the 219 counties or other large municipal areas into which the provinces are divided. The number of children between the ages of 7 and 14 at school and not at school in the same counties has also been compiled, whilst the provincial education reports give the school grades of pupils in 1921 for 150 out of the 219 counties.

These different sets of figures offer materials for some interesting correlations. Taking first the school attendance figures and the illiteracy returns (Canadian-born), it may be noted that the former are for ages 7 to 14, whilst the latter are for ages 10 and over, the only element common to both being children of ages 10 to 14 years, a negligible factor, as they number in illiterates only 16,920 for the whole of Canada, *i.e.*, 7 per cent. of the total Canadian-born illiterates. When we correlate school attendance with illiteracy, therefore, in the different counties we are avoiding the pitfall of identity in the data. It seems reasonable to add that a correlation between school attendance and illiteracy under these circumstances, and with a sufficient number of cases to guard against chance, is meaningless except on the assumption that the illiteracy returns are accurate.

The illiteracy returns of 1921 were correlated with school attendance by the Pearsonian product moments and found to be .70, a rather high correlation considering the number of cases and the mutual exclusiveness of the elements.

Care was taken to correlate homogeneous elements, and the rural and urban parts were correlated separately, the coefficients coming out almost exactly the same for both, *i.e.*, around .70

The illiteracy returns and the grades at school were next correlated. It was possible to do this only in the case of four provinces, and as "grade" may have a slightly different meaning in different provinces it was deemed desirable to correlate each province separately. This unfortunately made the number of cases in two of the provinces rather small, namely, 19 and 15 counties, respectively, but the remaining provinces had respectively 66 and 50 counties. The percentage of illiteracy of the population over 10 years of age was correlated with the percentage of children at school below Grade IV and also below Grade V in 1921, and the coefficients found to be high. The two large provinces showed coefficients of .60 and .80 respectively. In the province with only 15 counties it was .98. The province with 19 counties showed a definite but not very high correlation for that particular year, but when the percentage below Grade IV was taken of the children at school according to the census returns, the correlation was .79. There is no doubt that the correlations would have been still better if the average grades for the ten previous years instead of for 1921 had been taken, but this would have involved an excessive amount of labour.

Now there is no conceivable reason for a correlation between the illiteracy of a community and the grade of the children at school, except on the assumption that both sets of figures are accurate; as in the case of school attendance, inaccurate figures would tend to destroy the correlation. Further, it is reasonable to conclude that the illiteracy figures have not only a consistent but also a definite meaning. The grade at school is a definite measurement of educational status; of children who have reached a certain grade it is known not only that they can read but also how much they can read—a point that the most skilfully devised census question on illiteracy could not ascertain. There is a strong probability, however, that in an illiterate community (*i.e.*, illiterate as shown by the census), (1) the school attendance of the children will be lower than in a literate community; and (2) the children actually at school will not reach the same standard before leaving school as the children in a literate community. In other words, the illiteracy figures of the Census as shown by this correlation are an index not only of the inability to read but of the degree of educational status likely to prevail in a community.

It may be mentioned as further tests of the reliability of the census illiteracy figures, (1) that a curve showing the proportion of illiterates in each percentage of the population, beginning with the worst percentage and proceeding in order to the best, is remarkably smooth; (2) that the standard deviation between different communities is largest of all (relatively to the mean) in the case of the British-born, second largest in the case of the Canadian-born, and least in the case of the foreign-born. The British-born show the least illiteracy in the Census, next the Canadian-born, and next the foreign-born. The British-born are comparatively homogeneous as to race; the Canadian-born are of widely different racial origins, including Indians; whilst the foreign-born represent a still larger admixture of origins. If the Census enumerator had prejudged illiteracy upon impression, assuming that the person of his own race and language was able to read and that the person who could not understand or speak that language was unable to read, then it is apparent that the foreign-born, varying all the way from the English-speaking American to the Ukrainian, Ruthenian and Italian, would show a greater variability in illiteracy than the Canadian-born, and the latter again than the British-born. But the facts are the opposite, though on the assumption of their accuracy quite intelligible. In a large number of communities, each with only a small number of British-born, the illiteracy of the latter was nil; in other communities, where the British-born were preponderatingly recent arrivals in Canada (so that there were not many old persons among them) their percentage of illiteracy was again very small and in correspondence with the percentage of illiteracy of the English-speaking Canadian-born; whilst in still other communities where the occupations were mining or fishing and where the British-born contained a large element of unskilled laborers and old persons their percentage of illiteracy was large. In other words, the illiteracy of British born communities varied all the way from zero to a relatively high percentage. In the case of the Canadian-born, on the other hand, English-speaking communities had a considerable proportion of old people so that there were no communities with zero illiteracy. In the case of the foreign-born, although certain races show an educational status practically as high as that of English-speaking Canadians, these had as neighbors in the community an illiterate race which brought down the general educational status; consequently the illiteracy of the foreign-born tended more towards a dead level as between communities than that of the Canadian-born, as did the illiteracy of the latter when compared with that of the British-born.

If the above conclusions hold, the census returns of illiteracy are of greater value than has hitherto been claimed for them.

(2) *Weighting.*

False weighting is a fruitful source of error in the interpretation of statistics. For example, if it is desired to compare two observations, A and B, and A is found in the form W_1A, whilst B is found in the form W_2B, the public frequently base their conclusions on what are really the weights and not the observations.

Again using the illiteracy figures for illustration: In 1921, as in former censuses, the percentage of illiteracy of females was considerably lower than that of males, namely, 4.4 per cent., compared with 5.7 per cent. A hasty conclusion is that females as a sex have a higher educational status than males, the explanation being sought in the fact that the former attend school more regularly and longer.

But is this really a phenomenon of sex? Let us consider the "weights" involved, namely, (1) that race is a potent factor in illiteracy, the most illiterate elements in the population being amongst the foreign-born; and (2) that illiteracy is greater in rural than in urban parts.

Now it so happens that in our sparsely settled districts and in rural parts generally the males are in the majority, also in the illiterate races, whilst in the urban centres and in the less illiterate races the proportion of females is much greater. That is, under the conditions that chiefly accompany illiteracy, the males are in the majority, whilst they are either in the minority or about equal to the females where the conditions make for literacy. In comparing the illiteracy of males and females, therefore, it is not the sexes but these other conditions that are actually under interpretation. Taking the unweighted average of the percentages of illiteracy of males and females under all the conditions mentioned, or giving them equal weights, it will be found that the difference between the sexes in this respect is unimportant.

A striking illustration of the same fallacy was encountered in studying illiteracy in the different counties. In county A the illiteracy in urban centres was 5.41; in rural centres 3.37. In county B the illiteracy in urban centres was 5.70; in rural centres 3.39. In county A as a whole the illiteracy was 4.77; in county B as a whole, 4.49. In other words, the rural person in A was less illiterate than the rural person in B, and the urban person in A was also less illiterate than the urban person in B; and yet A on the whole was more illiterate than B. What might almost be considered a mathematical paradox was here caused by the different proportions of population in rural and urban centres in the two counties. The question is, which county was the more illiterate, A or B? We think B.

The following table shows the percentages of illiteracy in the different provinces of Canada in two ways: 1. the actual Census figures; 2. as the figures would appear if each province had the same proportions of Canadian, British and foreign-born, also of rural and urban populations, and also of males or females, as the whole of Canada. It should be noticed that the only re-adjustment made here is in the weights—the percentages of illiteracy of Canadian-born, etc., are left untouched:

	P.C. of Population over 10 Years Illiterate According to the Census	P.C. Illiterate if the Population were Distributed by Nativity, Rural and Urban Residence and Sex in each Province in the same Proportion as in all Canada
P. E. Island	3.1	3.0
Nova Scotia	5.1	5.1
New Brunswick	7.6	6.1
Quebec	6.2	6.0
Ontario	2.9	3.7
Manitoba	7.1	5.2
Saskatchewan	5.9	3.6
Alberta	5.2	4.4
British Columbia	6.2	6.9

(3) *Age Distribution, Montreal and Toronto.*

Age distributions differ widely in communities with varying characteristics and standards of life. Thus the population of Quebec is a much "younger" population than that of Ontario. On the basis of the 1921 Census, the median inhabitant in respect of age in Quebec was only 20.79 years of age, while the median person in Ontario was 26.76 years of age—nearly 6 years older.

This phenomenon may be further illustrated by a comparison of the populations of the two leading cities, Montreal and Toronto. Both of these moved up during the past decade beyond the half-million mark—the first time Canada witnessed such a phenomenon—itself significant of much. Montreal had in 1921 a total population of 618,506, of whom 614,604 were of known ages; Toronto had a population of 521,893, of whom 520,991 were of known ages. The difference between the populations of known age was 93,613 in favour of Montreal, which would naturally be presumed sufficient to give Montreal a surplus in all the quinquennial age-groups.

We find, however, on examination, that practically the whole excess of the Montreal population over that of Toronto is in the quinquennial age groups under 30—the population over 30 years of age (of known ages) being 242,082 in Montreal as compared with 241,188 in Toronto. Indeed if we consider age 35 as the middle point of human life—*mezzo del cammin di nostra vita*—the persons of known age above this point numbered 192,239 in Toronto as compared with 191,391 in Montreal. It is in the age groups under 15 that two-thirds of Montreal's excess is made up. The question is far from having a merely theoretical interest; it has a profound influence upon the social and economic life of the two cities. The student of statistics would expect the following phenomena to be present in Montreal as compared with Toronto, (1) higher birth rate, accounting for the greater number of young people; (2) higher general death rate, as shown by the lower percentage of people at the more advanced ages; (3) higher infant mortality; (4) proportionally larger educational problem; (5) earlier school leaving and more child labour; (6) lower general standard of living, the percentage of non-workers to total population being larger. These deductions are in the main borne out by independent statistics as follows: (1) the crude birth rate in Montreal in 1921 was 34.2 as against 25.6 in Toronto; as for (2) the general death rate in Montreal in 1921 was 16.6 as compared with 11.3 in Toronto; (3) the infant mortality was in 1921, 158 per 1,000 living births in Montreal as compared with 91 in Toronto; (4) records from the census of 1921 are not yet available, but in 1911 there were 2,911 child workers under 15 years of age enumerated in Montreal as compared with 1,875 in Toronto; (5) hourly wages and hours worked per week in leading trades show on the average smaller earnings and longer working hours in Montreal*. Again, the youthful character of Montreal's population means, generally speaking, that a smaller percentage of the population are at ages where they can be gainfully employed. Thus, out of Montreal's population of 470,480 in 1911, 184,257 or 39.2 per cent., were returned as gainfully employed; out of Toronto's population of 376,538, 169,520 or 45.0 per cent., were returned as gainfully employed. Necessarily it must make a difference to the well-being of a community whether 39 out of every 100 or 45 out of every 100 of its population are producers. On the other hand, from a far-sighted point of view, it may in the long run prove the better

*See *Canada Year Book*, 1922-23, pp. 737-8.

economic policy to invest capital in the rearing of numerous children rather than in things material. As Ruskin has said, "there is no wealth but life".

The table of the age distribution of the 1921 populations of Montreal and Toronto, by quinquennial age groups follows:

POPULATION OF THE CITIES OF MONTREAL AND TORONTO BY AGE GROUPS, CENSUS, 1921

Age Groups	Population of Montreal	Population of Toronto	Excess of Montreal over Toronto	Excess of Toronto over Montreal
Under 5 years	69,607	46,933	22,674	..
5-9 years	70,166	49,867	20,299	..
10-14 "	62,513	42,957	19,556	..
15-19 "	58,817	41,269	17,548	..
20-24 "	57,271	47,137	10,134	..
25-29 "	54,148	51,640	2,508	..
30-34 "	50,691	48,949	1,742	..
35-39 "	45,845	47,394	...	1,549
40-44 "	38,009	37,826	183	..
45-49 "	29,873	29,549	324	..
50-54 "	24,082	24,819	..	737
55-59 "	17,578	17,505	73	..
60-64 "	14,318	14,664	..	346
65-69 "	9,537	9,023	514	..
70-74 "	6,395	5,873	522	..
75-79 "	3,254	3,149	105	..
80-84 "	1,649	1,630	19	..
85-89 "	635	640	..	5
90-94 "	170	135	35	..
95-99 "	39	28	11	..
100 and over	7	4	3	..
Total population of known ages	614,604	520,991	96,250	2,637
Age not given	3,902	902	3,000	
Grand Total Populations	618,506	521,893	96,613	

Postcensal and Intercensal Estimates—The Future

But of all topics suggested by the Census, the greatest is that of population aggregates and the place of Canada in that sudden and enormous expansion of the human family which made the nineteenth century the "magnificent episode" of Mr. Keynes's phrase—with of course, the revulsion of the present war-stricken day. For six centuries before 1600 the growth of Europe was negligible. In the seventeenth century the rate advanced to, say, $33\frac{1}{3}$ per cent., and in the eighteenth century to, say, 50 per cent. In the nineteenth, this slow fire burst into a marching conflagration, and we had an increase in one hundred years of over 300 per cent.

Canada grew into existence whilst the rate was still moderate, though she was a symptom of the new expansion from the beginning. In 1791, the year of our Constitutional Act, which marks our first non-French immigration,—the United Empire Loyalists—we had reached about 200,000. (The first census of the United States taken the year previously gave 3,929,214.) This, it may be noted, was seven years earlier than Malthus, and before the Industrial Revolution. We began the nineteenth century with a quarter of a million, and long before 1840 had increased this to a million and a half. Our first regular census in 1851 showed 2,384,919, a total which had grown to 3,689,257 in 1871, the first Census after Confederation. Not to prolong the recital, we ended the nineteenth century with 5,371,315. If the world increased by three times during the nineteenth century, Canada increased by over twenty times; she grew her three times in the first twenty-five years, and added another three times on the top of that before 1850; reckoning backwards, she grew three times since 1840 alone.

It has been in the twentieth century, however, that the most spectacular expansion of the Canadian population has taken place, the underlying factor being of course the opening to settlement of the "last best West", which became the chief magnet for the attraction of old-world peoples. With an accompanying importation of probably $2½ billions of capital to finance the large constructive undertakings required (railway, municipal and industrial), 1900-1910 became the *decas mirabilis* of Canadian progress. The growth in population was 35 per cent., the highest of any country in the world. Again, in 1911-1921, in spite of the war, our increase was nearly 22 per cent., once more the highest of any country except Australia and that by only a fractional amount. We seem, in fact, during the twentieth century thus far, to have approximated to the experience of the United States throughout the nineteenth, when the latter's increase ran undeviatingly at 35 per cent. from decade to decade until 1860, falling to 27 per cent. in the next three decades, and to 21 per cent. for the ensuing two. It was the astounding growth of the United States, we may recall, that moved Darwin in 1870 to deduce that in 657 years at a similar rate of progress four men would have to stand on every square yard of the whole terraqueous globe*. So Knibbs more recently notes that at the rate of progress immediately preceding 1911 the earth would in 500 years contain 246 billions of people†.

Such statements are beside the mark, save as picturesque and startling descriptions of current facts. Nevertheless, so practically important are the consequences of population growth or decrease, and so necessary are population aggregates for mere shop purposes, as a basis of general statistical measurements from day to day, that the statistician cannot avoid the problem of post-censal estimation—especially in a country that is "new" like Canada, and organized on a federal basis. On a larger view, the questions whether we are as a race to go on and if so, what share is Canada to have in that progress, are, we may agree with Mr. Keynes, the most interesting questions of the present day to which time will bring an answer.

Can we, then, or can we not predict the growth of population? On the one hand it is held that there is no simple law of population growth. True,

**Descent of Man*, vol. I, Part I, c. 4.
†Scientia, vol. XII.

population tends to increase, and that in constant ratio, but (*a*) the increase even in itself causes the rate to change, and (*b*) the increase cannot be isolated from other phenomena. In other words, the intervention of social and economic tendencies unpredictable with our present statistical apparatus except within brief limits of time—"imponderables" like the progress of invention, the relation of migration to natural fecundity, changes in the standard of living, etc.—allow any theory to work only by invalidating it altogether. On this basis the predictions of Elkanah Watson for the United States in 1815, which came true for half a century and which were once thought a marvel for prescience, were really borne out by accident. Even the multiple correlation method which Mr. E. C. Snow has so ably advocated as an attempt to take as many factors as possible into account must fail for lack of data and laboriousness of application. On the other hand, in the full light of all these difficulties, and indeed interpretative of them, we are told that there does exist a law of population growth capable of mathematical formulation—that there is an optimum population to which mankind tends to adjust its numbers, that, in brief, "growth of population is fundamentally a phenomenon like auto-catalysis". The phrase is that of Messrs. Pearl and Reed*, of whom more in a moment.

The present writers are concerned only (1) to explain the methods of post-censal and inter-censal calculation used for practical purposes in the Bureau of Statistics, and (2) to apply a formula or two to our past population records, with a view to seeing what horoscope these may cast for Canada for, say, around 1950.

(1) *Post-censal and Inter-censal Estimates.*—We are still without entirely comprehensive vital statistics for Canada, and likewise statistics of the net annual increase or decrease from migration. In any case it is doubtful whether records of the kind can be sufficiently inclusive to measure population from year to year by provincial areas, valuable though the figures would be as a check. We must fall back therefore on the application of a formula to Census results.

In the older countries the method of arithmetic progression is widely adopted for post-censal estimates, and the practice is also followed in the United States. It is not yet applicable to Canada. Neither, it may be added, does the method of geometrical progression apply, there having been only two decades out of our six in which it did not run wild. The Pearl method, which consists of fitting observed population data to the curve $\left(y=d+\dfrac{b}{e^{-ax}+c}\right)$, whatever its theoretical foundations, cannot be successfully fitted as yet to the Canadian situation. We submit for provisional adoption therefore the method of fitting by least squares a system of parabolas to the provinces, using the order of parabola that is found on trial to fit each province best, and adding these to represent the Dominion. Pearl himself used the parabola $Y=a+bx+cx^2+d\log x$, with satisfactory results prior to his introduction of biological law and the location of asymptotes. It frees us from theory of any kind as to the law of population growth, and it has the merit of reacting more favourably than any other we

*See Proc. Nat. Acad. Sci., vol. VI, No. 6, p. 287.

have tried to every test in the past. It can be continued in use, fitting it anew from census to census without much labour until a better consensus on the whole point has been arrived at.

With regard to inter-censal estimates and the correction of our yearly post-censal estimates in the light of the following Census, the formula used for the Dominion as a whole is the ordinary one of geometric increase, the tenth root of the ratio of increase between the censuses being taken from year to year. As the geometric ratio of increase for each Province would of course not add up to the geometric ratio for the Dominion, the yearly arithmetic increase of each Province is taken and made to agree *pro rata* with the geometric total for Canada by being multiplied by the ratio between the geometric increase of the Dominion over the total arithmetic increases of the provinces.

(2) *Quo vadimus?*—And now let us essay prophecy by the light of such mathematics as we have been able to invoke. The results will of course have merely an *intérêt de curiosité*, being but a projection of the rate of past growth into the future, whereas, as everyone knows, the future of a new country like Canada is veritably on the knees of the gods.

(1) At the rate of growth which prevailed between 1871 and 1921 the population of Canada in 1951 would be 14,794,000. At the rate of increase between 1881 and 1921 the total would be slightly higher, viz., 14,957,000. By similarly taking the rate between each successive Census and 1921, we have estimates for 1951 ranging from the above to 18,393,000. The geometric mean of these is 15,962,000. There is no good reason for taking this mean, except the usual one of seeking a *via media*. If instead of using two extreme points, as above, a straight line is fitted to the logarithms of all the Censuses, thus striking an average geometric rate of increase, the line will cross 1951 at over 14 millions, but the fit for the observed Censuses is bad.

(2) The Pearl and Reed method may be experimented with as a variant, though, as stated, it is not yet fairly applicable to Canada. Without going into details, the curve was fitted through three points to the Canadian population since 1851, namely, 1851, 1881, and 1911. This gives a prediction of 8,570,000 for 1921 compared with the Census figure of 8,788,483, which is low by less than 3 per cent., and much nearer than by the method of geometrical progression. In fact the aggregate difference between the figures obtained by this method and the Census is considerably less than the similar difference between the Census and results obtained by geometrical progression for every decade since 1851. When, however, the curve is produced to 1950 a population of only 12,274,000 is shown. This we think represents a flattening of the curve beyond all probability, for the reason that it brings the ultimate population of Canada to only 23 millions, whereas on the theory on which the argument for the reliability of the method is based, Canada should ultimately have as large a population as she can support, which is undoubtedly beyond 23 millions. This is not so much a condemnation of the method as an indication that we

have as yet an insufficient number of points of observation in Canada to enable it to be applied*.

(3) We incline, therefore, in the state of our present data, to the method of using the curve which fits best the observed population since our earliest records. So far as our experiments have gone, this has been found to be the system of curves already described. For 1921 the method gave a total of 8,823,000, which is nearer than Pearl's, and very close indeed to the actual Census. Working it out until 1950 a population for Canada approaching 15,000,000 is shown. Here the calculator may rest his case, unless he adds some infusion of theory as to population growth, which the above makes a point of avoiding.

"Be ye fruitful and multiply; seize the good places of the earth and use them." This advice has been followed by mankind in the past; Canada having so many "good places" still unseized may well see it followed in the future to a degree that will shatter any prophecy based merely upon her past. That, however, is a subject beyond the scope of this paper.

*The curve flattens much less if the Canadian-born population alone is considered, and the fit is also better prior to 1921.

USE OF MATHEMATICS IN ECONOMIC, SOCIAL AND PUBLIC STATISTICS

BY PROFESSOR A. L. BOWLEY,
London School of Economics, University of London, London, England.

My object in this paper is not to contribute any new methods or analysis of statistics nor to attempt to survey the whole field of mathematical statistics, but rather to raise a general discussion on a more limited question. On the one hand there are many methods used by experts in special subjects, which involve intricate mathematical analysis of a kind that is continually adapted to particular problems; the results are intelligible only to students of these subjects, and laymen must accept their findings with that mixture of credulity, scepticism, appreciation and interest with which the work of experts is generally regarded. Such subjects include biology, heredity, insurance and actuarial science, some branches of demographical statistics and other work which cannot be popularized. On the other hand there are methods, simple in analysis or in interpretation, which are common to many types of public statistics, and are or should be used in those sociological and economic investigations which are of general interest, and can be appreciated by ordinary educated people. It is to the use of mathematical methods in this latter class that I wish to direct attention, it being understood that there is no hard and fast line between the two groups I indicate.

I have in mind in particular the work that is properly undertaken by official statisticians, whose business it is to inform a government and the public on all matters relating to a nation or society that are susceptible to numerical measurement. Let us assume that there is adequate knowledge of the principles and methods of definition, classification, tabulation, graphic presentation and direct arithmetical processes, and ask what more is needed for the interpretation of the material which is collected.

The subject may be divided into two parts; in the first, the information is complete and covers the whole field; in the second the information is partial, depending on the examination of samples or on the "representative" method, and questions of error or precision arise, involving the use of algebraic probability.

In both there is a cross division into results that can be expressed in words intelligible to the ordinary reader, and results that need technical terms or mathematical symbols. The mathematics may not uncommonly be regarded as a scaffolding to be removed when the edifice is complete. The results may depend on mathematical analysis and yet be expressible in simple terms for the layman. But by way of an appendix or otherwise, the methods by which the results are obtained from the data and the formulae used should be accessible to the scientific student.

A. Methods not Involving Errors or Probability

Under the first part I include the trend and fluctuation of time series, the measurement of seasons, the presentation of frequency groups, double tabulation and correlation, and index-numbers.

A. 1. *Time Series*

Consider first annual averages, so as not to be concerned with seasonal movements.

The question whether graphic representation should be on a natural scale or on the ratio or logarithmic scale has been well ventilated and I need not discuss it, except to remark that *a priori* it is only when there is a presumption that a given change in the forces affecting the numbers produces a constant proportional change in those numbers that the ratio scale is justified. Such a presumption is proper in index-numbers of prices and in population totals.

If ratios are chosen, then the numbers should be replaced by their logarithms for further work. It seems to be assumed in popular presentation that the meaning of the word ratio is appreciated, but that the word logarithm is esoteric.

The task in presenting time series is to distinguish the trend from the fluctuations. Whatever method is taken the result can be expressed graphically, and the ordinary reader need not be involved in the arithmetic.

It is recognized now that the fluctuations due to the "trade cycle" are of such irregular period that one cannot begin by estimating them and then proceed to study residuals. We have in fact in very many economic series to deal with two classes of variation at once, namely, those which rise in two or more steps from a minimum to a maximum and then fall for two or more years, and those whose movements are completely sporadic or which tend to rise and fall in alternate years.

The methods in use (excluding the non-mathematical method of smoothing with a free-hand line) are the fitting of a straight line or a parabola of the second or higher degree, and the method of moving or progressive averages.

The rationale of the first method is the assumption that the long-period or cumulative changes are continuous and not capable of producing a rapid change of curvature. Consequently the fourth* or even lower differences of the curve representing the trend are negligible, and a formula $y=a+bx+cx^2+dx^3$ can be used. We do not escape from arbitrariness by this method, since the choice of the number of terms sensibly affects the result, unless the series is in fact rapidly convergent.

It has been found in several cases that when a line calculated by such a formula is drawn on a chart showing the observations, that the fit is obviously bad. In fact there have been sudden reversals of tendencies, especially in prices

*In some studies parabolas of a higher degree have been used.

(circâ 1874 and 1895) and measurements depending on value, and the smoothing of them is unjustifiable. Again a parabola even of high degree is not capable of representing the phenomena of the war crisis in the same curve as the more leisurely movements in pre-war years. Investigators have therefore in some cases substituted two or more straight lines or even two second-degree parabolas for a single curve, and this opens the gate to such arbitrariness that one might perhaps with more reason smooth by a free-hand curve.

The choice of constants is generally made by the method of least squares. It is in any case empirical, and there is something to be said for minimizing instead the total of the (signless) deviations of the first degree; Professor Edgeworth has developed methods by which this can be done in simple cases, and the possibility of finding a simple process might well be further explored.

For myself I prefer the method of moving averages in many cases. The problem being to separate a total into two parts, one of which is attributable to long period causes, the other to the temporary circumstances of a particular year, the direct process is to estimate the first part by averaging the results of a period in which the year in question is central. There are three objections to or difficulties in the method. The choice of the length of the period to be averaged is arbitrary, and since it ought not to contain two maxima and one minimum of the trade cycle (if that affects the figures) it may be advisable to vary the length; whether this is done or not small irregularities result which are irrelevant to the trend; and where the true trend is curved the averages lie within the curve, not on it.

Of these objections the first is implicit in all graduation formulae, and I see no theoretic objection to varying the period to be averaged. To the second, so far as popular presentation is concerned I see no objection to smoothing off small irregularities arbitrarily, or to using a short interpolation formula in more refined work. To overcome the third, when it is serious, I suggest the following method.

Suppose the curve to be convex to the axis for 30 years. Let $\bar{y}_{-10}$, $\bar{y}_0$, $\bar{y}_{10}$ be the averages of the first, second and third decades. Then if t is the number of years from the centre of the period, the parabola

$$y=\bar{y}_0+\frac{1}{20}(\bar{y}_{10}-\bar{y}_{-10})t+\frac{1}{200}(\bar{y}_{10}+\bar{y}_{-10}-2\bar{y}_0)t^2$$

passes through the points representing these averages.

Consider the average of the $(2m+1)$ years from $(s-m)$ to $(s+m)$ after the centre. For the points on the arc of the parabola named it is easily shown to be $y_s+\frac{m(m+1)}{600}(\bar{y}_{-10}+\bar{y}_{10}-2\bar{y}_0)$. That is, in a parabola the average is above the point at the year in question by $\frac{m(m+1)}{600}(y_{-10}+\bar{y}_{10}-2\bar{y}_0)$ for all values of s.

Subtract this quantity from the computed average for the year, and we have a reasonable correction for curvature.

Whatever the method of smoothing used, the deviations of the observations from the positions on the trend line can be computed and represented in a separate diagram.

The correlation of two such series involves so many difficulties, still the subject of discussion, that it may be left in the group of special problems outside the public province of the official statistician.

A. 2. *Seasons*

I have recently dealt with measurements of seasonal movements in a memorandum issued by the London and Cambridge Economic Service,* and therefore only give an outline here.

The problem that should be solved in official statistics is the measurement of the seasonal variations found in monthly or weekly records over a series of years. An example is to be found in the English statistics of unemployment. The direct method, and the only one (so far as I know) officially employed, is to compile the arithmetic averages of the records for January, of the records for February, and so on. It ought to be decided, however, whether the movement is absolute or relative, that is whether we should average quantities or ratios; whether, for example, we should say that wheat imports were 10 mln. cwt. or 10% below the average in January. If the latter, we should probably take the geometric mean. Further there is the question whether we should compare each month with that preceding or with the annual average, *e.g.*, whether we should take as our datum that January imports are 20 mln. cwt. below December imports, or 10 mln. cwt. below the annual average. Finally if there is any trend during the period, the December average will be affected by 11 months average trend over the January average, and some means must be used for correcting this.

Whichever of (at least) seven possible methods are used the results can be stated in simple non-mathematical language.

The measurement is important in retail (and perhaps wholesale) prices, foreign exchanges, banking statistics, production and foreign trade, and many other branches of common statistics.

A. 3. *Statistical Groups*

Let us assume that we have a group of persons or things, reasonably selected as having in common a number of characteristics, which are adequately defined, and that it is required to study their distribution as regards some one measurable variable. Common examples of such groups are the distribution by age or by wage in an occupation. A more complicated example is to be found in the variation in expenditure on rent or on meat in a number of family budgets.

In such cases the required information can be given in a well graded table, from which students could make any measurement they wished. But considerations of space and expense often make it impracticable to give full detail, and therefore one or more measurements of variation should generally be given. In fact, a great part of very important information, collected at great expense, is often lost because only averages are stated.

*Special memorandum No. 7, *Seasonal Variations in Finance, Prices and Industry,* by A. L. Bowley and K. C. Smith.

The mathematical statistician commonly uses the standard deviation for this purpose. Here it may be remarked that much confusion arises from the fact that some writers use probable error and others standard deviation for the same purpose, often without a statement of their choice. It would be well if the phrase "probable error" was dropped out of descriptive statistics, and used only when the theory of error is definitely involved; even then the word "probable" is an unfortunate choice. If this quantity is obtained by the factor ·674... from the calculated standard deviation, it is simply misleading, unless it is definitely known that the distribution is normal. In descriptive statistics, which are now in question, the term "quartile deviation" is preferable, and it should be obtained from the computed quartiles.

Standard deviation is not specially appropriate for the description of wage and similar distributions. There is much to be said in favour of computing quartiles and deciles, both because they involve no hypotheses (except some working rules for their localisation), and because they can be expressed in non-mathematical language. Further, since in general it cannot be assumed that the distribution is approximately normal, and that if the standard deviation is known the distribution throughout the scale is known approximately, the mere statement of the standard deviation does not give enough information. In practical questions quite as much depends on the numbers at the extremity of the scale, as the average and the numbers near it. It is the number whose wages are low, or the districts in which infant mortality is high, that concern the reformer.

I find a statement such as the following, which depends on the average, the quartiles and two deciles, of great utility:

The average weekly wages of a group are 58*s*. One half of the men received within 12*s*. (or 20 per cent.) of 60*s*.; four-fifths within 18*s*. (or 30%) of 60*s*. Or in more detail, one-tenth received less than 42*s*., one-quarter less than 48*s*., one-quarter 72*s*. or more, one-tenth 78*s*. or more.

I wish to emphasize that statements of family expenditure lose a great part of their value, because averages only are commonly stated. We should be told not that the average bread consumption is 33 lbs. weekly, but that it is 33 ± 6 lbs. weekly, where 6 is the quartile deviation; or preferably we should have two measurements of variation such as I have suggested.*

A. 4. *Double tabulation and correlation.*

It is doubtful how far measurements of correlation or regression should be used in common statistics. The general reader is of course ignorant of the meaning of the words, and even mathematical statisticians are apt to read into the measurements much more than they really imply.

The difficulty can be evaded in part as follows. Suppose we have a common double tabulation in two variables, such as income and rent. Grades being suitably chosen, the averages of columns and lines can be computed, so that the

*It is awkward that unless the distribution is symmetrical the median is not midway between the quartiles. The actual phrasing of the non-technical statement requires care on this account.

average rent for families in each given income grade is stated, and the average income in each rent grade. These can then be graphed, so as to show the unadjusted lines of regression.

If the lines are approximately rectilinear, and not otherwise, their gradient can be computed, by the full product-sum formula or some approximation to it. For the non-mathematical reader the result can then be expressed in the form "On the average an increase of 1*s*. in weekly rent is found with an increase of 5*s*. in weekly income, over the range of incomes 25*s*. to 50*s*." The graph would show how nearly true the statement was.

For such purposes the idea of regression is more important than that of correlation, and the expression in the form I have given is free from the vagueness and looseness of thinking that is common in connection with a correlation coefficient.

Even in this guarded form there is risk of misapprehension. The statement of average rent may be taken without regard to the variation in the array, so that it might be assumed as generally true that, with an income of 40*s*., 7*s*. was the rent, with an income of 45*s*., 8*s*. the rent and so on. The variation in an array is* $s \times \sqrt{1-r^2}$, whatever the nature of the regression, when s is the standard deviation for the whole group not distributed according to income, and r the product-sum correlation coefficient between rent and income. If r is as great as ·8, $\sqrt{1-r^2}$ is no greater than ·6. If s was, for example, 2*s* 6*d*. and r ·8, we should know that the variation of rent from the general average was (on the measurement by standard deviation) 2*s*. 6*d*.; if in addition we knew the income, the variation from the average for that income would still be as great as 2*s*.

A. 5. *Index-numbers*

I hope that my introduction of this subject in a subsection will not lead to any discussion on the best formula for index-numbers, since all, perhaps more than all, that can usefully be written on that subject is common property. The three points to which I wish to call attention are common to many formulae.

The use of the geometric mean, if it is preferred on theoretical grounds, need not be rejected because the unlearned do not understand it, nor because it takes long to compute. The use of logarithms is common in engineering and several other professions, it is taught in England to children 14 years old, and it is familiar to all who have any right to deal with statistics. For others it can be used without confession or explanation. The additional time in calculation, as compared with an arithmetic mean, can hardly be two hours in any ordinary index-number.

Where weights are used, they should be reconsidered at as frequent intervals as possible, with a view to determining whether their adjustment would make any significant difference. If Laspeyres' and Paasche's formulae give practically the same results, there is no need to change the weights. Laspeyres' formula is commonly in use and can be easily expressed in non-mathematical language, and this is an advantage over a formula using double weights. Whether an

*Sometimes called the "coefficient of alienation".

adjustment is called for can often be judged by general statistics; thus for the English Cost of Living Index, the figures of national consumption of grain, meat, sugar, tea, etc., will show whether any marked change in the distribution of expenditure is taking place. If there is a significant change, then it is well to compute and publish the results both by Laspeyres' and by Paasche's formula, and also to adopt Professor Irving Fisher's "Ideal Index Number" or some other that is practically equivalent to it. The difficulty so far has been the ignorance of appropriate weights at more dates than one.

A very great deal depends on the precision of index-numbers, and to aid in judgment about it a measure of the dispersion of the price-changes of the different commodities should be given. Since the dispersion differs according to the base year taken, it should be calculated on two or more hypotheses, for example with reference to 1914 and with reference to twelve months before the current date. The quartile deviation expressed as a percentage of the average is a possible measurement.

A. 6. *Standardization*

Though the use of a standardized death-rate and other measurements involving similar processes is arithmetical, yet it may suitably be included among mathematical processes. Essentially it is a method of eliminating the variation of one factor in order to study in isolation the variation of one or more other factors, and it is a particular example of weighted averages. Insufficient attention has been given to the computation of standardized marriage and birth rates, and material is or can be made generally available for the process.

Standardization is, of course, used in other connections, and in particular in adjusting budgets of expenditure. Further work is needed on food requirements according to age, sex and occupation. After the minima of calories, carbohydrates, proteids, etc., have been ascertained, there remains the application of the results to customary expenditure, to the expense of different foods and the cost of cooking. But these are matters for the specialist rather than for the official statistician. I wish to mention, however, that the requirement for house-room according to age and sex needs investigation, and that satisfactory definitions and measurements of crowding or overcrowding have not yet been made.

B. Methods Involving the Theory of Probability

In many enquiries it is impossible to collect complete information; for example, budgets of family expenditure can only be obtained from a limited number of persons. Even when universality is conceivable, it is often out of the question owing to expense. Further, it is very often important to reach results in a very short time, if they are to be of any utility, and it is possible to handle 1000 or even 10,000 cards in a few weeks, while 1,000,000 or 10,000,000 would with the best organization take several months.

In a very wide region of investigation two conditions are necessary for pure sampling; first, that the universe to be examined should be exactly defined; second, that every person or thing in the universe shall have an equal chance of being included in the sample. The second condition commonly requires in sociological statistics, that there should be an adequate list of the objects in the universe, from which those to be examined can be selected. The condition of equal chance can be obtained by numbering all the objects and then selecting by a random list of numbers (*e.g.*, from a seven figure table of logarithms); or it can be obtained by marking throughout the list one in one hundred (or whatever the proportion the sample bears to the whole). If the objects are in strata in which the attributes or magnitudes in question are associated, Mr. Yule shows that the latter method gives a greater precision.* Whichever the method it is essential that no alteration from the objects chosen should be allowed owing to difficulties of observation. The universe of which the sample is in fact representative is that over which the rule of selection holds good.

Pure sampling, however, is often impracticable, and then we have to form a judgment, often on non-mathematical grounds, of the probable effect of a breach of the conditions.

I follow Mr. Yule's useful classification and distinguish between variables and attributes.

B. 1. *Average by sample*

Take for example the investigation in the Census for Scotland of 1911† about the number of children according to the occupation of their fathers. The contents of the sample was determined by the condition that the families should be completed and the details available. The universe may be roughly defined as that of Scotchmen married for more (say) than 20 years and not emigrated. The number of families in all was 133,960; the average number of children per marriage was 5.82 with standard deviation 3.099. For those 923 families in which the father was a boiler-maker the average number of children was 6.00, with standard deviation 3.039. The question put to be answered from these figures is, is the difference shown, *i.e.*, 6.00—5.82, one that might be found in this system of selection, though in general the number of boiler-makers' children is no greater than in all occupations. The answer given, in effect, is that the standard deviation of the difference to be expected from averages so found if this occupation is quite unconnected with number of children is .10, and the actual difference .18 is not so far up the normal scale of error as to make the divergence significant.

Two points are to be noted. First it is assumed that the normal scale is roughly applicable. This I am not at all inclined to question when the numbers on which the averages are based are so large. Secondly, there is no appeal to the principal of inverse probability. If there were no difference, an event has happened the chances against which may be measured as about 13 : 1.

If now we put a quite different question: what was the average number of children in Scotch families defined as in the Census?—the selection being regarded as a random sample of a larger universe (which included, for example, all the Scottish and not merely those for whom records were available), the answer would be in the form $5.82 \pm \frac{3.099}{\sqrt{133,960}}$ or $5.82 \pm .01$, approx. If in the universe

An Introduction to the Theory of Statistics, by G. U. Yule, 1922 edition, pp. 285 and 349.

†Cd. 7163, p. 288.

the average were 5.82, then in a sample the odds would be about 2 to 1 against differing from this by more than .01 is (within certain limitations) a correct statement. But to make the result useful we must invert this, and the very difficult question arises whether under the circumstances we can say that it is 2 to 1 against the average in the universe being more than .01 different from the average found in the sample.* This question cannot be discussed in a parenthesis and I propose only to indicate some points about it under my next heading. Here I wish to emphasize the necessity of making some inference from the sample to the universe, if the sample is to be of value, and to call attention to the fact that in so doing we step into quite a different region of the theory of probability.

In publication of such results for the general reader, an indication of the precision of the result ought to be given. How lucidity is to be combined with accuracy is a difficult problem.

B. 2. *Frequency of attributes by sampling*

If p is the proportion in a universe of persons or things possessing a certain attribute, and from it n things are taken independently and at random, then if n and pn are sufficiently great, the chance of obtaining a proportion in the sample differing from p by as much as any assigned quantity x is given by the table of the normal error function, with standard deviation $\sqrt{\{p(1-p)/n\}}$.

If n is sufficiently great we can ignore the term that comes from the second approximation to Bernoulli's Laws, which involves the factor $(1-2p)/\sqrt{pqn}$. But it has not been sufficiently noticed that, if we do not neglect this term, it still disappears if we express our result in the form—the chance of differing from p by as much as the standard deviation whether in excess or defect is .317; for the excess in the integral due to this term on one side of the average is exactly balanced by the defect on the other side whatever the range taken.

In practicable investigations, where it is desired to draw all permissible inferences from the sample, the condition that pn should be great is important. If pn is small, we obtain Euler's series, *i.e.*, the approximation known as the "law of small numbers". After some rather careful examination I find that the normal table can be applied with very fair accuracy when pn is as great as 10.

Thus if $pn=10$, and n is sufficiently great to allow us to neglect $\sqrt{1-p}$, then the chances of finding numbers within certain limits is as follows:

Number of successes	*Chances*	
	Law of small numbers	*Normal table*
0 to 5	.067	.078
6 to 14	.850	.844
15 or more	.083	.078
	1.000	1.000

For smaller values the following table may be of use:

*See Metron, vol. II, N° 3. *The precision of measurements estimated from samples.*

Number of Successes	Value of pn: 1	2	3	4	5	6	7	8	9
0	.3679	.1353	.0498	.0183	.0067	.0025	.0009	.0003	.0001
1	.3679	.2707	.1494	.0733	.0337	.0149	.0064	.0027	.0011
2	.1839	.2707	.22405	.1466	.0842	.0046	.0223	.0107	.0050
3	.0613	.1805	.22405	.1954	.1404	.0892	.0521	.0286	.0150
4	.0153	.0902	.1680	.1954	.1755	.1339	.0912	.0573	.0337
5	.0031	.0361	.1008	.1563	.1755	.1606	.1277	.0916	.0607
6	.0005	.0120	.0504	.1042	.1462	.1606	.1490	.1221	.0911
7	7 or over	.0034	.0216	.0595	.1044	.1377	.1490	.1396	.1171
8	.0002	.0009	.0081	.0297	.0653	.1033	.1304	.1396	.1318
9	..	9 or over	.0027	.0132	.0363	.0688	.1014	.1241	.1318
10	..	.0002	.0008	.0053	.0181	.0413	.0710	.0993	.1186
11	..	..	11 or over	.0019	.0082	.0225	.0452	.0722	.0970
12	..	..	.0003	.0006	.0034	.0113	.0264	.0481	.0728
13	..	..	..	13 or over	.0013	.0052	.0142	.0296	.0504
14	..	..	..	.0004	.0004	.0022	.0071	.0169	.0324
15	..	..	..	..	15 or over	.0009	.0033	.0090	.0194
16	..	..	..	..	.0003	.0003	.0014	.0045	.0110
17	..	..	..	..	..	17 or over	.0006	.0021	.0058
18	..	..	..	..	..	.0002	.0002	.0009	.0029
19	..	..	..	..	..	..	19 or over .0002	.0004	.0014
20	..	..	..	..	..	..	..	20 or over .0003	.0006
21	..	..	..	..	..	..	..	..	21 or over .0004

Generally the chance of reaching or exceeding $u+d$, where $u=pn$ is $< \dfrac{e^d(u+d+1)}{\left(1+\frac{d}{u}\right)^{u+d}(d+1)\sqrt{2\pi(u+d)}}$.

See Tables for Biometricians for a more complete table. The above is only roughly computed as regards the fourth decimal place.

It has been objected to this method that the existence of uncommon attributes may not be revealed in the sample. The table just given enables us to test how far this is the case. If, for example, $n=100$ and $p=.01$ (one per cent.), the chance of finding no instances is as great as .368; if $p=.06$, the chance of finding none is only 1 in 400. If $n=900$ and $p=.01$, the chance of finding less than 3 is .0062. If we want to deal with so small a proportion as one per cent. we should increase our sample to 1000; even then our expected result would have a considerable standard deviation, and might be written $.01 \pm .0032$.

In such investigations, however, the problem is generally the inverse one: given the proportion in the sample, what can we infer about the universe? Suppose, for example, we find in a sample of 1000 four items that have a certain property. The table above gives the following information:

Proportion in universe	*Expectation of* 4 *successes in sample of* 1000
1 in 1000	.0153
2 '	.0902
3 '	.1680
4 "	.1954
5 "	.1755
6 "	.1339
7 "	.0912
8 "	.0573
9 "	.0337
10 "	.0189
	.9784

Now we should not be justified in saying that, in the absence of information to the contrary it was just as likely that the universe contained 1 case per 1000, as 2 cases, 3 cases, etc., and we are not entitled to add up the chances of obtaining 4 in the successive universes. We can, however, say that if the universe contained 2 or less per 1000 the chance of obtaining 4 was less than one-tenth, and if it contained 7 or more per 1000 the chance was again less than one-tenth. We can in fact state a range of the proportions in the universe beyond which the chance of finding 4 instances was small, and within which it was moderate.

With larger numbers also we can marshal hypothetical universes and assign the chance of obtaining the given number of successes to each. These chances tail off rapidly as the difference between the most likely number from each universe and the number realized passes from once to twice and three times the quantity $\sqrt{p(1-p)n}$, where p is the proportion found in a sample containing n things. Unless then there is some *a priori* reason for believing that the universe differs considerably from the sample, we have, I think, good reason to expect that the proportion in the universe is within, say, $2\sqrt{p(1-p)/n}$ of the proportion found in the sample.

I have endeavoured to carry the argument further, and shown that under hypotheses I think reasonable (where n is great) the normal table of error can be applied to determine the chances of possible values of p in the article already named,* and there also I deal similarly with averages as well as with proportions. It is evident that the problem requires very careful and clear thinking, and here I have done no more than give a partial solution in outline. We cannot, however, ignore it, if we are to make any inference from samples, and we cannot generally take so large a sample as to make it unimportant. But, finally, we must admit that to pass from the sample to the universe is in fact a mental process which every uninstructed person would follow without hesitation, and if the measurement of chance is fundamentally psychological our mathematics must tally with this process.

B. 3. *Correlation by sampling*

It is a very simple matter to calculate r by the product-sum formula, and to write down the value of $r \pm \frac{1-r^2}{\sqrt{n}}$, but it is by no means easy to interpret the result, or even to justify the expression. To discuss the meaning of r, if it is a measurement as well as the result of a numerical process, would need a paper to itself. I only here offer some remarks on the standard deviation.

If the universe is given, the standard deviation for the frequency curve of r based on a multitude of n-fold samples is not of the simple form generally used, but is much more complicated†. It may be written as follows:

$$\sigma_r^2 = \frac{1}{4n^2} S\{r(x^2+y^2)-2xy\}^2,$$

where the standard deviations of x and of y are taken as unity, x and y are measured from their averages, and the summation of the squares is extended over the n pairs. If $r = \pm 1$, the full formula gives the same result as the approximation generally used, it tends to do so when $r=0$, and the two become identical if the original distribution is normal. In some typical non-normal cases in which I have tried it, the full result differs very little from the approximation. There is in fact not much room for variation in ordinary cases, but the agreement ought to be tested, and in the form I have given it the full working is not arduous.

When the universe is not given, but is to be inferred from the sample, the difficulties I have already indicated are present in an aggravated form; we cannot suppose that universes in general are distributed normally with r as the variable. Safety can only be obtained if n is so large as to make the form of the distribution of the universe unimportant. I suggest that we may be able to establish the value of r in the universes correct to the first decimal place if n is as great as 1000, and r is not near 0 or 1. It is futile to work out values for r to three or four places in problems of the nature here in question.

*Metron, *loc. cit.*

†The formula is given by Dr. Sheppard, *Application of the Theory of Error*, Trans. Royal Soc., vol. 192, 1898, A. 229, p. 128; and is quoted in Bowley's *Elements of Statistics*, 1920, p. 422.

B. 4. *The representative method*

I distinguish the representative method from the method of pure sampling in that the rule of equal chance of inclusion for every person or thing in the universe is not obeyed, but whole sections or classes are chosen which separately or grouped together are held by some criterion to be representative of the whole.

For example the distribution of income might be examined by choosing a district or class in which (1) the average income was the same as the average of incomes in the country, (2) the proportion of persons paying income tax, and (3) the proportion receiving wages were the same.

The adequacy of the method depends very much on the extent to which the distribution is similar in different districts. If, for example, Pareto's Law applied over a certain range of incomes, and the number of incomes in the country were known, and it was known that the constant determining the gradient of the curve was the same in a particular district as in the whole country, then this constant (and therefore the whole distribution) could be found from the numbers of incomes above and below any assigned mark in that district.

If the law is not known at all, the increased precision obtained by the knowledge that certain averages or proportions were the same in the district as in the country can be considered as follows:

Let there be n values of a variable in a country, $a_1, a_2 \ldots a_n$. Let the proportions of the population at these values be $p_1, p_2, \ldots, p_n$. In the selected district let the proportions be p_1+x_1, $p_2+x_2, \ldots$, so that

$$p_1+p_2+\ldots+p_n=1$$
$$x_1+x_2+\ldots+x_n=0.$$

Let $\sigma_1, \sigma_2\ldots$ be the standard deviations of $p_1, p_2\ldots$ in the various districts in the country, and suppose p_1 to be normally distributed, and also p_2, etc. Let r_{st} be the coefficient of correlation between p_s, p_t over all the districts.

Then the chance of finding the groups of deviations $x_1, x_2\ldots$ in a district selected at random is $Ke^{-\frac{1}{2}\chi^2}$, where K is a constant and

$$\chi^2=\frac{1}{R}\left(\frac{x_1^2}{\sigma_1^2}R_{11}+\ldots+\frac{2x_1x_2}{\sigma_1\sigma_2}R_{12}+\ldots\right),$$

and R is the determinant formed from the terms r_{st}, and R_{11}, R_{12}, etc., its cofactors.

The chance of finding a composite discrepancy so great as that indicated by χ^2 is given by Professor Pearson's table of goodness of fit, by the quantity named P.

P is reduced* successively by every additional limitation on the values of $x_1, x_2, \ldots$.

If, for example, $n=8$, the chance of so great an error as that indicated by $\chi^2=10$, is .350. Now if we are given that the average of all in the district equals

*See: *On the interpretation of χ^2 from Contingency Tables*, R. A. Fisher. Journal Royal Stat. Soc., January 1922.

that in the country, so that

$$p_1x_1+p_2x_2+\ldots+p_nx_n=0,$$

the table must be used as if n were reduced by 1. Then $P=.265$ instead of .350. If three such limitations are given, $P=.125$, if six, $P=.019$.

As the number of controls approaches the number of divisions to be observed, the chance of a serious error diminishes with a rapidity indicated by these figures.

If there is no correlation between the proportions, additional controls do not give much help. Thus in the same notation, if $\sigma_1=\sigma_2=\ldots=\sigma_n$, I find that the fixing of the average* only reduces σ_1, which measures the precision in the first class, to

$$\sigma_1\sqrt{\frac{(n-1)\ (n-2)}{n(n+1)}}\ ;$$

if $n=10$, this is approximately $.81\sigma_1$.

These tentative examples will serve to indicate that it is not easy to measure the precision of samples selected in this way. The measurements depend to an unfortunate extent on the good judgment of the investigator. Confirmation from other sources can, however, often be found, and good results that have been tested have been obtained. The whole method is the subject of enquiry by a commission of the Institut International de Statistique†, and it may be hoped that useful rules may be laid down and appropriate tests devised.

*The result here is only true if a_1, $a_2\ldots$ are in arithmetic progression. The reduction to $\sigma_1\sqrt{\frac{n-1}{n}}$ is obtained simply from the condition that the sum of the x's is zero. The formula has only been verified up to $n=7$, but it appears to be general.

†Reported at the Congress at Rome, 1925, and published in the Bulletin de l'Institut International de Statistique, Tome XXII, 1[e] Livraison, p, 357 *seq.*

ABSTRACTS OF COMMUNICATIONS

SECTION V

SUR UNE MÉTHODE D'ÉVALUATION DES INTÉGRALES DE PROBABILITÉ

PAR M. L. E. PHRAGMÉN,
Académie Royale des Sciences, Stockholm, Suède.

Si $s(u)$ désigne l'ordonnée d'une courbe de fréquence, on sait que la fonction $S(u)$, dite *réciproque* ou *adjointe*, et liée à $s(u)$ par l'égalité

$$\int_{-\infty}^{+\infty} s(u)\, e^{2\pi itu}\, du = S(t) = A(t) + iB(t),$$

satisfait aussi à l'égalité

$$\int_{-\infty}^{+\infty} S(t) e^{-2\pi iut}\, dt = s(u),$$

de sorte que l'intégrale $\int_{o}^{u} s(u) du$ peut être représentée par l'expression

$$\int_{-\infty}^{+\infty} \left\{ \frac{A(t) \sin 2\pi ut}{2\pi t} + \frac{B(t)\ (1-\cos 2\pi ut)}{2\pi t} \right\} dt.$$

L'objet de ce travail est d'attirer l'attention sur le fait un peu inattendu que la valeur approchée de cette intégrale obtenue en appliquant le schéma de calcul

$$\int_{-\infty}^{+\infty} F(t)\, dt = \sum_{\lambda=-\infty}^{+\infty} \frac{1}{k} F\left(\frac{\lambda}{k}\right)$$

donne une méthode très commode et très générale pour évaluer l'intégrale $\int_{o}^{u} s(u) du$ avec une grande approximation.

A GENERALIZED LAW OF ERROR

By Professor P. R. Rider,
Washington University, St. Louis, Missouri, U.S.A.

In an article entitled "First and Second Laws of Error" (Journal of the American Statistical Association, volume XVIII, new series, number 143, September, 1923, pages 841-851) Professor E. B. Wilson has shown that certain data give evidence of being distributed according to Laplace's first law of error which states that the frequency of an error or deviation can be expressed as an exponential function of the absolute value of the error, rather than according to Laplace's second law,—called the normal law or Gauss's law,—which states that the frequency is an exponential function of the square of the error. As it hardly seems reasonable to suppose that there is a sharp line of demarcation between data that follow the first law and those that follow the second, the author suggests the more general law that the frequency of an error can be expressed as an exponential function of the mth power ($m>0$) of the absolute value of the error. This law reduces to Laplace's first law when $m=1$ and to his second law when $m=2$. The author's analysis of the data considered by Professor Wilson indicates that certain data may be found to conform more closely to the general law proposed, where m has some other value than 1 or 2, than to Laplace's first or second law.

A METHOD OF ESTIMATING THE SIGNIFICANCE OF THE DIFFERENCE BETWEEN TWO AVERAGES BY MEANS OF BAYES' THEOREM ON THE PROBABILITY OF PROPORTIONS

By Dr. G. F. McEwen,

Scripps Institution for Biological Research, La Jolla, California, U.S.A.

There is an acknowledged need on the part of experimentalists, especially in biology, of a method of obtaining readily an estimate of the significance of the difference between two averages of a small number of measurements. The variety of frequency distributions presented by such measurements together with the small number of observations usually available often render the "probable error" method inapplicable. This paper presents an extension of Bayes' theorem on the probability of obtaining differences in proportions in two samples of given numbers to averages of measurements.

In each of the two series to be compared the measurements are arranged according to magnitude, and the difference between the averages of the two series is assumed to depend upon the difference between the proportions of the measurements in two groups for each series, one group exceeding an empirically determined intermediate value, and the other falling short of this value. Sub-groups of these groups may also be used for greater accuracy. The method is based upon the idea of direct "random sampling" of the measurements, rather than upon their distribution according to some theoretical frequency curve. Tables have been computed from Pearson's convenient formulation of Bayes' theorem, thus rendering the method readily applicable.

ON THE DEVELOPMENT OF FORMULAE FOR GRADUATION BY LINEAR COMPOUNDING, WITH SPECIAL REFERENCE TO THE WORK OF ERASTUS L. DE FOREST*

By Mr. Hugh H. Wolfenden,
Consulting Actuary, Grimsby, Canada.

"Graduation" by "linear compounding" is effected by replacing an observed value u_0 (where the true value $U_0=u_0+e$) by a linear compound, v, of u_0 and terms $u_1, u_2, \ldots, u_{-1}, u_{-2}, \ldots$ on either side of it, on the assumption that differences of U beyond a certain order may be neglected. This may be done (I) by interpolation, from single values, or groups, or averages thereof, without any criterion of "fit" or "smoothness" beyond the above assumption, (II) by fitting—making the mean square error of v a minimum (least squares), or (III) by reduction of error—making the mean square error of $\Delta^n v$ a minimum. The names usually associated with the development of (I) are Davies, Berridge, Woolhouse, Higham, Hardy, Karup, Spencer, Lidstone, King, Wickens, and others; of (II), Landré, and W. F. Sheppard; of (III), Hardy, W. F. Sheppard, Henderson, and Larus. It is not generally known that much of the theory of (I), the "least square formulae" of (II), and those of (III) which make the mean square error in $\Delta^4 v$ a minimum had previously been discussed very fully by Erastus L. De Forest, in Smithsonian Reports 1871 and 1873, in a pamphlet "Interpolation and Adjustment of Series" (1876), and in the Analyst (Des Moines), 1877-79. This paper traces the development of the above methods, and classifies them, with the particular object of showing the relation of De Forest's important work to that of more recent investigators.

*This paper has since been published in full in the Transactions of the Actuarial Society of America, Vol. XXVI. pp. 81-121.

COMMUNICATIONS

SECTION VI

PHILOSOPHY, HISTORY, DIDACTICS

UNIFORMITY OF MATHEMATICAL NOTATIONS—RETROSPECT AND PROSPECT

By Professor Florian Cajori,
University of California, Berkeley, California.

In mathematical notations mathematicians are not profiting by the teachings of history. As one surveys mathematical writings of the last five centuries, certain facts arrest the attention It is noticeable, for example, that no one individual can invent an extended system of symbols which all mathematicians will adopt. W. Oughtred* used one hundred and fifty symbols, many of his own design. Of the latter only one, the St. Andrew's cross for multiplication, is still in general use. The long lists of symbols framed by P. Hérigone† in the seventeenth century, and by C. F. Hindenburg‡ in the eighteenth century have passed away. These and more recent experiences indicate that mathematical symbols, being for community use, must be adopted by the community; they cannot be forced upon it. Mathematicians constitute a democracy; each individual has been inclined to exercise a personal right to express his ideas in any manner he pleases. Consequently in some fields there has been no generally adopted sign language for the intercommunication of ideas, and progress has been retarded. We still are in the shadow of the Tower of Babel.

It is my purpose to illustrate the great diversity of symbolisms in the past, to exhibit the extreme slowness of unification and to stress the need of greater organized effort.

In the latter part of the fifteenth century $\tilde{p}$ and $\tilde{m}$ became symbols for "plus" and "minus" in France§ and Italy||. In Germany the Greek cross and the

*W. Oughtred's symbols, as found in the several editions of his *Clavis mathematique*, and in his other publications, have been listed in a paper by the present writer, printed in the University of California Publications in Mathematics, Vol. I, 1920, pp. 171-186.

†P. Hérigone's *Cursus mathematicus*. Paris, 1634, 2. ed. 1644, Vol. I, "Explicatio notarum."

‡Of C. Hindenburg's publications on combinatorial analysis, the following may be specially cited: *Infinitinomii dignitatum exponentis indeterminati, historia, leges, ac formulae*, Göttingen 1779; *Novi systematis permutationum, combinationum ac variationum primae lineae*. Lipsiae, 1781; *Sammlung combinatorisch-analytischer Abhandlungen, erste Sammlung*, Leipzig, 1796, *zweyte Sammlung*, 1800. See also Ernst Gottfried Fischer, *Ueber den Ursprung der Theorie der Dimensionszeichen*. Halle, 1794.

§N. Chuquet: *Le Triparty en la science des nombres*, 1848, published by A. Marre in B. Boncompagni's Bullettino, Vol. XIII, 1880, Vol. XIV; Estienne de la Roche, *Larismethique*. Lyons, 1520; 2[e] ed., Lyons, 1538.

||L. Pacioli: *Sūma de arithmetica geometria proportioni et proportionalita*. Venice, 1494. Second ed., Toscolano, 1523, under the title *Summa*, etc.

dash were introduced*. The two rival notations competed against each other on European territory for many years. The $\tilde{p}$ and $\tilde{m}$ never acquired a foot-hold in Germany. The German + and − gradually penetrated different parts of Europe. It is found in Scheubel's algebra† published in Paris in 1551, in Robert Recorde's *Whetstone of Witte* published at London in 1557 and in the algebra of C. Clavius published at Rome in 1608. In Spain the German signs occur in a book of 1552‡, only to be superseded by the $\tilde{p}$ and $\tilde{m}$ in later algebras§ of the sixteenth century. The struggle lasted about one hundred and thirty years when the German signs won out everywhere except in Spain. Organized effort, in a few years, could have ended this more than a century competition.

If one takes a cross-section of the notations for radical expressions as they existed in algebra at the close of the sixteenth century, one finds four fundamental symbols for indicating roots, the letters R and l, the radical sign $\surd$ proper, and the fractional exponent. The letters R and l were sometimes used as capitals and sometimes as small letters||. The student had to watch his step, for at times these letters were used to mark, not roots, but the unknown quantity x and, perhaps, also its powers¶. When R stood for "root", it became necessary to show whether the root of one term or of several terms was meant. There sprang up at least six different symbols for the aggregation of terms affected by the R, namely one of L. Pacioli**, two of H. Cardan††, the round

*J. Widman: *Behēde vnd hubsche Rechnung*. Leipzig, 1489. Later editions have other forms of title, and appeared at Pforzheim, 1500 and 1508; at Hagenau, 1519; at Augsburg, 1526.

†J. Scheubel: *Algebrae compendiosa facilisqve descriptio*. Paris, 1551.

‡*Libro primero de Arithmetica Algebratica* . . . por Marco Aurel, natural Aleman. Valencia, 1552. *See* J. Rey Pastor's *Los matemáticos españoles del siglo XVI*. Oviedo, 1913, p. 36 foot-note.

§J. Perez de Moya: *Arithmetica practica, y especvlativa*. Salamanca, 1562. This work appeared at Madrid in the fourteenth impression, in the year 1784, with no variation from the original edition except occasional modernization of the spelling; the fifteenth century symbols p, m, co., ce., cu. still occur in the impression of 1784. Another Spanish sixteenth century work was Antich Rocha's *Arithmetica*, Barcelona, 1564, also 1565.

||The letter R usually occurs as a capital, but sometimes is found as a small letter, for instance, in L. Pacioli's *Summa*, Part I, 1523, folio 86A. On the other hand the letter l usually occurs as a small letter, but is given as a capital, for instance, by Guillaume Gosselin in his *De arte magna*, Paris, 1577.

¶The use of the letter ℞ to signify root is seen in Leonardo of Pisa's *Practica geometriae* (1220), printed in Scritti di Leonardo Pisano, edited by B. Boncompagni, Rome, Vol. II, 1862, p. 209. The use of the word "radix" to signify x is found in Leonardo of Pisa's *Liber abbaci*, edited by B. Boncompagni, Rome, 1857, Vol. I, p. 407. The use of ℞ for root is common in L. Pacioli's *Summa* of 1494 and 1523; the sign ℞ for x and its powers occurs in the *Summa*, Part I, folio 67B. The representation of our x by ℞ is found also in J. Peletier's *Algebra*, Paris, 1554, bk. I, p. 8. The use of the letter l as a radical sign and also for x is explained by L. Schöner, a follower of Ramus, in the *Petri Rami arithmetices libri duo, et algebrae totidem: à Lazaro Schonero*, Frankfurt, 1592, and by Bernardus Salignacus, another follower of Ramus, in his *Algebrae libri duo*, Frankfurt, 1580. *See* P. Treutlein in *Abhandlungen zur Geschichte der Mathematik*, Vol. 2, 1879, p. 36.

**L. Pacioli marks by "℞. v." (radix vniversalis) that the root of two or more terms is indicated; thus, in his *Summa*, Part I, folio, 149A, "℞ v. ℞. $20\frac{1}{4}$. $\tilde{m}$. $\frac{1}{2}$. " stands for $\sqrt{\sqrt{20\frac{1}{4}}-\frac{1}{2}}$.

††H. Cardan in his *Practica arithmeticae* as printed in *Opera*, Vol. 4, 1663, pp. 14 and 16 lets "℞, v." stand for "radix vniversalis", but once writes also "(℞) 13 . $\tilde{p}$℞9" for $\sqrt{13+\sqrt{9}}$. He introduces the "radix ligata" to express the roots of each of the terms of a binomial; he writes (p. 16) "*L*℞7$\tilde{p}$℞. 10 "for $\sqrt{7}+\sqrt{10}$. This "*L*" would seem superfluous, but was introduced to dis-

parenthesis of N. Tartaglia*, the upright and inverted letter L of R. Bombelli† and the "r bin." and "r trinomia" of A. Romanus‡. There were at least five ways of marking the orders of the root, those of N. Chuquet§, E. De la Roche||, L. Pacioli¶, Ghaligai** and H. Cardan††. Thus the letter ℞ carried with it at least fourteen varieties of usage. In connection with the letter l, signifying latus or root, there were at least four ways of designating the orders of the roots and the aggregation of terms affected‡‡.

A unique line symbolism for roots of different orders occurs in manuscripts of J. Napier§§.

tinguish between the above form and the "radix distincta", as in "℞. D. 9 $\tilde{p}$ ℞ 4", which signifies 3 and 2, taken separately. Accordingly, "℞. D. 4 $\tilde{p}$. ℞. 9", multiplied into itself gives $4+9=13$, while the "℞. L. 4 $\tilde{p}$. ℞. 9", multiplied into itself, gives $13+\sqrt{144}=25$.

*N. Tartaglia, in *La quarta parte del general trattato de nvmeri et misvre*, Venetia; 1560, folio 40B, writes in the running text "℞ v. (72. men ℞ 1036$\frac{4}{5}$)" for $\sqrt{72-\sqrt{1036\frac{4}{5}}}$, but on the margin the second half of the parenthesis is omitted. In some other expressions (folio 41A) the final part of the parenthesis is omitted both in the text and on the margin. The use of "V." and also of the parenthesis marks a redundancy.

†Rafael Bombelli, in his *L'Algebra*, Bologna, 1572, uses almost always "radix relata" for the root affecting only one term. In this respect he followed Cardan. To unite two or more terms into one term, Bombelli placed L right after the ℞, and an inverted ⅃ at the end of the expression to be radicated. Thus he wrote "℞$L7$ p. ℞. 14" for our modern $\sqrt{7+\sqrt{14}}$. In the manuscript of his algebra, prepared some twenty years before its publication, aggregation is indicated differently, namely by the use of horizontal lines with cross-bars at the ends; the lines are placed below the terms to be united, as was the case in Chuquet.

‡A. Romanus, in *Ideae mathematicae pars prima*, Antwerp, 1593, after the preface, writes "r bin. $2+r$ bin. $2-r$ bin. $2-r$ 2" to designate $\sqrt{2+\sqrt{2-\sqrt{2-\sqrt{2}}}}$.

§N. Chuquet's *Le Triparty*, etc., 1484 (B. Boncompagni's *Bullettino*, Vol. 13, p. 755), indicates the second, third, fourth roots, etc., by ℞^{2}, ℞^{3}, ℞^{4}, . . .

||De la Roche, *Larismethique*, Lyon, 1520, 2. ed., 1538, uses Chuquet's notation and also a second notation, ℞□ for cube root, ℞℞ for the fourth root, and ℞℞□ for the sixth root.

¶L. Pacioli indicated the orders of roots in his *Summa*, Part I, as in these examples: (Folio 70B) "℞. 200." for $\sqrt{200}$; (Folio 119B) "℞. cuba de. 64." for $\sqrt[3]{64}$; (Folio 131A) "℞℞. 120." for $\sqrt[4]{120}$; (Folio 182A) "℞. relata." for the fifth root; (Folio 182A) "℞. cuba. de. ℞. cuba." for the sixth root; (Folio 182A) "℞℞℞. cuba." for the seventh root.

**Fr. Ghaligai in *Practica d'arithmetica*, Firenze, 1552 (first edition 1521), writes (folio 74A) "3℞□10" for our $\sqrt[3]{10}$; (folio 77B) "℞ di ℞ di 16" for our $\sqrt[4]{16}$; (folio 81B) "℞ ▭ di 2048" for our $\sqrt[3]{2048}$; (folio 81B) "℞ ▭ di □ 729" for our $\sqrt[6]{729}$; (folio 82A) "℞ □ di 72 ⊟ "for our $\sqrt{72x^5}$.

††H. Cardano in his *Ars magna*, as reprinted in *Opera*, Vol. 4, 1663, p. 255, indicates the cube root of a polynomial by writing on its left "℞. v. cub.", or "℞. v. cubica", or "℞. v. cu.". The sign "℞℞℞. 3", on p. 303, indicates $\sqrt[8]{3}$, not $\sqrt[6]{3}$ as it would in Pacioli.

‡‡P. Rami, *Scholarvm mathematicarvm libri unus et triginta*, Basel, 1569, lib. XXIV, pp. 276, 277, gives l 27 for $\sqrt{27}$; ll 32 for $\sqrt[4]{32}$; also (page 288) lr. $l112-l76$ for $\sqrt{\sqrt{112}-\sqrt{76}}$, the r. signifying "residu a", remainder, and therefore "lr." signifying the square root of the binomial difference. In P. Rami *arithmetices libri duo, et algebrae totidem . . . à L. Schonero*, Frankfurt, 1592, p. 272 ff. one finds l_c 4 for $\sqrt[3]{4}$; l_{bq} 5 for $\sqrt[4]{5}$. G. Gosselin, in his *De Arte magna*, Paris, 1577, marks the second, third and fourth roots, L, LC, LL; he writes $\sqrt{24+\sqrt{29}}$ thus: $LV24$ P $L29$.

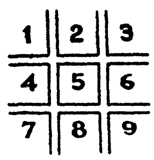
§§Napier's line symbolism for radicals is explained in *De arte logistica Joannis Naperi Merchistonii Baronis libri supersunt*. Edinburgh, 1839, p. 84, where ⊔ prefixed to a number meant

The radical signs for cube and fourth root had quite different shapes as used by C. Rudolff* and M. Stifel†. Though clumsier than Stifel's, the signs of Rudolff retained their place in some books for over a century‡. To designate the orders of the roots, M. Stifel placed immediately after the radical sign the German abbreviations of the words zensus, cubus, zensizensus, sursolidus, etc. S. Stevin§ made the important innovation of numeral indices. He placed them within a circle. Thus he marked cube root by a radical sign followed by the numeral 3 coralled in a circle. To mark the root of an aggregation of terms, C. Rudolff|| introduced the dot placed after the radical sign; sometimes two dots were used, one before the expression, the other after it¶. Stevin** and Digges†† had still different designations. Thus the radical sign carried with it seven somewhat different styles of representation. Stevin suggested also the possibility of fractional exponents‡‡, the fraction being placed inside a circle, and before the radicand.

Thus altogether there were at the close of the sixteenth century twenty-seven or more varieties of symbols for the calculus of radicals with which the student had to be familiar, if he desired to survey the publications of his time.

Decimal fractions were invented during the latter part of the sixteenth

square root, ⊐ its fourth root, □ its fifth root, ⌐ its ninth root. Napier's *notation* was derived from the figure



*C. Rudolff, in his *Behend vnnd Hubsch Rechnung . . . Coss*, Strassburg, 1525, lets $\surd$ be the square root, $\surd\!\!\surd\!\!\surd$ cube root, $\surd\!\!\surd$ fourth root.

†M. Stifel, in his *Arithmetica integra*, Nürnberg, 1544, lets $\surd z$ stand for square root, $\surd C^e$ for cube root, $\surd zz$ for fourth root, $\surd\beta$ for the fifth root, $\surd zC^e$ for the sixth root, etc.

‡Over a century after Rudolff's *Coss* of 1525, the Swiss J. Ardüser, in his *Geometriae theoricae et practicae XII. Bücher*, Zürich, 1627, folio 81A, used Rudolff's signs for square and cube root. J. Ozanam, in his *Nouveaux elemens d'algebre*, I^e Partie, Amsterdam, 1702, p. 82, still employs Rudolff's sign for the fourth root.

§With Stevin $\surd$③ was cube root, $\surd$④ or $\surd\!\!\surd$ fourth root, $\surd\!\!\surd$⑨ ninth root. *Oeuvres mathématiques*, edited by A. Girard, Leyden, 1634, p. 19.

||C. Rudolff, *Coss* 1525, p. 141, lets " $\surd.\,12+\surd140$" represent $\sqrt{12+\sqrt{140}}$.

¶M. Stifel in his *Arithmetica integra*, folio 135B uses two dots when one dot alone would introduce an ambiguity.

**S. Stevin, *Oeuvres*, Leyden, 1634, p. 19, wrote " $\surd$ bino $2+\surd3$" for $\sqrt{2+\sqrt{3}}$, "$\surd$ trino $\surd3+\surd2-\surd5$" for $\sqrt{\sqrt{3}+\sqrt{2}-\sqrt{5}}$; "quadrin" to mark the root of a quadrinomial.

††Leonard and Thomas Digges in a *Geometrical Practize, named Pantometria*, London, 1571 [pages unnumbered], write " $\surd z.\,v.\,98+\surd z\,1920\frac{4}{5}$" for our $\sqrt{98+\sqrt{1920\frac{4}{5}}}$. In the 1591 edition of the Pantometria, p. 106, one finds " $\surd z$. vni." in place of " $\surd z$. v."

‡‡S. Stevin, *Oeuvres, Arithm.*, p. 6, says: "$\frac{3}{2}$ en un circle seroit le charactere de racine quarrée de ③, par ce que telle $\frac{3}{2}$ en circle multipliée en soy donne produict ③, et ainsi des autres."

century. About thirty-four different notations* for such fractions were suggested of which several have survived to the present day. After three hundred years we still have some half dozen notations. If a dot is placed between the numerals two and five, we cannot tell what the symbolism means unless we know whether the writing is that of an Englishman, Austrian, or American. The English notation for two and five-tenths means two times five in America, and *vice versa.* It does not make a straw's difference which is adopted; the important thing is to select one only, in order to avoid confusion. On a primitive question of this sort an international committee could reach an agreement almost instantaneously.

Our Recordean sign of equality had at least half a dozen rivals, and when about one hundred and twenty years old it came very near perishing in the struggle with the rival sign proposed by R. Descartes†. It took a hundred and fifty years to determine the question of supremacy of different symbols for equality in writing ordinary equations‡. In marking proportions the struggle is still on. In England and America the four dots used in proportion by W. Oughtred, though obsolescent, are still found in some texts. After nearly three centuries this easy question remains unsettled, because there is no supreme court to render a decision.

One of the most extraordinary circumstances is that after two and a half centuries, there is still complete diversity of usage between the English speaking people and other peoples, on the symbol for division§.

One might think that such a simple matter as the marking of the hyperbolic,

*S. Stevin in 1585 gives 32⓪5①7② and $3\,\overset{\textcircled{0}}{2}\,\overset{\textcircled{1}}{5}\,\overset{\textcircled{2}}{7}$ for 32.57; Robert Norton in 1608 gives $3^{(1)}7^{(2)}5^{(3)}9^{(4)}$ for .3759; W. von Kalcheim in 1629 gives 693② for 6.93; J. H. Beyer in 1603 gives $14.\overset{\text{VIII}}{3}761$ for 14.00003761, also $12\overset{0}{3}\,.\,\overset{\text{I}}{4}\,.\,\overset{\text{II}}{5}\,.\,\overset{\text{III}}{9}\,.\,\overset{\text{IV}}{8}\,.\,\overset{\text{V}}{7}\,.\,\overset{\text{VI}}{2}\,.$ for 123.459872; F. Vieta in 1579 writes the decimal part in smaller type or else he places a vertical stroke between units and tenths; Kepler in 1616 writes 3(65; Joost Bürgi in 1620 gives 230270022 for 230270.022; Pitiscus in 1612 gives .05176381 for .5176381; also 02679492 for 0.2679492, also 5|269 for 5.269; J. Napier in 1617 wrote 1993,273 for 1993.273, his *Constructio* of 1620 contains our notation 25.803; J. Johnson in 1623 gives $3\,|\,\overset{1.}{2}\;\overset{2.}{2}\;\overset{3.}{9}\;\overset{4.}{1}\;\overset{5.}{6}$ for 3.22916; H. Briggs in 1624 wrote $5^{\underline{9321}}$ for 5.9321; W. Oughtred in 1631 gave $2|\underline{5}$; R. Jager in 1651 gave $16|\,\dot{7}\,\dot{2}\,\dot{4}\,\dot{9}$ for 16.7249; R. Balaam in 1653 wrote 3:04 for 3.04; Fr. van Schooten in 1657 gives 58,5① for 58.5; A. Tacquet in 1665 wrote $25.\overset{\text{I}}{8}\;\overset{\text{II}}{0}\;\overset{\text{III}}{0}\;\overset{\text{IV}}{7}\;\overset{\text{V}}{9}$; S. Foster in 1657 wrote $31.^{\underline{008}}$ for 31.008; J. Caramuel in 1670 adopts 22=3 for 22.3; C. F. M. Dechales in 1690 writes 12[345 for 12.345; S. Jeake states that 34 . 1 . 4 . 2 . 6 . is sometimes written for 34.1426; J. Whiston in 1707 writes 0;9985 for 0.9985; J. Raphson in 1728 puts $732,|\underline{569}$ for 732.569; Sherwin in 1741 gives (4)2677 for .00002677; Newton's *Arithmetica universalis*, 1707, contains 732′|569 for 732.569; Horsley's edition of *Newton's Arithmetica universalis*, 1799, contains 35′72 for 35.72; H. Clarke in 1777 gives 2˙5 for 2.5; Jonas Moore in 1688 writes 116 ⌒64 for 116.64; A. F. Vallin in 1889 has 2,5 for 2.5; in Scandinavian books of the present time, $2._5$ and $2,_5$ are found.

†René Descartes in his *La Géométrie* of 1637 gives ∞ as a sign of equality.

‡See F. Cajori's "Mathematical Signs of Equality", Isis, Vol. V, 1923, pp. 116-125.

§The sign ÷ for division, now used by English-speaking people, first occurs in J. H. Rahn's *Teutsche Algebra*, 1659; the sign : now used for division on the European continent, was suggested by G. W. Leibniz in Acta eruditorum, Leipzig, 1684, p. 470. See C. I. Gerhardt, *Leibnizens mathematische Schriften*, Vol. 5, Halle, 1858, p. 223.

functions would not give rise to a diversity of notation; but since the time of Vincenzo Riccati, there have been not less than nine distinct notations*. No better is our record for the inverse trigonometric functions. For these there were at least ten varieties of symbols†. Two varieties, namely J. Herschel's sign looking like exponent minus one, as in $\sin^{-1}x$, and the "arc sin", "arc cos", etc., of the continent, still await the decision as to which shall be universally adopted.

The second half of the eighteenth century saw much experimentation in notations of the calculus, on the European Continent. There was a great scramble for the use of the letters D, d, and ∂. The words "difference", "differential", "derivative" and "derivation" all began with the same letter. A whole century passed before any general agreement was reached. Particular difficulty in reaching uniformity prevailed in the symbolism for partial derivatives and partial differentials. A survey has revealed thirty-five different varieties of notation for partial derivatives‡. For about one hundred and fifty years, experimentation was carried on before some semblance of uniformity was attained. Such uniformity could have been reached in fifteen years, instead of one hundred and fifty, under an efficient organization of mathematicians.

In vector analysis, a subject most vital in mathematical physics, there is lack of uniformity at the very start in the representation of a vector. The diversity of symbolism increases when one passes to the different kinds of pro-

*In the *Institvtiones analyticae a Vincentio Riccato . . . et Hieronymo Saladino*, Vol. 2, Bologna 1767, p. 152, Riccati writes Shx, Chx for hyperbolic sin x and cos x. In the *Histoire de l'Académie de Berlin*, 1768, Vol. 24, p. 327, J. H. Lambert gives "sin h", "cos h". L'Abbé Saurin in his *Cours complet de mathématique*, Tome IV, Paris, 1774, pp. 222, 223, writes s. h, c. h, t. h, cot. h. C. Gudermann, in Crelle's Journal, Vol. 4, 1829, p. 295, adopts German type for Cos, Sin, Tang, Cot, Arc Cos, Arc Tang, Arc Cot, but states that Latin forms may be used also. Joh. Frischauf, in his *Absolute Geometrie*, Leipzig, 1876, p. 54, prefers small letters of German black-faced type. W. B. Smith, in his *Infinitesimal Analysis*, New York, Vol. I, 1898, pp. 87, 88, writes hs x, hc x, ht x, hsc x, hcsc x. Prefixing the syllable "hy" was suggested by G. M. Minchin, in Nature, Vol. 65, London, 1902, p. 531.

†In Commentarii academiae Petropolitanae, 1737, Vol. 9, p. 209, L. Euler wrote "$A \sin\frac{b}{c}$" for the arc whose sine is $\frac{b}{c}$. J. H. Lambert, in Acta Helvetica, Vol. III, Basel, 1758, p. 141, gives in full "arcus sinui b". C. Scherffer in his *Institutionum analyticum pars secunda*, Vienna, 1772, p. 144, gives "arc. tang." Bérard in Annales de Gergonne, Vol. 4, 1813-14, p. 263, writes "$\text{Arc.}\left(\text{Tang.} = \frac{dy}{dx}\right)$". J. Herschel, in the Philosophical Transactions, London, for the year 1813, p. 10, introduces $\sin^{-1}x$, $\tan^{-1}x$, etc. Martin Ohm, in his *System der Mathematik*, Vol. 2, Berlin, 1829, p. 372, gives $\frac{1}{\text{Sin}}y, \frac{1}{\text{Cos}}y$. John West in *Mathematical Treatises*, edited by John Leslie, Edinburgh, 1833, p. 237, writes "$\text{arc}\left(\tan.\frac{x}{c}\right)$". F. D. Covarrubias in his *Elementos de análisis*, 2. ed., Mexico, 1890, p. 48, gives "arco (sec$=x$)". Maurice Godefroy in his *Théorie des Séries*, Paris, 1903, p. 192, gives "Arg ch x" for the argument whose hyperbolic cosine is x. Several recent texts prefer a union of the syllables, like "arctan".

‡See F. Cajori, *History of Notations of the Calculus* in Annals of Mathematics, 2. S., Vol. 25, pp. 16-35.

ducts of two or more vectors*. It is difficult to find two prominent writers who use the same notation, except in cases where the two stand in the relation of teacher and pupil. Three attempts at unification have been made, one by an International Association organized in 1895 for the study of vector analysis, the second by a commission which Felix Klein appointed in 1903, the third a Commission created in 1908, at the International Congress at Rome. The International Association and the Commission of 1908 have not reached results, partly because of the death of some of the leaders and partly because of the Great War. The Commission of 1903, instead of reaching an agreement on some system of notation, added by their labours three new systems to those previously existing†. The primitive source of failure seemed to be the attempt on the part of each individual to secure a *perfect notation*, which was, of course, the one he himself proposed, rather than to *reach an agreement* on a notation. Years ago physicists encountered a similar problem in the selection of electrical units. But they were wiser than the mathematicians. Laying aside the immediate attainment of the ideal and resting contented with the best compromise securable, they established a system of theoretical and practical units which has aided all the wonderful modern advances along theoretical and practical lines. Is it not possible for the mathematicians to equal the performance of the physicists?

It would be easy to advance many other instances of "muddling along" through decades without getting together and agreeing upon a common sign language. Cases are found in descriptive geometry, the modern geometry of the triangle and circle, elliptic functions, the theory of potential and elasticity. An investigator wishing to familiarize himself with the researches of other workers in his own field must be able not only to read several modern languages, English, French, Italian and German, but he must also master the various notations employed by different authors. The task is unnecessarily strenuous‡.

A French cynic once said: "Language is for the suppression of thought." Certainly the diversity of language tends to the isolation of thought. The existence of heterogeneous notations has been frequently deplored by mathematicians. In analysis C. Babbage§ speaks of "a profusion of notations (when we regard the whole science) which threaten, if not duly corrected, to multiply our difficulties instead of promoting our progress." A committee composed of

*For references to the different notations which have been proposed in Vector Analysis, see C. Burali-Forti and R. Marcolongo in Rend. Circ. Mat. Palermo, Vol. 23, p. 324, 1907; Vol. 24, pp. 65, 318; Vol. 25, p. 352; Vol. 26, p. 369; C. Burali-Forti in l'Enseignement mathématique, Vol. 11, 1909, p. 50; Jahresbericht d. d. Mathematiker Vereinigung, Vol. 13, 1904, pp. 39, 222, 229; Vol. 14, 1905, p. 211; J. B. Shaw's paper in International Association for the Study of Quaternions and Allied Systems of Mathematics, Lancaster, Pa., 1912; *Unification des notations vectorielles* in Isis, Vol. 2, pp. 173-182.

†F. Klein, *Elementarmathematik vom höreren Standpunkte aus*, Teil I, Leipzig, 1908, p. 157.

‡The fundamental agreements were reached at the International Electrical Congress held at Paris in 1881, and at the International Conference at Paris in 1882. See address by Lord Kelvin, Nature, Vol. 28, 1883, p. 91; H. T. Eddy in Science, Vol. I, 1883, p. 87; H. S. Carhart in Science, Vol. 21, 1893, pp. 86, 87.

§C. Babbage *On the Influence of Signs in Mathematical Reasoning* in Trans. Cam. Phil. Soc., Vol. II, 1827, p. 326.

W. Spottiswoode, G. G. Stokes, A. Cayley, W. K. Clifford and J. W. L. Glaisher conservatively said*: "Anything which tends towards uniformity in notation may be said to tend towards a common language in mathematics, and . . . must ultimately assist in disseminating a knowledge of the science of which they treat." These representative men saw full well that God does not absolve the mathematician from the need of the most economic application of his energies and from indifference to the well-tried wisdom of the ages.

Our mathematical sign language is still heterogeneous and sometimes contradictory. And yet, whatever it lacks appears to be supplied by the spirit of the mathematician. The defect of his language has been compensated by the keenness of his insight and the sublimity of his devotion. It is hardly worth while to indulge in speculation as to how much more might have been achieved with greater symbolic uniformity. With full knowledge of the past it is more to the point to contemplate the increasingly brilliant progress that may become possible when mathematicians readdress themselves to the task of breaking the infatuation of extreme individualism on a matter intrinsically communistic, when mathematicians learn to organize, to appoint strong and representative international committees whose duty it shall be to pass on the general adoption of new symbols and the rejection of outgrown symbols, when in their publications mathematicians, by a gentlemen's agreement, shall abide by the decisions of such committees.

*British Association Report for 1875, p. 337.

PAST STRUGGLES BETWEEN SYMBOLISTS AND RHETORICIANS IN MATHEMATICAL PUBLICATIONS

BY PROFESSOR FLORIAN CAJORI,
Univerity of California, Berkeley, California, U.S.A.

For many centuries there has been a conflict between individual judgments, on the use of mathematical symbols. On the one side are those who, in geometry, for instance, would employ hardly any mathematical symbols; on the other side are those who insist on the use of ideographs and pictographs almost to the exclusion of ordinary writing. The real merits or defects of the two extreme views cannot be ascertained by *a priori* argument; they rest upon experience and must therefore be sought in the study of the history of our science.

The first printed edition of Euclid's elements and the earliest translations of Arabic algebras into Latin contained little or no mathematical symbolism.* During the Renaissance the need of symbolism disclosed itself more strongly in algebra than in geometry. During the sixteenth century European algebra developed symbolisms for the writing of equations, but the arguments and explanations of the various steps in a solution were written in the ordinary form of verbal expression.

The seventeenth century witnessed new departures; the symbolic language of mathematics displaced verbal writing to a much greater extent than formerly. The movement is exhibited in the writings of three men, Pierre Hérigone† in France, William Oughtred‡ in England, and J. H. Rahn§ in Switzerland. Hérigone used in his Cursus mathematicus of 1634 a large array of new symbols of his own design. He says in his Preface: "I have invented a new method of making demonstrations, brief and intelligible, without the use of any language." In England, William Oughtred used over one hundred and fifty mathematical symbols, many of his own invention. In geometry Oughtred showed an even

*Erhard Ratdolt's print of Campanus' Euclid, Venice, 1482. Al-Khowarizmi's algebra was translated into Latin by Gerard of Cremona in the twelfth century. It was probably this translation that was printed in Libri's *Histoire des sciences mathématiques en Italie*, Vol. I, Paris, 1838, pp. 253-297. Another translation into Latin, made by Robert of Chester, was edited by L. C. Karpinski, New York, 1915. Regarding translations of Al-Khowarizmi, see also G. Eneström, *Bibliotheca mathematica*, 3 S., Vol. 5, 1904, p. 404; A. A. Björnbo, *Bibliotheca mathematica*, 3 S., Vol. 6, 1905, pp. 239-248.

†Pierre Hérigone: *Cursus Mathematicus*, Vol. I-IV, Paris, 1634; 2. Ed., 1644.

‡W. Oughtred: *Clavis mathematicae*, London, 1631, and later editions, also Oughtred's *Circles of Proportion*, 1632; *Trigonometrie*, 1657 and minor works.

§J. H. Rahn: *Teutsche Algebra*, Zürich, 1659. Thomas Brancker: *An Introduction to Algebra*, translated out of the High-Dutch, London, 1668.

greater tendency to introduce extensive symbolism than did Hérigone. Oughtred translated the tenth book of Euclid's Elements into language largely ideographic, using for the purpose about forty new symbols*. Some of his readers complained of the excessive brevity and compactness of the exposition, but Oughtred never relented. He found in John Wallis an enthusiastic disciple. At the time of Wallis, representatives of the two schools of mathematical exposition came into open conflict. In treating the Conic Sections†, no one before Wallis had employed such an amount of symbolism. The philosopher Thomas Hobbes protests emphatically‡: "And for . . . your Conic Sections, it is so covered over with the scab of symbols, that I had not the patience to examine whether it be well or ill demonstrated." Again Hobbes says§: "Symbols are poor, unhandsome, though necessary, scaffolds of demonstration," and ought not to "appear in public." Hobbes explains further||: ". . . Symbols, though they shorten the writing, yet they do not make the reader understand it sooner than if it were written in words. For the conception of the lines and figures . . . must proceed from words either spoken or thought upon. So that there is a double labour of the mind, one to reduce your symbols to words, which are also symbols, another to attend to the ideas which they signify. Besides, if you but consider how none of the ancients ever used any of them in their published demonstrations of geometry, nor in their books of arithmetic . . ., you will not, I think, for the future be so much in love with them. . . ." Whether there is really a double translation, such as Hobbes claims, and also a double labour of interpretation, is a matter to be determined by experience.

Meanwhile the Algebra of J. H. Rahn appeared in 1659 in Zurich and was translated by Brancker into English and published with additions by John Pell, at London, in 1668. The work contained some new symbols and also Pell's division of the page into three columns. He marked the successive steps in the solution so that all steps in the process are made evident through the aid of symbols, hardly a word of verbal explanation being necessary. In Switzerland the three column arrangement of the page did not receive an enthusiastic reception. In Great Britain it was adopted in a few texts, John Ward's *Young Mathematician's Guide*, parts of John Wallis' *Treatise of Algebra* and John Kirby's *Arithmetical Institutions*. But this almost complete repression of verbal explanation did not become widely and permanently popular.

In the great mathematical works of the seventeenth century—the *Géométrie* of Descartes, the writings of Pascal, Fermat, Leibniz, the *Principia* of Sir Isaac Newton—symbolism was used in moderation.

The struggle in elementary geometry was more intense. The notations of Oughtred also met with a friendly reception from Isaac Barrow, the teacher of Sir Isaac Newton, who followed Oughtred even more closely than did Wallis.

*Printed in *Oughtred's Clavis mathematicae*, 3rd Edition, 1648, and in the editions of 1652, 1667, and 1693.

†J. Wallis: *Operum mathematicorum*, *Pars altera*, Oxford, De sectionibus Conicis, 1656.

‡The English works of T. Hobbes, by Sir William Molesworth, Vol. 7, London, 1845, p. 316.

§*Loc. cit.*, p. 248.

||*Loc. cit.*, p. 329.

In 1655, Barrow brought out an edition of Euclid in Latin and in 1660 an English edition. He had in mind two main objects: first, to reduce the whole of the Elements into a portable volume and, secondly, to gratify those readers who prefer "symbolical" to "verbal reasoning." During the next half century Barrow's texts were tried out. In 1713, John Keill of Oxford edited the *Elements of Euclid*, in the preface of which he criticized Barrow, saying: "Barrow's demonstrations are so very short, and are involved in so many notes and symbols, that they are rendered obscure and difficult to one not versed in Geometry. There, many propositions, which appear conspicuous in reading Euclid himself, are made knotty, and scarcely intelligible to Learners, by his Algebraical way of demonstration. . . . The Elements of all Sciences ought to be handled after the most simple Method, and not to be involved in Symbols, Notes, or obscure Principles, taken elsewhere." Keill abstains altogether from the use of symbols. His exposition is quite rhetorical.

William Whiston, who was Newton's successor in the Lucasian professorship at Cambridge, brought out a school Euclid (an edition of Tacquet's Euclid) which contains only a limited amount of symbolism. A more liberal amount of sign language is found in the *Elements of Geometry* of William Emerson, London, 1763.

Robert Simson's edition of Euclid appeared in 1756. It was a carefully edited book and attained a wide reputation. Ambitious to present Euclid unmodified, he was careful to avoid all mathematical signs. The sight of this book would have delighted Hobbes. No scab of symbols here!

That a reaction to Simson's Euclid would follow was easy to see. In 1795 John Playfair of Edinburgh brought out a school edition of Euclid which contains a limited number of symbols, and in 1831 R. Blakelock of Cambridge edited Simson's text in the symbolic form*. Oliver Byrne's Euclid in symbols and coloured diagrams was not taken seriously, but was regarded as a curiosity†. The Senate House examinations discouraged the use of symbols‡. Later DeMorgan wrote§: "Those who introduce algebraical symbols into elementary geometry, destroy the peculiar character of the latter to every student who has any mechanical association connected with those symbols; that is, to every student who has previously used them in ordinary algebra. Geometrical reasons, and arithmetical process, have each its own office; to mix the two in elementary instruction, is injurious to the proper acquisition of both."

The same idea is embodied in Todhunter's edition of Euclid which does not contain even a plus or minus sign, nor a symbolism for proportion.

*Another text of that period, the *Supplement to the Elements of Euclid*, by D. Cresswell, 2. ed., Cambridge, 1825, uses algebraic symbols and pictographs.

†O. Byrne: *The Elements of Euclid in which coloured diagrams and symbols are used.* London, 1847.

‡In *The Elements of Euclid . . . from the Text of Dr. Simson . . . introduction of Symbols.* By a Member of the University of Cambridge, London, 1827, one finds on page 104 the symbol ∽ for "is similar to", but the student is told that "in writing out the propositions in the Senate House, Cambridge, it will be advisable not to make use of this symbol, but merely to write the word short, thus *is sim.*" Moreover in the Preface it is stated that "more competent judges than the editor" advised that the symbol be eliminated, and so it was, except in one or two instances where "it was too late to make the alteration", the sheets having already been printed.

§A. De Morgan: *Trigonometry and Double Algebra*, 1849, p. 92, foot notes.

The view point of the opposition is expressed by a writer in the London Quarterly Journal of 1864: "The amount of relief which has been obtained by the simple expedient of applying to the elements of geometry algebraic notation can be told only by those who remember to have painfully pored over the old editions of Simson's Euclid. The practical effect of this is to make a complicated train of reasoning at once intelligible to the eye, though the mind could not take it in without effort."

English geometries of the latter part of the nineteenth century and of the present time contain a moderate amount of symbolism. The extremes as represented by Oughtred and Barrow on the one hand and by Simson on the other, are avoided. Thus a conflict in England, lasting two hundred and fifty years, has ended as a draw. It is a stupendous object-lesson to mathematicians on mathematical symbolism. It is the victory of the golden mean.

The movements on the Continent were along the same lines, but were less spectacular than in England. In France, about a century after Hérigone, Clairaut* used in his geometry no algebraic signs and no pictographs. Bézout† and Legendre‡ employed only a moderate amount of algebraic signs. In Germany, Karsten§ and Segner|| made only moderate use of symbols in geometry, but S. Reyher¶ and J. F. Lorenz** used extensive notation; Lorenz brought out a very compact edition of all books of Euclid's Elements.

The great mathematicians of the eighteenth century, Euler, Lagrange, Laplace, used symbolism freely, but expressed much of their reasoning in ordinary language. In the nineteenth century, one finds in the field of logic all gradations from no symbolism to nothing but symbolism††. The well-known opposition of Steiner to Plücker touches the question of sign-language. Some modern works on geometry make very restricted use of algebraic symbols. On the other hand, there are prominent investigators who would practically dispense with ordinary language and "express all propositions of mathematics by means of a small number of signs"‡‡ or express "mathematical propositions"§§, including

*A. C. Clairaut: *Elémens de géométrie*, Paris, 1753. First edition, 1741.

†E. Bézout: *Cours de mathématiques*, T. I, Paris, N. Ed., 1797, *Elémens de géométrie.*

‡A. M. Legendre: *Eléments de Géométrie*, Paris, 1794.

§W. J. G. Karsten: *Lehrbegrif der gesamten Mathematik*, I. Theil, Greifswald, 1767, pp. 205-484.

||I. A. de Segner: *Cursus mathematici*, Pars I: Elementa arithmeticae, geometriae et calculi geometrici, editio nova, Halle, 1767.

¶Samuel Reyhers . . . *Euclides*, Kiel, 1698.

**J. F. Lorenz: *Euklid's Elemente, auf's neue herausgegeben von C. B. Mollweide*, 5. Ed., Halle, 1824. First edition, 1781; 2. Ed., 1798.

††A. De Morgan: *Syllabus of a Proposed System of Logic*, London, 1860, p. 72. See also references in the two following foot-notes.

‡‡See G. Peano: *Formulaire de mathématiques*, Tome 1, Turin, 1895. Introduction, p. 52.

§§A. N. Whitehead and B. Russell, *Principia mathematica*, Vol. 1, Cambridge, 1910, Preface; Vol III, 1913, Preface. On this question see also E. H. Moore, *Introduction to a Form of General Analysis*, New Haven, 1910 (The New Haven Mathematical Colloquium), p. 150; A. Padoa: *La logique déductive dans sa dernière phase de developpement*, Paris, 1912; A. Voss, *Ueber das Wesen der Mathematik*, Leipzig und Berlin, 2 Aufl., 1913, p. 28; P. E. B. Jourdain, *The Logical Work of Leibnitz*, in Monist, Vol. 26, 1916, p. 515; P. E. B. Jourdain, *The Function of Symbolism in Mathematical Logic*, in Scientia, Vol. 21, 1917, pp. 1-12; H. F. Baker, *Principles of Geometry*, Vol. I, Cambridge, 1922, pp. 62, 137; Vol. II, pp. 4, 16, 153.

those of geometry, in the symbols of mathematical logic. The survey of modern periodicals indicates that the efforts of about fifteen, thirty and forty years ago to express all mathematics in ideographic symbolism is not meeting with the support of mathematicians in general at the present time; the efforts are hardly more successful than were similar ones in the earlier centuries.

Thus the experience of the past points to conservatism in the use of symbols. The rank and file of mathematicians have not taken kindly to a lavish use of signs. The inspiration of genius has led to a large output of new symbols intended to express all mathematics in purely ideographic form. But this is a phase in the development of our science on which the inspiration of genius is not sufficient; there must be brought to bear upon this problem the wisdom of many minds, and that wisdom discloses itself in the history of the science. New symbols are necessary as the science grows, but the experience of the past warns us against the too exclusive use of ideographic signs.

The conclusion reached in this address may be stated in terms of two schoolboy definitions of salt. One definition is, "Salt is what, if you spill a cup-full into the soup, spoils the soup." The other definition is, "Salt is what spoils your soup when you don't have any in it."

LA MEMORIA "DE INFINITIS HYPERBOLIS" DI TORRICELLI

DEL PROFESSORE ETTORE BORTOLOTTI,
R. Università di Bologna, Bologna, Italia.

1. Una delle scoperte geometriche che al loro apparire destarono più viva ammirazione e che maggiormente contribuirono a promuovere nuovi avanzamenti ed a dischiudere nuovi orizzonti alle ricerche geometriche, fu quella relativa alla misura finita di spazi generati dalla rotazione di figure infinitamente estese in lunghezza ed in superficie.

La proprietà di aver volume finito e misurabile, scoperta dal Torricelli negli ultimi mesi del 1641 per il solido acutissimo iperbolico generato dalla rotazione di un ramo di iperbola intorno ad un asintoto, e la misura esatta di questo volume, parvero cose meravigliose, divine, appena credibili, allo stesso Cavalieri maestro di ardimenti nel campo degli infiniti e creatore della *Geometria degli indivisibili.*

"Quel solido iperbolico infinitamente lungo (egli scriveva a Torricelli il 17 dicembre 1641) ed eguale ad un corpo, quanto a tutte e tre le dimensioni finito, mi è riuscito infinitamente ammirabile. Ed, avendolo io comunicato ad alcuni miei scolari filosofi, hanno confessato parergli veramente meraviglioso e stravagante che ciò possa essere, onde se la dimostrazione almeno del caso quando gli asintoti fanno angolo retto, potesse esser capita da loro, che hanno pur visto i sei primi libri di Euclide, mi saria caro poter loro dare in questa parte soddisfazione, oltre il gusto che ne avrei anch'io."*

Il Torricelli prontamente accondiscendeva al desiderio del Maestro; ma, anche dopo letta la dimostrazione, il Cavalieri pareva quasi non poter credere ai propri occhi, e rispondeva (7 gennaio 1642):

"La dimostrazione del solido acuto iperbolico è veramente divina, e non so come abbi pescato, nell' infinita profondità di quel solido, così facilmente la sua dimensione; poichè veramente a me pare infinitamente lungo, parendomi infinitamente lungo lo spazio piano che lo genera, ed ogni spazio di esso generando parte di solido. Sebbene non ci ho pensato molto, e potrei anco ingannarmi."

2. Questi dubbi non smossero certamente il Torricelli dalla sua ben fondata convinzione, poichè le proposizioni sul conoide acuto iperbolico, sono fra quelle che egli inviò ai matematici di Francia nel giugno 1643† e che al dire di Michel Angelo Ricci, "camminavano colà per le mani di tutti con molta lode delle belle invenzioni".

**Opere di Torricelli* (Faenza 1919), Vol. III, p. 65.
†*Op. Torricelli,* Vol. III, p. 129.

3. Nè il Fermat nè il Roberval risparmiarono al Torricelli le lodi più aperte. Ad emulazione della proposizione torricelliana il Fermat proponeva un problema sulla teoria dei numeri* (quod tuo de Conoideo acuto infinito aequivaleat) che solo ai nostri giorni si è saputo risolvere; ed il Roberval ricercava una sua propria dimostrazione della proposizione torricelliana (omnium elegantissima).†

Più tardi (nel 1646) lo stesso Roberval escogitava una trasformazione piana che, in certi casi, faceva corrispondere ad una figura finita di data area, un'altra figura ad essa equivalente ma dotata di punti impropri (infinitamente lunga).

Ed, in riscontro, il Torricelli, guidato dalla analogia col suo solido acuto iperbolico, francamente affrontava il problema diretto, di determinare le condizioni perchè una figura geometrica avente punti all'infinito avesse area (o volume) finita; cioè, nel senso moderno, di *determinare le condizioni di esistenza di integrali impropri estesi ad intervalli infiniti;* e lo risolveva completamente per le curve $x^m y^n = k$, e per i solidi generati dalla rotazione di esse intorno ad un asintoto.

4. Egli stesso ci racconta come giunse a questo fortunato scoprimento:

"Prima io dimostrai che infinite quantità in proporzione geometrica *majoris inequalitatis* sono misurabili. Poi mostrai che un solido infinitamente lungo era misurabile. E sapevo che descrivendo una figura piana le cui linee andassero continuamente decrescendo come i circoli del suddetto solido, anche quella era misurabile."‡

Le prime curve che per tal modo egli ebbe a considerare furono le iperbole spurie $x^2 y = k$; che immediatamente si presentano nell'interpretazione in figura piana del conoide iperbolico. Le altre infinite iperbole vennero poi quasi spontaneamente; anzi, come egli stesso ebbe a dire,§ lo tirarono contro sua voglia. E subito, con intuito geniale e con impeccabile raziocinio, egli sapeva trovare, per le condizioni di convergenza, quella "raison naturelle et immediate de cette merveilleuse propriété", che, 11 anni più tardi, il Wallis domandava al Fermat,|| quando questi si era secolui vantato di aver egli scoperto la quadratura delle infinite iperbole, e di averla egli stesso comunicata al Torricelli¶.

5. Ho rivendicato al Torricelli questa scoperta, con prove che ritengo inoppugnabili, in una memoria che può servire di introduzione alla presente,

**Ibid.*, p. 158, "Invenire triangulum rectangulum in numeris, cuius latus maius sit quadratum, summaque duorum aliorum laterum etiam sit quadratum, denique summa maioris et medij lateris sit etiam quadratum."

†*Ibid.*, p. 135.

‡*Op. di Torricelli*, Vol. 3°, p. 361 lettera a Michelangelo Ricci del 17 marzo, 1646.

§*Op. di Torricelli*, Vol. III, p. 373.

||*Œuvres de Fermat*, III (Paris, 1896), p. 408: "Cependant si votre très noble correspondant voulait bien indiquer, soit sa methode de quadrature des paraboles et des hyperboles, soit encore ce criterium, qui distingue dans les figures de ce genre les finies des infinies, j'entends la véritable raison de cette propriété, cela me ferait le plus grand plaisir. Car si ma méthode m'a suffi, comme je l'ai dit, pour ce même objet, je n'ai pas coutume d'avoir pour mes découvertes tant de prétension, ni tant de partialité non plus, que je croie pour elle devoir négliger celles des autres.

¶*Œuvres de Fermat*, II, Paris, 1894, p. 337-338.

pubblicata nell'ultimo fascicolo dell' *Archivio di Storia della scienza** ed ho anche fatto osservare che il Torricelli nelle sue lettere agli amici d'Italia ed agli scienziati di Francia, si pregiava, più ancora che delle scoperte da lui fatte, del metodo universalissimo che egli teneva per dimostrarle.

Ma non ho ancora fatto conoscere in che veramente consistesse questo suo metodo.

I.—Il Manoscritto Torricelliano e l'Edizione Faentina

6. Tutto occupato nel "lavorar occhiali", il Torricelli non trovava tempo per dar forma definitiva alle cose sue, le quali erano da lui affidate a certe *cartucce e fogliucci* con così poca cura, da *stimar più difficile il capire una scrittura tanto sporca che l'inventar di nuovo.*

Queste carte, dopo la morte improvvisa ed immatura di lui, furono pietosamente raccolte dal Serenai, che diligentemente le trascrisse, perchè nulla di esse andasse perduto. Ma, digiuno affatto di cose matematiche, il Serenai non seppe ordinare gli sparsi foglietti secondo la successione logica degli argomenti, e li dispose così come gli vennero sottomano, o come furono—alla spezzata—vergati dal Torricelli. Da ciò è risultata una illogica disposizione degli argomenti, resa più grave, nell'edizione faentina†, dalla arbitraria numerazione dei paragrafi e dalla imperfezione delle figure.

7. La memoria che ora vogliamo studiare è pubblicata nel vol. I parte II, occupa le pag. dalla 229 a 274, ed ha per titolo *De infinitis hyperbolis.*

Non è un' esposizione coordinata e connessa; ma una serie di appunti che il Torricelli buttava giù per preparare il materiale del suo lacoro. Dal confronto col manoscritto originale, conservato a Firenze nella Biblioteca nazionale‡, si riconosce che l'edizione faentina riproduce fedelmente, nel suo caratteristico disordine, il testo torricelliano.

Il riordinamento, secondo la successione logica degli argomenti, di quegli oscurissimi frammenti, così capricciosamente trasposti e frammischiati, pare a prima vista cosa pressochè inattuabile; poichè domanda una preventiva cognizione del concetto generale dell'opera e dei particolari svolgimenti ad essa dati dal Torricelli; la qual cognizione certamente non si può ricavare dalla semplice lettura di quegli appunti, così come si trovano nell' edizione faentina.

8. Ho visto per altro che, ricorrendo all' esame diretto dell' autografo, con la scorta del carteggio e degli altri scritti torricelliani, e con la paziente applicazione del metodo delle approssimazioni successive, si può ricostruire nella sua interezza

*Cf. E. Bortolotti. *La scoperta e le successive generalizzazioni di un teorema di calcolo integrale.* Archivio de Storia della scienza, Vol. V, No. 3 (1924).

†*Opere di Evangelista Torricelli*, edite in occasione del III centenario delle nascita, col concorso del comune di Faenza da Gino Loria, e Giuseppe Vassura, Vol. I, pubblicato per cura di Gino Loria, Parte II, Faenza 1919.

‡*Discepoli di Galileo*, Vol. XXXI.

il pensiero torricelliano e ricomporre, anche nei particolari, la trattazione della memoria, senza lacune essenziali; pur senza nulla aggiungere e nulla mutare: in modo da giustificare l'appassionata protesta che il Torricelli faceva nel suo letto di morte, di "aver scritte le dimostrazioni e ogni cosa di quello che ho promesso agli amici di dimostrazione".

Dirò anzi che si trova più di quello che il Torricelli aveva promesso. Poichè, non solo tutte le proposizioni da lui enunciate sono con metodo generalissimo ricondotte ad un unico lemma fondamentale semplicissimo; ma molte idee generali, che solo più tardi entrarono nel dominio della scienza, trovano qui la loro origine ed il loro principio. Onde verrà che questa memoria dovrà esser considerata come una delle fonti principali per la storia del metodo infinitesimale.

9. Mi propongo ora di far conoscere il contenuto essenziale di tale memoria, usando per maggior perspicuità il linguaggio e le notazioni moderne.

Ed esprimo l'augurio che in tempo non lontano, venga fatta un'edizione critica, che riproduca nella veste originale, ma nel naturale ordinamento logico della materia, le opere lasciate inedite dal Torricelli.

II—Esposizione Sistematica del contenuto della Memoria "De Infinitis Hyperbolis"

10. Lemma Fondamentale—

Sia la curva $x^m y^n = k$, m, n, *interi qualunque (positivi o negativi) e su di essa i punti* $P_1(x_1 y_1)$, $P_2(x_2, y_2)$, $(x_2 > x_1)$. *Si considerino le proiezioni* Q_1, Q_2; R_1, R_2 *di questi punti sugli assi di riferimento e si completino i rettangoli* P_1Q_2, P_2Q_1, P_1R_2, P_2R_1. *Nella ipotesi (sempre ammissibile) che sia* $|n| > |m|$,

1° *Il rapporto dei rettangoli inscritti è maggiore del rapporto degli esponenti, e quello dei rettangoli circoscritti minore:* precisamente (nel caso iperbolico)

$$\frac{P_2Q_1}{P_1R_2} > \frac{n}{m} > \frac{P_1Q_2}{P_2R_1}. \tag{1}$$

2° *Se facciamo rotare tutta la figura intorno all'-asse y,* (OR_1), *il rapporto dei solidi generati dalla rotazione dei rettangoli inscritti è maggiore del doppio del rapporto degli esponenti, quello dei solidi generati dai circoscritti minore.* Precisamente (stesso caso)

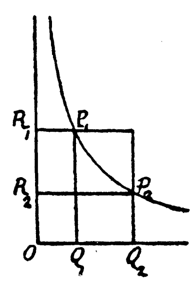


Fig. 1a

$$(2) \qquad \frac{\text{solido da } P_2Q_1}{\text{solido da } P_1R_2} > \frac{2n}{m} > \frac{\text{solido da } P_1Q_2}{\text{solido da } P_2R_1}.$$

3° *Analoghe relazioni si hanno per i solidi che si ottengono facendo rotare la figura intorno all'asse x:* (OQ_2), e cioè (sempre nello stesso caso):

$$(3) \qquad \frac{\text{solido da } P_2Q_1}{\text{solido da } P_1R_2} > \frac{2m}{n} > \frac{\text{solido da } P_1Q_2}{\text{solido da } P_2R_1}.$$

L'enunciato, sotto forma generale, pel 1° caso, si trova alla pag. 248, nelle ultime 11 linee del parag. 33 colle parole: *Esto ut dignitas AB ad BC, ita dignitas BD ad BE.... Dico parallelogrammum DF ad FA majorem habere rationem quam exponentes dignitatum.*

Se ne trova una dimostrazione completa al paragrafo 28 per la iperbola $xy^3 = k$; e, limitata al caso dei solidi, per l'iperbola $x^3y^7 = k$, al parag. 27.

Si vedano anche i parag. 1 a 7, 17, 29 a 36 ove sono trattati altri casi particolari. Il Torricelli avverte inoltre che i medesimi ragionamenti valgono anche per le parabole $y^n = kx^m$.

Se si tien conto del fatto che Torricelli sostituisce la operazione aritmetica della estrazione di radice n^{ma} con la inserzione di $n-1$ medie proporzionali, e la operazione dell'innalzamento a potenza, con la costruzione di linee continuamente proporzionali, si potrà dare ai ragionamenti che egli esprime sotto forma geometrica, la seguente interpretazione algebrica:

Suppongasi, nella $x^my^n = k$, $n > m > 0$ (caso iperbolico), onde

$$x = k^{\frac{1}{m}} y^{-\frac{n}{m}}.$$

Si avrà (Fig. 1):

$$\frac{P_2Q_1}{P_1R_2} = \frac{(x_2-x_1)y_2}{(y_1-y_2)x_1} = \frac{(y_2^{-\frac{n}{m}} - y_1^{-\frac{n}{m}})y_2}{(y_1-y_2)y_1^{-\frac{n}{m}}} = \frac{(y_1^{\frac{n}{m}} - y_2^{\frac{n}{m}})y_2}{(y_1-y_2)y_2^{\frac{n}{m}}}$$

$$= \frac{\left(\frac{y_1}{y_2}\right)^{\frac{n}{m}} - 1}{\frac{y_1}{y_2} - 1} = \frac{\left(\frac{y_1}{y_2}\right)^{\frac{n}{m}} - 1}{\left(\frac{y_1}{y_2}\right)^{\frac{1}{m}} - 1} : \frac{\frac{y_1}{y_2} - 1}{\left(\frac{y_1}{y_2}\right)^{\frac{1}{m}} - 1}.$$

Pongo $\left(\frac{y_1}{y_2}\right)^{\frac{1}{m}} = h$, onde, per le nostre ipotesi, sarà

$$h > 1,$$

ed avremo:

$$\frac{P_2Q_1}{P_1R_2} = \frac{h^n-1}{h-1} : \frac{h^m-1}{h-1} = \frac{h^{n-1}+h^{n-2}+\ldots+1}{h^{m-1}+h^{m-2}+\ldots+1}$$

$$\frac{P_2Q_1}{P_1R_2} - 1 = \frac{h^{n-1}+h^{n-2}+\ldots+h^m}{h^{m-1}+h^{m-2}+\ldots+1} > \frac{(n-m)h^m}{mh^{m-1}} > \frac{n-m}{m}.$$

Onde appunto:

$$\frac{P_2Q_1}{P_1R_2} > \frac{n}{m}.$$

Con analogo ragionamento si procede negli altri casi.

11. Quadrature—Dal lemma fondamentale il Torricelli deduce (con una elegante dimostrazione per exaustione) al parag. 45 (pag. 256-257) che *i quadrilinei che risultano dalle proiezioni dell'arco P_1P_2 sugli assi stanno fra loro nel rapporto medesimo degli esponenti;* cioè che

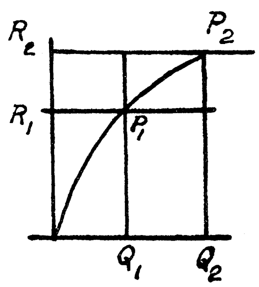


Fig. 2

(4) $$\frac{P_1Q_1Q_2P_2}{P_1R_1R_2P_2} = \frac{n}{m}.$$

Di qui, nel caso parabolico componendo, si ha

(5) $$P_1Q_1Q_2P_2 = \frac{n}{m+n}(R_2Q_2 - R_1Q_1)$$

e, nel caso iperbolico dividendo,

(6) $$P_1Q_1Q_2P_2 = \frac{n}{n-m}(R_2Q_2 - R_1Q_1).$$

Volendo interpretare con la notazione integrale la dimostrazione fatta scriveremo anzitutto la formula

(7) $$\int_{x_1}^{x_2} ydx = \frac{n}{m}\int_{y_1}^{y_2} xdy,$$

che è l'espressione del teorema (4); le operazioni del comporre o del dividere corrispondono alla integrazione per parti

(8) $$\int_{x_1}^{x_2} ydx = [xy]_{x_1}^{x_2} - \int_{y_1}^{y_2} xdy;$$

combinanto questa con la (7) si avrà:

$$\left(\frac{m}{n}+1\right)\int_{x_1}^{x_2} ydx = [xy]_{x_1}^{x_2}.$$

ossia:

$$\int_{x_1}^{x_2} x^{\frac{m}{n}}\, dx = \frac{1}{\frac{m}{n}+1}\left(x_2^{\frac{m}{n}+1} - x_1^{\frac{m}{n}+1}\right). \tag{9}$$

per m, n interi positivi o negativi qualunque. Questa formula corrisponde appunto alle (5), (6). Onde si vede che il metodo di Torricelli offre la più generale dimostrazione della *formula di Cavalieri* $\int_0^x x^\alpha dx = \frac{1}{\alpha+1}\, x^{\alpha+1}$, con α razionale qualunque.

12. Il Torricelli ha piena cognizione della generalità del suo procedimento; infatti egli fa precedere alla sua dimostrazione la nota seguente*:

NOTA—"Nota che la dimostrazione della seguente pagina è comune anco alle parabole e alle spirali, e con pochissima mutazione può adattarsi a ciascuna delle tre predette sorti di figure, e fors' anco potrebbe comprenderle tutte in una sola proposizione. Ma noi doviamo aborrire più la oscurità che la lunghezza."

A questa proposizione nell' autografo torricelliano, seguono quattro pagine bianche dove il Torricelli forse intendeva di svolgere le facili conseguenze. Da questo teorema segue infatti immediatamente la dimostrazione delle proposizioni da lui enunciate nelle lettere spedite nel luglio 1646†.

13. Tangenti—Dal lemma fondamentale segue immediatamente la *costruzione delle tangenti* alle curve considerate. Tale costruzione è indicata succintamente al parag. 46 (Tangentes omnium). Anche a questo paragrafo seguono, nell'autografo torricelliano, parecchie pagine bianche, ove forse egli intendeva di meglio sviluppare il suo ragionamento. Ciò che qui è accennato, insieme con la avvertenza che il *suo metodo procede per gnomoni*,‡ permette di sviluppare la semplicissima dimostrazione.

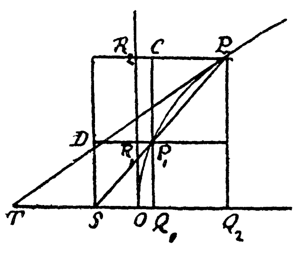


Fig. 3

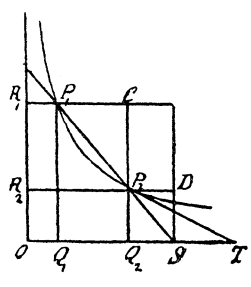


Fig. 4

Si prolunghi la corda P_1P_2 fino ad incontrare l'asse delle x nel punto S, e si completi il rettangolo P_1Q_1S (nel caso iperbolico, oppure P_2Q_2S nel caso para-

**Op. di Torricelli*, Vol. I, Parte II, pg. 256.

†Cfr. E. Bortolotti: *La scoperta . . . loc. cit.*, pg. 218 *e sqq.*

‡Cfr. lettera a Cavalieri del 21 settembre 1647.

bolico). Pel *teorema del gnomone* avremo (caso iperbolico) l'equivalenza dei rettangoli P_2Q_1 e CD onde per il lemma (formola 1)

$$\frac{CD}{P_1R_2} > \frac{n}{m} > \frac{P_1D}{P_2R_1},$$

ed, osservando che i rettangoli che qui si considerano sono posti fra le medesime parallele, si ha ancora

$$\frac{\overline{Q_2S}}{\overline{OQ_1}} > \frac{n}{m} > \frac{\overline{Q_1S}}{\overline{OQ_2}} \tag{10}$$

Se ora si considera che quando P_1 muovendosi sulla curva viene a coincidere con P_2 anche i punti Q_1Q_2 vengono a coincidere, indicando con T la intersezione della tangente coll' asse x, avremo:

Se la retta P_2T è tangente alla curva in P_2, si ha.

$$\frac{\overline{Q_2T}}{\overline{OQ_2}} = \frac{n}{m}. \tag{11}$$

Pel caso parabolico si ha similmente

$$\frac{\overline{TQ_2}}{\overline{OQ_2}} = \frac{n}{m}. \tag{11'}$$

14. Le Linee Supplementari—Del teorema ora dimostrato il Torricelli dà una notevole interpretazione nella sua memoria *De infinitis parabolis*, al paragrafo intitolato: *Delle tangenti alle parabole infinite per lineas supplementares*' (pag. 320 a 323.)

Dopo aver dimostrato che le aree dei quadrilinei inscritti $P_1Q_1Q_2P_2, P_1R_1R_2P_2$ stanno fra loro come gli esponenti cioè come n: m, il Torricelli osserva che lo stesso interviene per i quadrilinei che si ottengono prendendo invece dell'ascissa Q_1, il punto di mezzo del segmento Q_1Q_2: "*et se faremo, o supporremo fatta questa divisione in infinite volte, resteranno in cambio di figure, due linee, le quali, non secondo la longitudine ma secondo la quanità, stanno nel medesimo rapporto come l'esponente all'esponente.*"

Se *gli spessori di queste linee aventi quantità* (indivisibili) *si indicano con* dx, dy (come più tardi fece il Leibniz) e le lunghezze loro con y, x, si avranno dette quantità (cioè gli elementi d'area) espresse da ydx, xdy, ed il trorema potrà enunciarsi così:

Gli elementi d'area ydx, xdy, sono fra loro come $n : m$.

Il Torricelli chiama *linee supplementari* due segmenti le cui lunghezze stanno fra loro nel rapporto degli elementi d'area, cioè come $n : m$. Per il teorema dimostrato al numero precedente, *sono dunque linee supplementari la sottotangente TQ_2 e l'ascissa OQ_2*. In altri termini si ha

$$\frac{\overline{TQ_2}}{\overline{OQ_2}} = \frac{\overline{Q_2P_2}dx}{\overline{OQ_2}dy}. \tag{12}$$

Di qui immediatamente si ricava (ma ciò non è detto nella memoria di Torricelli).

$$\frac{dy}{dx}=\frac{Q_2P_2}{TQ_2}=\mathrm{tg}\,\overline{P_2TQ_2}, \tag{13}$$

che *esprime il coefficiente angolare della tangente, mediante il rapporto dei differenziali delle variabili* (degli spessori degli indivisibili).

15. La costruzione delle tangenti ed un teorema sui massimi e minimi — Il Torricelli ha in più luoghi osservato che *il suo metodo dava ad un tempo le quadrature, le tangenti ed i massimi.* La ragione di ciò è fugacemente accennata al parag. 46 della memoria che stiamo esaminando, ma si trova più chiaramente sviluppata alle pag. 307, 308, nella memoria "*De infinitis parabolis*" ove si tratta appunto *Delle tangenti dell' infinite hyperbole senza quadratura ex sola definitione**.

16. Il Torricelli vuol provare che se $\dfrac{\overline{OQ_2}}{\overline{Q_2T}}=\dfrac{m}{n}$, ($m$, n positivi) *il punto* Q_2 *divide il segmento* $\overline{OT}$ *in modo che il prodotto* $\overline{OQ_2}^m.\ \overline{Q_2T}^n$ *è massimo.*

Si prenda un punto Q_1, diverso da Q_2 (per fissare le idee facciamo $\overline{OQ_1}<\overline{OQ_2}$, cioè $\overline{Q_1Q_2}>0$), poi si conducano le perpendicolari Q_1P_1, Q_2P_2, e su queste si prendano i punti P_1, P_2, tali che

$$\overline{OQ_1}^m.\ \overline{Q_1P_1}^n=\overline{OQ_2}^m.\ \overline{Q_2P_2}^n=k, \tag{14}$$

e si conducano le rette P_1P_2S, TP_2M.

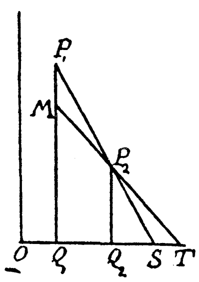


Fig. 5

Essendo $\qquad \overline{Q_2T}:\overline{Q_1T}=\overline{Q_2P_2}:\overline{Q_1M},$

avremo

$$\frac{\overline{OQ_2}^m.\ \overline{Q_2T}^n}{\overline{OQ_1}^m.\ \overline{Q_1T}^n}=\frac{\overline{OQ_2}^m.\ \overline{Q_2P_2}^n}{\overline{OQ_1}^m.\ \overline{Q_1M}^n}=\frac{\overline{OQ_1}^m.\ \overline{Q_1P_1}^n}{\overline{OQ_1}^m.\ \overline{Q_1M}^n}=\left(\frac{\overline{Q_1P_1}}{\overline{Q_1M}}\right)^n.$$

Si vede di qui che,

$$\text{se } \overline{Q_1P_1}>\overline{Q_1M},\ \text{sarà anche } \overline{OQ_2}^m.\ \overline{Q_2T}^n>\overline{OQ_1}^m.\ \overline{Q_1T}^n. \tag{15}$$

e reciprocamente, onde:

*Questo pezzo è evidentemente fuor di posto. Vedi anche il paragrafo 16 alla, pag. 239.

Se il punto T sarà determinato in modo che il prodotto $\overline{OQ_2}^m \, \overline{Q_2T}^n$ sia massimo (o minimo) tutti i punti M della retta TP_2 (tranne il punto P_2 stesso) saranno esierni alla curva $x^m y^n = k$, cui appartengono i punti P_2, cioè la retta TP_2 sarà tangente in P_2 alla curva, e reciprocamente.

17. Ma si può provare direttamente che, nelle condizioni poste, sussiste il Teor. 16, cioè che si ha veramente $\overline{Q_1P_1} > \overline{Q_1M}$ (d'onde anche una nuova dimostrazione per la *costruzione delle tangenti*).

Infatti, pel Lemma (n⁰. 13, form (10)), si ha

$$\frac{\overline{Q_2S}}{\overline{OQ_1}} > \frac{n}{m} > \frac{\overline{Q_1S}}{\overline{OQ_2}}$$

e, per essere

$$\frac{\overline{Q_2T}}{\overline{OQ_2}} = \frac{n}{m},$$

$$\frac{\overline{Q_2S}}{\overline{OQ_1}} > \frac{\overline{Q_2T}}{\overline{OQ_2}} > \frac{\overline{Q_1S}}{\overline{OQ_2}},$$

da ció

$$\overline{Q_2T} > \overline{Q_1S};$$

ed, essendo

$$\overline{Q_2T} = \overline{Q_2Q_1} + \overline{Q_1S} + \overline{ST},$$

$$\overline{Q_2Q_1} + \overline{ST} > 0,$$

cioè

$$\overline{ST} > \overline{Q_1Q_2} > 0.$$

Ma

$$\frac{\overline{Q_1P_1}}{\overline{Q_2P_2}} = \frac{\overline{Q_1S}}{\overline{Q_2S}}; \quad \frac{\overline{Q_1M}}{\overline{Q_2P_2}} = \frac{\overline{Q_1T}}{\overline{Q_2T}} = \frac{\overline{Q_1S}+\overline{ST}}{\overline{Q_2S}+\overline{ST}},$$

ed avendosi $\overline{Q_1S} > \overline{Q_2S}$*, ne verrà:

$$\frac{\overline{Q_1T}}{\overline{Q_2T}} < \frac{\overline{Q_1S}}{\overline{Q_2S}},$$

cioè:

$$\frac{\overline{Q_1M}}{\overline{Q_2P_2}} < \frac{\overline{Q_1P_1}}{\overline{Q_2P_2}},$$

e infine

$$\overline{Q_1M} < \overline{Q_1P_1}.$$

Da questa dimostrazione si vede che *la medesima costruzione serve ad un tempo per tracciare la tangente in un punto dato P* della curva $x^m y^n = k$, e, quando sia dato il segmento $\overline{OT}$, per trovare il punto Q che divide tale segmento in due parti $\overline{OQ}$, $\overline{QT}$ tali che rendono massimo il prodotto $\overline{OQ}^m . \overline{QT}^n$.

*Il caso $\overline{Q_1S} < \overline{Q_2S}$ si ha quando n ed m hanno segni contrari (caso parabolico); ed in questo caso risulta $\overline{Q_1M} > \overline{Q_1P_1}$ e si ha *minimo*.

18. Criterio di convergenza per integrali estesi ad intervalli infiniti—

"*Necesse est exponentem asymptotalium majorem esse exponente applicatarum; alias figura, non solum longitudine, sed etiam magnitudine infinita esset.*" (Parag. 23, pag. 242).

Cioè se $m \leqq n$, l'area compresa fra la curva $x^m y^n = k$ e l'asse x, è infinita. La dimostrazione è data ai paragrafi 54, 55 pag. 263, 264 ed è la seguente: Si prendano sull'asse x i punti Q_1, Q_2, Q_3, ..., tali che

$$\overline{OQ_2} = 2\overline{OQ_1},\ \overline{OQ_3} = 2\overline{OQ_2},\ \overline{OQ_4} = 2\overline{OQ_3}, \ldots,$$

Si conducano le ordinate corrispondenti della curva: Q_1P_1, Q_2P_2, Q_3P_3, ..., e le rette R_1P_1, R_2P_2, R_3P_3, ..., parallele all'asse x; si chiami con T_1 il punto d'incontro delle P_1Q_1, P_2R_2, con T_2 il punto di incontro delle P_2Q_2, P_3R_3, ...

Si considerino *i* rettangoli iscritti

$$T_1Q_2,\ T_2Q_3,\ T_3Q_4,\ \ldots .$$

Pel lemma fondamentale avremo

$$\frac{T_1Q_2}{R_1T_1} > \frac{n}{m}, \quad \text{cioè} \quad \frac{T_1Q_2}{R_1T_1 + T_1Q_2} > \frac{n}{n+m},$$

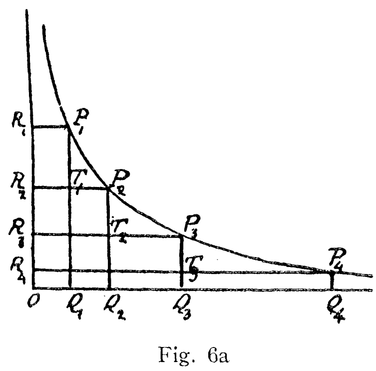


Fig. 6a

e, per essere $\overline{OQ_1} = \overline{Q_1Q_2}$, potremo scrivere

$$\frac{T_1Q_2}{R_1T_1 + R_2Q_1} > \frac{n}{n+m}, \quad \text{ossia} \quad \frac{T_1Q_2}{R_1Q_1} > \frac{n}{n+m}.$$

Similmente,

$$\frac{T_2Q_3}{R_2Q_2} > \frac{n}{n+m},$$

ma $R_2Q_2 = 2T_1Q_2$, onde:

(16)
$$\left\{\begin{array}{l} \dfrac{T_2Q_3}{T_1Q_2} > \dfrac{2n}{n+m}, \quad \text{e similmente:} \\ \dfrac{T_3Q_4}{T_2Q_3} > \dfrac{2n}{n+m}, \\ \dfrac{T_4Q_5}{T_3Q_4} > \dfrac{2n}{n+m}, \dots \\ \dots\dots\dots\dots \end{array}\right.$$

Da cio si vede che, se $2n \geqq n+m$, cioè se $n \geqq m$, si ha:

$$T_1Q_2 < T_2Q_3 < T_3Q_4 < \dots .$$

Dunque *nell'area compresa fra l'asintoto, la curva e l'ordinata* Q_1P_1, *si possono inscrivere infiniti rettangoli, tali che ciascuno di essi è maggiore del precedente.* L'area stessa è perciò infinita.

19. Questa dimostrazione *per il caso di* $m=n=1$, *cioè per la iperbola apolloniana, coincide con quella* (nota col nome di Bernouilli) *comunemente usata per provare la divergenza della serie armonica.*

20. Se invece si suppone n minore di m, considerando rettangoli circoscritti, ed applicando il medesimo lemma, si dimostrerà che l'area è finita (minore della somma degli infiniti termini di una progressione geometrica decrescente*.

21. Calcolo di integrali estesi ad intervalli infiniti—

"*Se sarà una delle infinite iperbole le cui asymptoti siano ox, oy e sia ox l'asymptoto delle dignità maggiori, se la figura sarà segata dall'applicata* P_2Q_2, *e si faccia il parallelogramma* $OR_2P_2Q_2$:

1⁰ *la figura* $Q_2P_2PR_2O$ *infinitamente lunga, sta alla figura* P_2PR_2 *infinitamente lunga, come l'esponente maggiore al minore, cioè come m ad n.*

2⁰ *il parallelogramma* $OR_2P_2Q_2$ *sta alla figura infinitamente lunga* P_2PR_2, *come la differenza degli esponenti al minor esponente*†.

La prima parte discende immediatamente con un facile passaggio al limite, dalla considerazione che, comunque sia presa l'ascissa OQ_1 il rapporto dei quadrilinei $Q_1P_1P_2Q_2$, $R_1P_1P_2R_2$, è sempre eguale al rapporto $n : m$; e dal fatto che, nelle nostre ipotesi, l'area totale $Q_2P_2PR_2O$ è limitata.

La seconda parte si ricava dalla prima "convertendo".

La proposizione così dimostrata è espressa dalla formola

(17)
$$\int_{y_2}^{\infty} y^{-\frac{n}{m}} dy = \frac{m}{n-m} x_2 y_2 = -\frac{1}{-\frac{n}{m}+1} y_2^{-\frac{n}{m}+1}$$

*Cf. lettera a M. Ricci del 17 marzo, 1646. Op., Vol. III, 361.

†Cfr. lettera al Cavalieri del 5 maggio, 1646: *Racconto di alcuni problemi*, parag. XLI (III, p. 25).

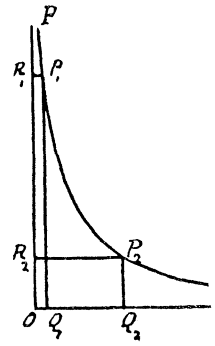


Fig. 7

22. Criteri per conoscere se *i* solidi (acutissimo o latissimo) generati dalla rotazione dell'iperbola intorno ad un asintoto sono infiniti.

Questi criteri, enunciati ai n^i3, 4, della lettera di Torricelli a M. Ricci del 29 Giugno 1647, sono dimostrati colla applicazione del medesimo procedimento seguito per le aree piane, cioè: *ostendendo quod in nostro solido compraehenditur series quaedam infinitorum multitudine cylindrorum, quorum sequens semper praecedente major est.**

Ciò ai parag. 52, 56, 65.

23. Calcolo di integrali impropri mediante integrazione per parti e per sostituzione.

Il calcolo del volume des solido acutissimo iperbolico, che si può dedurre dai lemmi precedenti, è fatto dal Torricelli anche *con una transformazione che riduce il calcolo di un integrale esteso a campo infinito, al calcolo di un integrale esteso a campo finito*, con procedimenti eleganti e rigorosi, fondati sulla applicazione del *metodo degli indivisibili, per indivisibili curvi*".

Questa materia è sviluppata per esteso ai parag. 37 a 40 e può riassumersi cosi: data la curva $x^m y^n = k$ si può calcolare il volume del solido generato dalla rotazione di essa intorno all'asse delle y, cioè calcolare l'integrale $\pi \int_{y_0}^{\infty} x^2(y)\, dy$, partendo dalla considerazione che l'equazione della data iperbola può scriversi $x^n y^n = k x^{n-m}$, ossia

$$xy = k^{\frac{1}{n}} x^{\frac{n-m}{n}}$$

ed osservando che per $m < 2n$, *il solido iperbolico acutissimo è equivalente al solido generato dalla rotazione intorno all'asse di una parabola (associata alla data iperbola) la cui equazione è data dalla relazione*

$$Y^2 = 2xy = 2k^{\frac{1}{n}} x^{\frac{n-m}{n}}. \tag{18}$$

**Op di Torricelli*, I, Parte II, p. 262.

Ciò equivale, nel linguaggio moderno, a trasformare l'integrale proposto mediante *la formola di integrazione per parti*

$$2\pi \int_0^{x_0} xy dx = \pi x_0^2 y + \pi \int_{y_0}^{\infty} x^2(y) dy \tag{19}$$

con la quale esso si riduce al calcolo dell' integrale esteso a campo finito $\int_0^{x_0} x^{\frac{n-m}{m}} dx$.

24. Il calcolo del *volume del solido iperbolico latissimo* cioè dell' integrale

$$\pi \int_0^{y_0} x^2(y) dy,$$

dove $x(y)$ è infinita per $y=0$, con trasformazione analoga a quella indicata precedentemente è ridotto a quello del *conoide iperbolico acutissimo* generato dalla rotazione della iperbola associata $y^{2n}x^{m-n}=2^n k$.

Questo poi è trattato come al caso precedente.

Tutto ciò è sviluppato per esteso nei paragrafi 41 a 44. Il medesimo argomento è ripreso, con metodo poco diverso e son più ampi sviluppi, negli ultimi paragrafi della memoria, cioè nei paragrafi 62 a 67.

25. Integrali indefiniti—(parag. 48 a 51, pag. 259-261 e pag. 309-310).

Il Torricelli osserva anzitutto che, se nello stesso tempo due mobili percorrono gli spazi s, S, gli spazietti elementari (indivisibili) percorsi in uno stesso tempo elementare dt, sono fra loro come le velocità v, V, che in quell'istante hanno i due mobili.

Gli spazi totali s,S, sono dunque fra loro come gli integrali delle velocità

$$s : S = \int_0^t vdt : \int_0^t Vdt. \tag{20}$$

Negli esempi da lui recati egli suppone V costante, cioè lo spazio S percorso con moto equabile: e precisamente suppone che sia S lo spazio che per legge d'inerzia percorre il punto mobile in un tempo t_1 (susseguente a t_1 ed eguale a t_1) nella ipotesi che all'istante $t=t_1$ cessi l'azione della forza che prima agiva sul punto. Lo spazio S è dunque dato dal rettangolo t, $v(t)$.

Il Torricelli considera ad un tempo le due curve; quelle che nel dato moto percorre il punto mobile, e quella che rappresenta la velocità di tale punto, come funzione del tempo. E dimostra che, *se la curva che rapprente la variazione delle velocità, è une iperbola quadratica* (cioè $v^2t=k^2$), *gli spazi sono in ragione subdupla dei tempi* ($s^2=4k^2t$).

E, che *se la curva delle velocità è una iperbola semi-cubica* ($v^3t^2=k^3$), *i cubi degli spazi sono fra loro come i tempi* ($s^3=ct$).

Egli calcola cioè gli *integrali indefiniti*

$$\int \frac{dt}{\sqrt{t}}, \quad \int \frac{dt}{\sqrt[3]{t^2}}.$$

Il procedimento è generale, poichè si fonda sulla applicazione dei teoremi da lui dimostrati per le iperbole generali. Ne viene così la determinazione *dell'integrale indefinito:*

$$\int \frac{dt}{t^{\frac{m}{n}}}.$$

Questo medesimo procedimento è da lui applicato per le parabole generali $v^n = kt^m$ nella memoria "*De Infinitis Parabolis*" (pag. 309-311).

26. Il teorema di inversione—

Uno dei principali meriti che gli storici riconoscono al Torricelli è quello di *aver fatto conoscere il carattere inverso che presentano le operazioni di integrazione e di derivazione* (la quadratura e la costruzione delle tangenti* ma ordinariamente si ammetteva che il solo caso considerato dal Torricelli fosse quello della parabola apolloniana†.

Vedremo invece che egli ha considerato tutte le curve rappresentate da equazioni della forma $y^n = kx^m$, cioè tutte le *parabole generali,* con procedimento che si può applicare anche alle iperbole generali.

Ed infatti egli *desume il teorema d'inversione dalla considerazione delle linee supplementari*‡ osservando che *il rapporto del segmento parabolico ABDV al suo trilineo BCVD è eguale al rapporto della sottotangente all'ascissa* $\overline{AT} : \overline{AV}$ cioè che

$$\int_0^x ydx : \left(xy - \int_0^x ydx\right) = ydx : xdy = n : m$$

onde conclude che: *dalla conoscenza della quadratura si ricava la regola per la costruzione delle tangenti, e reciprocamente.*

"Post quadraturam certe concludit, et dicit . . . tangens fiat ut parabola as trilineum, ita *TA* ad *VA*§."

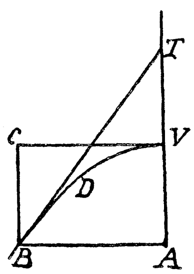


Fig. 8

*Cfr. per es. Zeuthen, H. G.: *Notes sur l'histoire des Mathématiques* (Bull. Acad. de Danemark a. 1897, pp. 572-576).

†Nella sua Mem. *De moto gravium naturaliter descendentium* prop. XVIII (Op., Vol. II pp. 122 e segg.).

‡Nella mem. *De Infinitis Parabolis*, p. 313.

§*Op.* I, parte II, p. 313.

Infatti, se si sappia quadrare la curva, cioè se si conosca il rapporto

$$\frac{\int_0^x ydx}{xy} = \frac{n}{n+m},$$

che l'area della curva ha a quella del rettangolo circoscritto, basterà in questa proporzione *dividere*, per avere il rapporto

$$\frac{\int_0^x ydx}{xy - \int_0^x ydx} = \frac{n}{m} = \frac{\overline{AT}}{\overline{AV}},$$

della sottotangente alla ascissa; e con ciò la costruzione della tangente alla curva data.

Reciprocamente, se sarà nota la costruzione della tangente, cioè il rapporto della sottotangente alla ascissa

$$\frac{\overline{AT}}{\overline{AV}} = \frac{n}{m} = \frac{\int_0^x ydx}{xy - \int_0^x ydx},$$

componendo, si troverà il rapporto dell' area della curva a quella del rettangolo circoscritto, e perciò si saprà quadrare la curva.

27. Ma il "*principio di inversione*" risulta nel modo più generale e perspicuo dall' applicazione della *dottrina dei moti*, fatta dal Torricelli per le infinite parabole nella memoria "De infinitis parabolis". E, più precisamente, dal contenuto delle pag. 309 a 316 Vol. I, parte II, delle "*Opere*".

L'argomento è estremamente attraente, ma lo non voglio più oltre indugiarmi, riservando ad altra mia comunicazione lo studio della Memoria "*De infinitis parabolis*".

28. E, se queste mie fatiche interesseranno i cultori della storia della matematica, continuerò poi, cercando di mettere in luce con successive comunicazioni le ricerche di Torricelli:

Sulle spirali, le cicloidi, la logaritmica,
sulla determinazione dei centri di gravità,
sui paradossi degli indivisibili.

E spero che anche altri, di me più degno, vorrà por mano a quest' opera di reintegrazione: in modo che finalmente si giunga a conseguire la piena e sicura conoscenza dell' opera geometrica di Torricelli, cioè di una delle più importanti fonti per la storia del calcolo infinitesimale.

HISTORY OF SEVERAL FUNDAMENTAL MATHEMATICAL CONCEPTS

By Professor G. A. Miller,
University of Illinois, Urbana, Illinois, U.S.A.

One of the most fundamental practices in mathematics is the utilization of a symbol for an unknown number, both as an operand and also as an operator. In the work of Ahmes there appear various examples in which an unknown is used as an operand, giving rise to equations of the first degree in which we find practically a special symbol for the unknown with the suggestive meaning of heap. When the unknown is squared, or two unknowns are multiplied together, the unknown is evidently used as an operator as well as an operand. In these forms it is found in Egyptian papyri which may be as old as the work of Ahmes itself.

The significant fact is that the mathematician often trails an unknown until it entangles itself by assuming a position in an equation which can be held by only one or by at most n distinct numbers, where n is usually a small positive integer. In the former case he says that the unknown satisfies a linear equation, while in the latter case he says that it satisfies an equation of degree n. This method of watching the performances of an unknown until it either discloses its identity by its environments, or assumes a position which only a small number of numbers can possibly occupy and thus throws light on itself is a mathematical trick to which students of our subject have resorted practically from the time when the earliest of the extant literature relating thereto was developed.

The disentangling of the unknown when it presents itself in the form of an algebraic equation has been an objective of mathematical endeavours extending through several thousand years, and the fruits of these endeavours are among the most marvelous which adorn the history of our subject. Among these fruits are the extension of the number concept so as to include the ordinary complex numbers of our days and the later far-reaching theory known by the name of Galois. In fact, the vast subject of algebra is due mainly to the efforts to disentangle the unknowns of algebraic equations.

It might have been thought that the more extensively the unknown becomes involved in the form of an equation the more fully would it disclose its identity but just the reverse is actually the case since $n-1$ other numbers can satisfy the same condition when it presents itself in the form of an algebraic equation of degree n. The fact that there is really no algebraic equation of degree n in one unknown when n exceeds unity but that $n-1$ others unite themselves with this one in a somewhat mysterious manner, and thus turn the study of the algebraic equation into a problem in sociology is one of the most interesting phenomena of mathematics, and it naturally has served to excite the curiosity of many mathematicians. This is a very important fact in the history of our subject since mathematical developments have been very largely inspired by curiosity.

We have little evidence that the Greeks recognized this social problem in their study of the algebraic equation. In the case of the special numerical quadratic equations which they solved they gave only one root even when the two roots are positive and unequal. On the contrary, Archimedes did consider the condition under which a certain problem which is equivalent to a cubic has two positive solutions, but this seems to be an exceptional case judging from the Greek literature which has been preserved in a more or less reliable form. We direct attention to the adjective reliable because we possess comparatively little of Greek mathematical literature in the original form. Even Euclid's *Elements*, which constitute such a corner stone in the history of Greek mathematics, are now known only from transcriptions.

The fundamental position occupied by the unknown in the history of mathematics is illustrated not merely by its influence on the development of algebra and the powerful number concept as exhibited in the modern theory of algebraic numbers, but the fruitful concepts of function and group are also closely related thereto. It is difficult to determine just when these concepts first appeared in mathematics. In an implicit form they certainly are present in the work of Ahmes and more emphatically in Greek mathematics. In fact, the group concept, when the term group is used with its wider meaning, appeared already when it was noted that the sum of two positive integers is always a positive integer and that the product of two positive integers is always a positive integer. In this form it is probably older than the concept of an unknown, but when the latter appeared and the symbol denoting it was subjected to operational laws, the group concept presented itself more explicitly, especially when a theory of the disentanglement of this unknown was undertaken after it had embedded itself in the form of an algebraic equation of degree n, $n>1$.

A very important epoch in the history of a mathematical concept is reached when it receives a special name. This name may not be a permanent one. As noted above the first name for the unknown which has as yet been found was the temporary name *heap*. Much later Diophantus denoted it by *number* and about two and a half centuries after this Aryabhata used the term *small sphere* for the same purpose. While all of these terms are suggestive they fail to bring out explicitly an acknowlegdment of actual ignorance as is done by our modern name unknown. In fact, we use a special symbol to represent a certain kind of ignorance which we now commonly represent by the letter x. It is interesting to note such evidence of increasing openness and honesty among the mathematicians of the world. Mathematics has always tended to make the thinking world more honest as well as more accurate.

Although the concept of group with its most general meaning may be older than the concept of unknown, we have no evidence of a special name for the former concept until long after it had been used with its more fruitful restricted meaning, and thus had received a special name near the close of the eighteenth century when Ruffini denoted it by *permutation*. Somewhat earlier Lagrange denoted the theory of substitutions by *calculus of combinations*. A. L. Cauchy used the term *system of conjugate substitutions* to represent a group of substitutions and this term was also used by J. A. Serret in his widely read algebra, but the modern term, due to Galois, made its appearance only about thirty years after this concept had first received a special name, and the general adoption of this term is largely due to C. Jordan.

The concept of function, which is also closely related to that of unknown, received a special name somewhat earlier than that of group, but it was also used implicitly long before such a name aided the mathematicians to handle it more conveniently and to enrich their grasp of the far-reaching rôle which this concept occupies in the development of mathematics. John Bernoulli seems to have been the first to use the term function with its modern meaning in 1698 but Leibniz had used it with a somewhat different mathematical meaning about six years earlier*. A very important stage in the history of the unknown was reached when the symbol representing it was allowed to represent a variable as well as a constant. In general, these two notions are not assigned to this symbol at the same time, but the reader is allowed the privilege of passing from one of them to the other without any change in the formal language. Traces of this change are as old as traces of the concept of function, and they appear in the Greek equations for certain curves as well as in the work of Oresme, but one fails to find much along this line before the founding of analytic geometry through the labours of Descartes, Fermat, and others.

The equation is an important trap for catching some mathematical unknowns. This trap cannot properly be said to have been invented by the mathematician for the sake of catching certain unknowns but in his study of his surroundings the mathematician often found an unknown thus entrapped and he naturally was interested in its disentanglement. As evidence of this fact we may direct attention to the *golden section* of the ancient Greeks and to their problems of duplicating the cube and of trisecting an angle, as well as to a much older problem found in Ahmes relating to the average daily income when the annual income is known. After becoming acquainted with this intellectual trap the mathematician naturally set it himself with a view to catching particular unknowns in which he became interested. After they were thus caught the disentanglement became a serious problem which often required greater skill than the process of entrapping.

The Indians made decided progress in the study of the equation as a tool by means of which we find more than we were explicitly seeking. In particular, they observed that certain quadratic equations have two roots, but even they, as well as the Arabs, failed to make the very important contribution to mathematics known as the complete solution of the quadratic equation. Such a solution had to await the development of the number concept so as to include the ordinary complex numbers. We may, however, notice among the Indians

*Cf. H. Wieleitner, *Bibliotheca Mathematica*, 13 (1913), p. 145.

a tendency to study the equation as a powerful intellectual trap. This tendency has become more and more pronounced in the later development of our subject. In a mild form this tendency appears already in the work of Ahmes, as may be seen from such problems as "heap its 2/3, its 1/2, its 1/7, its whole, it makes 33". The early mathematical life was largely supported by intellectual trapping just as the early physical life of the human race was largely dependent on physical trapping in securing food and clothing.

The present writer has protested elsewhere* against the common statement that the Greeks *solved* the quadratic equation and that the algebraic solution of the cubic and the biquadratic was contributed by the Italian mathematicians during the first half of the sixteenth century. It seems to him that such incomplete statements relating to questions of fundamental importance should no longer be tolerated in general histories of our subject since they tend to obscure the actual situation. On the present occasion he feels compelled to protest equally strongly against a common statement which has recently been expressed in the following words, "The geometric solution of the quadratic equations is accomplished by Euclid's construction," page 38 of Volume 3, 1922, Tropfke's *Geschichte der Elementar-Mathematik.* On page 69 of the same volume it is stated that the geometric solution of the cubic was found by the Arabs.

By *solution* the modern student naturally understands much more than what was accomplished by these ancient methods. Complete geometric solutions of these equations are of great interest, but the Greeks and the Arabs had not reached the stage when complex roots could be considered and hence they could not have given complete geometric solutions of these equations. The cultivation of precision in general historical statements is of fundamental importance to the beginner in this field. The expression solution of an equation naturally implies more and more as our knowledge as regards the equation increases.

The concept of unknown is a prehistoric mathematical concept. It appears with its modern meaning and with an appropriate name in the work of Ahmes. On the other hand, the closely related concepts of function and group did not reach the same stage of development until after the centre of mathematical progress had been established in western Europe. The former of these two concepts assumed a fundamental rôle during the founding period of analytic geometry and the calculus, while the latter did not play such a rôle before the second half of the nineteenth century. An interesting negative property of a group element is the fact that it has no prime factors.

The close contact between group and unknown is exhibited not only by the fact that the former was first systematically studied in connection with the behaviour of the n unknowns in a so-called equation in one unknown, but also by the equational definition of a group; *viz.*, that if any two of the symbols in $xy=z$ represent known elements of a group the third or unknown element is uniquely determined thereby and that the symbols obey the associative law.

*Cf. School Science and Mathematics, Vol. 24 (1924), p. 509.

This useful definition seems to have first appeared explicitly in Weber's *Kleines Lehrbuch der Algebra*, 1912, page 181.

One of the most attractive features of the history of mathematics is that many of the developments centre in dominant concepts. In fact, the real life of this history for many of us is involved in these dominant concepts, which have not always exercised the same relative influence. The concept of the natural numbers is doubtless older and more fundamental than that of the unknown. In fact, the former is often regarded as the source of all mathematics, but the time when it was the only fundamental mathematical concept which was explicitly recognized had ended before our earliest extant mathematical literature had its origin. Not only does this early literature present to us an extension of the number concept so as to include rational fractions and the unknown, but also elementary geometric concepts had already made their appearance in an explicit form.

When one has named the concept of natural number, the concept of unknown as it appears in the equation, and the concepts of function and group, one has a category of four dominating concepts in the history of our subject. Each of these, except the first, evolved out of its predecessors and served to throw additional light thereon. Notwithstanding the great importance of these concepts it seems possible to exaggerate this importance, and it would seem to be one of the duties of the historian to seek the proper modifications. For instance, it is often stated that the group of an equation reflects *all* the properties of the equation*. One need only recall the many properties of the quadratic equation and the relatively few properties of its group to see that this must be an exaggeration.

The first three of these fundamental concepts deal mainly with quantity and hence they relate principally to matter. On the other hand, the last relates to pure form. This fact was clearly expressed by H. Poincaré shortly before his death when he said "the theory of groups is, so to say, entire mathematics divested of its matter and reduced to pure form.†" This does not imply that it is desirable, in general, to divest mathematics of its matter, nor does it say anything about the relative importance of matter and of pure form in mathematical work. It is, however, a historical statement of the greatest significance when it is coupled with the fact that group theory received comparatively little attention from mathematicians until the second half of the nineteenth century.

One of the characteristic features of a group is the fact that it centres attention on totalities and hence it tends to large views. For instance, if one multiplies two rational functions in a given number of unknowns the product is a rational function in these unknowns. Some might be inclined to call this a new function in these unknowns but from the group theory point of view all such unequal functions, excluding those which are identically 0, constitute a group with respect to the operation of multiplication, and hence our product is not a new rational function but simply an element of our group which had

*Cf. Pascal, *Repertorium der höheren Mathematik*, Vol. 1 (1910), p. 169.

†H. Poincaré, Acta Mathematica, Vol. 38 (1921), p. 145.

been clearly in mind as one of the elements of the totality under consideration. This example may also serve to illustrate what is meant by pure form in the quotation noted in the preceding paragraph.

Since two of the four dominating mathematical concepts noted above received special names before the time of the Greeks while the other two did not receive such names until long after this time, some mathematical historian will naturally ask, what rôle did the Greeks play as regards these concepts? This question is the more natural since mathematics is sometimes called a Greek science. A partial answer to this question is furnished by the statement made above that the concepts of function and group appear implicitly in the Greek mathematical literature and even earlier. The ancient Greeks did not extend explicitly the number concept. They received the pairs of positive rational numbers, known as fractions, from their predecessors and passed the number concept practically thus limited to their followers. It is true, however, that some of the later Greeks did calculate already with irrational numbers represented by square roots and that the Greek rules for finding a root of a quadratic equation prepared the way for a further extension of the number concept, since these rules are equivalent to the modern formula which often represents complex numbers.

Greek mathematics was dominated by a spirit of timidity as is exhibited by their constant appeal to geometric interpretations. This dominance was perhaps natural in view of the fact that such a small part of the mathematical sea had as yet been explored. It does, however, not create an atmosphere in which new fundamental concepts are apt to receive explicit recognition. The creation of a need of such recognition is perhaps most important. At least, it is an essential step to insure the development and permanence of such concepts. This need the Greeks created and herein lies one of their greatest mathematical contributions. In particular, the statement that a set of numbers is a point would probably have had little significance if a more concrete conception of point had not paved the way for the usefulness of this view.

The explicit introduction of a system of postulates for geometry and the marvelous geometric construction erected thereon by the ancient Greeks constitute their best known mathematical contributions. This does, however, not involve any mathematical concepts which are as important as the four noted above. The concept of postulate system is rich in mathematical fruition and is still unusually attractive to many mathematicians. It has served to clarify and to inspire confidence in results already obtained, as well as to point to new methods of attack. In bringing it out explicitly in advance of the concepts of function and group the Greeks failed to travel the best road to success. They found, however, a good road and deserve great credit for having pursued it so far. They utilized explicitly the concepts of number and unknown to great advantage. The introduction of a system of postulates was an important step in divorcing mathematics and philosophy. This divorce is clearly exhibited in Euclid's *Elements*, but some of the later Greeks tried already to effect a reconciliation.

One of the boldest postulates of the ancient Greeks is the so-called postulate of Archimedes, which had been used before the time of Archimedes, and made it possible for the Greek mathematicians to overlook the actually infinitely small quantities and to say, in particular, that the diameters of small circles make the same angle with the circles as the diameters of larger circles although the contrary appears to be obviously true. In the form of the angle of incidence the concept of the actually infinitely small gave rise to a large amount of controversial literature and it had a decided influence on the development of the differential calculus. These facts enhance the historical interest in the Greek conception of a tangent line to a curve and in the fact that the extant Greek literature does not exhibit this general conception as explicitly as might appear desirable.

The statement made by M. Chasles on page 57 of the second edition of his well-known *Aperçu historique des méthodes en géométrie*, 1875, that "the ancient geometers defined tangent to a curve as a straight line which having a point in common with the curve is such that one cannot draw through this point any other straight line between this one and the curve" seems to be based partly on indirect evidence. In the special case of the circle, which is the only curve considered in Euclid's *Elements*, it is possible to define a tangent line as a straight line which has one and only one point in common with the curve, and the Greeks seem to have defined the tangent line to a circle by this property, but such a definition does not apply to the tangent lines of some of the other curves considered by them. In particular, according to the Greek view a diameter of a parabola has only one point in common with this curve, but they did not regard this line as a tangent to the parabola.

Each of the five fundamental mathematical concepts; number, unknown, system of postulates, function, and group; has a very extensive history. In fact, the history of the development of these concepts entails the history of the development of the entire subject of mathematics. Among these five concepts that of system of postulates occupies the middle position in the historical order of development and it is evidently quite distinct from the others. As was noted above it represents the Greek contribution to this category of explicit concepts. While we ventured to state above that it is, from the standpoint of mathematics, the least important of the five, since it is a logical rather than a mathematical concept, it must be admitted that this is only the expression of an opinion which can evidently not be proved to be correct. Each of these concepts has played, and will probably continue to play, a very important rôle in the development of our subject. Possibly none of them has as yet yielded its richest treasures and each seems to represent a mine of inexhaustible wealth. They are not the only fundamental mathematical concepts, but none of the others seems to have had such wide mathematical influence. The influence of many other concepts, in particular the concept of limit, has also been extensive.

It may be desirable to emphasize the fact here that there is a wide historical difference between having names for the first few natural numbers and having a general name for number. The latter represents a name for a totality and is an important step towards the concept of group with its modern technical meaning. The fact that a special name for the general concept of natural number appears in the work of Ahmes is very significant. It is this fact that led us to note above that the fundamental concept of number had received a special name already in prehistoric times, and that the real history of our subject starts after two of the five fundamental mathematical concepts noted above had been developed. We shall probably never know much about the important mathematical developments which preceded the naming of these two concepts.

The work of Ahmes contains not only names for these two fundamental concepts, besides many special illustrative examples thereof, but it also involves examples of each of the other three fundamental concepts under consideration. From the fact that the area of a rectangle was found by multiplying the numerical measures of two of its adjacent sides it is clear that the postulate that all right angles are equal was assumed, although not expressed. Similarly, the method employed in dealing with arithmetic series in this work seems to imply a formula for the sum of such a series in terms of the first term, the number of terms, and the common difference, and hence the implicit notion of function.

One of the most important groups of mathematics, namely, the one formed by the different positive rational fractions, when these fractions are combined by multiplication, is also used implicitly in this work, and the very elementary group of period two, formed by identity and the operation of finding the reciprocal of a natural number is also represented by numerous examples found therein. The concept of group seems to appear more distinctly in the work of Ahmes than that of function.

The middle historical position occupied by the fundamental concept of system of postulates among the five dominant mathematical concepts under consideration is perhaps suggestive of the rôle of logic in the development of our subject. The early developments in a mathematical subject have frequently not been very rigorous as is illustrated, in particular, by the history of the differential calculus. Then comes a period when rigour received especial attention. Later again the investigator goes ahead with perfect security but often without any direct concern as to questions of rigour since he has cultivated habits based upon reliable processes. In presenting his results in a final form he is, of course, often compelled to consider explicitly the question of rigour but this does not always affect his feeling of security. It is obvious that different fields of mathematics present very different aspects as regards this particular question, so that these remarks are true only in a general way.

It may be worth noting here that these five fundamental concepts enter explicitly in their historical order into many of our modern mathematical courses. The concept of postulate systems seems to have received the least explicit attention in this connection and constitutes a domain common to philosophy and mathematics. In recent years the concept of function has received much more direct attention than formerly in the more elementary subjects. This has been largely due to the influence of F. Klein who was also largely responsible for the greater emphasis on the explicit concept of group in many of the more advanced recent developments of mathematics. In view of the fact that important special groups present themselves implicitly already in the work of

Ahmes it seems obvious that it would be possible to enrich our elementary instruction by stressing explicitly this concept whenever we meet it naturally, but thus far this has seldom been attempted. The elementary definition of group noted above is evidently within the reach of somewhat immature students.

It was noted above that in the work of Ahmes we find the number concept sufficiently extended so that the numbers under consideration constitute a group with respect to multiplication, where the term group is used with its modern restricted meaning. The extension of the number concept so as to constitute a group with respect to addition came much later and was not completed until the beginning of the nineteenth century when we first find a satisfactory theory of negative numbers. Steps towards this extension appear in the work of Diophantus where we find some operations with numbers to be subtracted. The Indians used negative numbers as such but they, of course, failed to develop a clear theory thereof. On page 78 of volume 2 of Tropfke's *Geschichte der Elementar-Mathematik*, 1921, we find the statement that the eighteenth century suffered as a result of the fact that a general satisfactory introduction of negative numbers was missing. After the extension of the number concept so as to constitute a group with respect to addition, as well as with respect to multiplication, when the identity of addition is omitted, we arrive at the important concept of domain of rationality, early in the nineteenth century.

From the preceding paragraph it results that while the extension of the number concept, so as to constitute a group with respect to multiplication, is prehistoric, the extension so as to constitute a group with respect to addition comes within the historic period of our subject. The latter extension implies both negative numbers and the use of zero as a number. In fact, since zero is the identity of the addition group it constitutes the most conspicuous element of this group, and hence its slow acceptance as an actual number becomes of greater historic interest. Each of the five fundamental concepts whose historical development is here briefly sketched is related to elementary mathematics as well as to the more advanced parts thereof. These developments exhibit the continual enrichment of elementary mathematics through the advances in the higher parts of our subject.

VILFREDO PARETO

THE MATHEMATICIAN OF THE SOCIAL SCIENCES

By Professor James Harvey Rogers,
University of Missouri, Columbia, Missouri, U.S.A.

In his *Mind in the Making*, Professor James Harvey Robinson makes the following prophetic remark: "And now the astonishing and perturbing suspicion emerges that perhaps almost all that had passed for social science, political economy, politics, and ethics in the past may be brushed aside by future generations as mainly *rationalizing*. John Dewey has already reached this conclusion in regard to philosophy. Veblen and other writers have revealed the various unperceived presuppositions of the traditional political economy, and now comes an Italian sociologist, Vilfredo Pareto, who, in his huge treatise on general sociology, devotes hundreds of pages to substantiating a similar thesis affecting all the social sciences. *This conclusion may be ranked by students of a hundred years hence as one of the several great discoveries of our age.*"

With the death of Vilfredo Pareto, has passed one of the few mathematicians working in the field of Economics and Sociology and one of the most refreshing and picturesque scholars of all times. A marquis by birth, his long academic career never robbed him of either the tastes or the charm of manner of an Italian nobleman. Although professor at the University of Lausanne, for half of his long life he made his home in the quiet country village of Céligny, where, on the shores of the Lake of Geneva, he built the beautiful little Villa Angora. Here, with all the seclusion and comforts of a miniature country estate, and surrounded by numerous and beautiful specimens of Angora cats—his favourite pets and perpetual companions, which, in painted form, ornamented even the walls of his villa—he divided his time between writing books, and entertaining his friends. From his study, thus secluded, but filled with books of all ages and in all languages—a veritable watch tower—he observed and studied the actions, customs, prejudices, motives, feelings, and thinking of his fellow men of all times, much as would an entomologist, at a safe distance, investigate the characteristics of a society of bees industriously shaping to their whims and needs, the social and material world around them.

His objectivity and freedom from bias have rarely been surpassed in the academic field. While the English and German scholars were issuing equally partisan and contradictory manifestos on the causes of the war, he calmly discussed the same great event as though it were taking place on the planet Mars, carefully pointing out to the writer, at that time his student, the absurd lengths to which even the best of humanity will go when sufficiently moved by emotional and sentimental appeals. If he had any preference whatever for either side in the great conflict, the writer, although intimately associated with him for ten months during that historic period, was unable to discover it. Pareto could see little difference in either the motives or the methods of one side and of the other, and besides to him it seemed dangerous, from the point of view of his judgment, to judge of events in the outcome of which he was keenly interested; so, inasmuch as he wished to be an unprejudiced critic, he chose the logical but, to most judges, the impossible course of becoming disinterested.

So great was his respect for experimental truth (he recognized no other), and so frank was he in facing all problems, that to be with him during the war period was much like having one's abode on another planet. From such a point of vantage could be observed the great masses of humanity wandering aimlessly and blindly, the victims equally of their own sentiments, and of the pious, patriotic, superstitious, and other sentimental appeals of intriguing leaders, who sought ever to use them for selfish and nationalistic purposes. Nor in war time did those Lilliputian Bismarks stop at aught. Socialists, Catholics, anarchists, and atheists were all, by their ill-disguised propaganda, brought into the common fold of nationalism. Urged on by their leaders—some sincere and, because ignorant even of the methods of ascertaining the truth, stupidly believing the lies they vehemently retailed, others cynically constructing the *eternal* truths to be passed on to their less sophisticated, more gullible, patriotic brothers, who in the measure of their devotion and of their sincerity proved useful in inspiring confidence in the falsehoods they proclaimed—the honest, hard-working masses, prince and pauper, atheist and believer, landlord and peasant, scholar and imbecile, saint and crook—went arm in arm to the common slaughter, all equally sure of the sacredness of the cause for which they were fighting.

To him, the creeds of the socialist, the anarchist, the pacifist, the republican, and the democrat were little different in character from those of the Mussulman and of the Catholic. "Their theology has changed," he would say, "but their religion is all left." Man is indeed a faithloving being. Rob him of one of his gods, he will invent another. To Pareto, this was one of the many characteristics of the animal we call man, to be observed and noted along with his gregarious instincts, his love of life, and the sex impulse. "Even the physical and natural scientists," he would point out, "who in their researches have freed themselves almost completely from the 'Great Truths,' which have proved such insurmountable obstacles to the thought of non-scientific investigators, have in many cases all but deified their own laws." So accustomed has man become to living by authority rather than by thought, that in the rare cases where the latter is employed, the fiction of the former is often by invention maintained.

On the other hand, he had no illusions about the dangers of finding and of spreading the truth. There is certainly no necessary and, with the limited data available, no observable connection between the scientifically (*i.e.*, experimentally) true and the socially good. Unlike Robinson, he thought that the discovery and (probably impossible) dissemination of sociological truth was just as apt to prove our ruin as our salvation, though this opinion in no way affected his interest in discovering it.

Rationalization, the exclusive tool of the logician, the simplest and therefore perhaps the most useful device of the mathematician and of the scientist, he regarded as a frequent and, because of its intellectual appeal, an extraordinarily effective means of *concealing* the truth. Activated by instincts, sentiments, and numerous other influences, man takes the countless actions—some logical but most non-logical—which go to make up life. The logical, as the non-logical, he always explains to himself, and often to others; and being a rational animal he rarely gives other than a rational explanation. But the correspondence between the explanations, or "derivations"—as Pareto called them—and the real reasons, or "residues," is generally but slight. Ask a child why he wears a hat, and he will answer rightly, "Everybody wears them"; but ask an adult the same question and, unless on his guard, he will invent some foolish rationalization such as to keep his head warm, to protect it from the rays of the sun, or any other which bears the stamp of plausibility and of reason. This habit of rationalization is likewise one of the human characteristics to be noted by the scientist, used by the intriguer, and discounted by the objective investigator, attempting to apply to scientific research, that prejudiced, inaccurate, distorting machine, known as the human intellect. If one would understand human actions it is more important to study the "residues" than the "derivations."

The American radical with his socialist, pacifist, anarchist, or other programme appeared to him much as the early Christian with his new religion, the social reformer with his chosen panacea, the college freshman with his football team—another worshipper at a renovated and ill-disguised shrine, another child confusing faith with insight. His contempt for such radicals, who insist on rating themselves as thinkers, was almost as pronounced and as unconcealed as for the old-fashioned conservatives, venerating the time-worn institutions of a forgotten age. Both, to him, appeared equally devout worshippers, the shrines only being changed, but both equally important to study as influences shaping our society. Temperamentally, they are different, but intellectually, not far apart. Their religions pass, carrying with them the faithful, to give place to new fads with adherents just as loyal to their *self-invented*, *eternal* ideals, and just as scornful of those ill-willed obstructionists recently passed on, as their successors will be of the misguided mortals who continue to offer up sacrifices to demonstrably false gods.

Pareto's training was that of an engineer and after graduation from the University of Turin, he spent twenty years in successful business as a manager in the iron works of the Val d'Arno before turning to the academic profession. This early practical experience undoubtedly accounts for much of the poise and objectivity which characterize his work as that of few other investigators in the field of theoretical economics and sociology. With him no theory was so good but that it must yield completely to the authority of observed facts, and no facts were so trivial as to merit being disregarded. Nevertheless, every problem was a practical one requiring a practical answer—not a philosophic explanation—and as most problems were much too complicated for exact solution, and too serious for mere logical discussion, he early began attacking them in the way he had found most effective in his engineering undertakings. Economic laws—as all other scientific laws—were to him mere convenient generalizations of observed uniformities, mere approximations more or less good according as they fit the facts, and important to the extent that they were useful in adapting complicated situations to the limitations of our own intelligence. Such laws, must, on account of their very nature, be displaced if the much desired scientific progress is to continue. He expected and hoped that they would give way to broader generalizations on the one hand and to closer approximations on the other. This attitude led Pareto to the extensive use of the method of approximation, which characterizes much of his later work. It was his firm conviction that the experimental method, which has yielded such brilliant results in the natural sciences, could be applied with equal success to the social sciences, and I think few who will read understandingly his great work on sociology will be inclined to debate his contention. Furthermore he had little faith in any other method, at least until such time as worthwhile results should be shown.

Most economic and sociological theory appeared to him as mere useless rationalization or else as a vain and superficial search for non-existent great principles, which the philosophically inclined expected ever to discover. He believed, and I think rightly, that this attitude toward scientific laws as great principles was all but universal, except with a small minority of the natural scientists. He took great delight, therefore, in the revelations of Einstein in the field of relativity, because his researches dissipated forever the absolute character of the greatest of the great principles, and reduced it to a mere generalization of observed uniformities, like most other laws with which we are familiar.

The avowed aim of his great work on sociology was to apply to that subject the experimental method. Accordingly he divided human activity into two principal branches, that of sentiment and that of experimental research. The first, he contended over and over again, could not be too much emphasized, as it is sentiment that drives to action, gives life to the laws of morality, to worship, to religion in all its forms socially so important and so complex and varied. "It is by aspiration to the ideal that human societies subsist and progress."* But the second branch, that of experimental research, though more modest in its claims, is also essential for all societies. "It furnishes the material which the first uses. We owe to it the knowledge which makes action effective.

All sciences, natural as well as social, have involved at their beginning a mixture of sentiments and of experiences. Centuries have been necessary to secure a separation of these elements; which separation, in our time, has been nearly entirely accomplished for the natural sciences, and has commenced and progressed for the social sciences."* And to advance this separation, Pareto devoted the last twenty years of his life.

*Taken from address of V. Pareto at the University of Lausanne in 1917 on the occasion of a jubilee commemorating the twenty-fifth anniversary of his appointment to the faculty of that university.

OBSERVATIONS PRATIQUES DE MÉTHODOLOGIE

PAR M. L. J. CRELIER,
Professeur à l'Université de Berne, Berne, Suisse.

Les Universités des petits États ont également pour but de préparer le corps enseignant secondaire nécessaire au pays. En conséquence, l'éducation professionnelle des candidats doit marcher de pair avec leur éducation scientifique.

A ce sujet, les professeurs de mathématiques de ces Universités ne peuvent pas se désintéresser de la formation professionnelle de leurs élèves. Au contraire, ils doivent s'efforcer d'exercer dans ce sens une influence bienfaisante de premier ordre.

Les étudiants de toutes les directions sont, en règle générale, astreints à suivre divers cours de pédagogie et de philosophie de caractère général. A côté de cela viennent la méthodologie et les exercices pratiques; c'est ici que l'influence des professeurs de mathématiques peut se faire sentir le plus efficacement.

Nous avons suivi depuis longtemps les essais tentés dans diverses Universités, surtout depuis les travaux de l'*Internationale mathematische Untersuchungskommission*. Nos expériences personnelles nous ont montré que l'on pouvait obtenir d'excellents résultats en discutant, d'une part, les diverses questions de méthodologie mathématique au séminaire, et, d'autre part, en suivant de près les exercices des étudiants dans les divers établissements d'instruction secondaire où ils ont l'occasion d'enseigner pratiquement à une classe réelle et d'après un programme normal.

I. Dans le *séminaire* où les jeunes gens présentent leurs conférences, toutes les questions de programme, de manuel, de mouvement pédagogique peuvent donner lieu à d'intéressants travaux de leur part, travaux que nous faisons toujours suivre d'une discussion fructueuse.

Diverses questions de mathématiques élémentaires et de mathématiques spéciales, considérées au point de vue de l'enseignement, et vues sous le jour des connaissances supérieures que les candidats ont acquises à l'Université, forment ensuite l'objet de nouvelles conférences.

II. L'*enseignement pratique*, auquel nous avons fait allusion, comprend un stage ininterrompu de quelques semaines dans un gymnase, sous la direction du même professeur de l'école, stage pendant lequel le candidat assiste aux leçons et en donne quelques-unes. Il comprend, en outre, une série de leçons d'épreuves dans la même école, leçons données cette fois en présence des professeurs de l'Université et des autres camarades d'études; la préparation pratique est complétée généralement par un deuxième stage où le candidat fonctionne comme remplaçant complet d'un professeur du gymnase empêché de donner ses cours pour une raison quelconque.

Les leçons d'épreuve ont chez nous une importance très considérable parce qu'elles sont exigées dans les examens d'État et que les Commissions scolaires les demandent également pour la nomination des nouveaux professeurs.

Nous y attachons, personnellement, une très grande importance et nous nous efforçons à ce que nos élèves fassent ressortir en cette occasion toutes les qualités pédagogiques que l'on est en droit d'attendre d'eux. Nous demandons que chaque leçon soit caractérisée par les traits essentiels suivants:

(*a*) *La préparation.* Le sujet doit être étudié consciencieusement. On doit retrouver la liaison avec les thèses précédentes, la concentration sur l'objet indiqué et les conclusions naturelles auxquelles conduit la leçon.

(*b*) *L'action.* A côté des qualités indispensables de tenue, de style et de langage, de clarté et de précision, la leçon doit être vivante. L'attention des élèves doit être constamment tenue en éveil, et, d'un bout à l'autre, leur intérêt doit être soutenu sans jamais fléchir.

(*c*) *Le cœur.* Le candidat qui donne sa leçon doit montrer qu'il sait être juste et bon, sévère quand il faut, mais non brutal. Il doit y aller de tout son cœur et les élèves doivent sentir l'affection réelle que leur porte celui qui leur parle.

L'enseignement, et surtout l'enseignement des mathématiques, *n'est pas un métier, c'est une vocation.*

PROPOSITIONS CONCERNANT L'UNIFICATION DE LA TERMINOLOGIE DANS LA NUMÉRATION PARLÉE

PAR M. L. GUSTAVE DU PASQUIER,
Professeur à l'Université de Neuchâtel, Neuchâtel, Suisse.

1. Les Sections de Mathématiques et d'Astronomie de l'Association Française pour l'Avancement des Sciences, sur ma proposition, émettaient au Congrès de Bordeaux en 1923 un voeu relatif à l'unification de la terminologie dans la numération parlée. Cette même question: comment unifier la terminologie de la numération arithmétique dans les principales langues, figure aussi cette année à l'ordre du jour du Congrès de Liège*. Il serait désirable qu'une autorité internationale reconnue par tout le monde exprimât un préavis qui permettrait de trancher la question. C'est pourquoi je la soumets au Congrès International de Mathématiques à Toronto. Avant de formuler mes propositions, je vais en expliquer la raison et indiquer très succinctement comment le problème a évolué et comment la question se pose actuellement.

2. En faisant une arithmétique comparée des divers peuples de la terre, j'ai trouvé qu'on peut répartir leurs systèmes de numération en deux grandes catégories.

(I) *Les systèmes purement additifs.* On dénomme certains nombres par des noms indépendants, puis, à l'aide de ces éléments constitutifs, on forme les noms de tous les autres nombres par composition additive uniquement, comme le fait par exemple la numération écrite romaine en partant des éléments I, V, X, L, C, D, M. On peut ranger ces systèmes additifs en trois classes sur lesquelles je ne m'étendrai pas davantage ici†.

(II) *Les systèmes multiplicatifs.* Ils admettent un principe de multiplication à côté du principe d'addition. Ils présentent dans la numération écrite des avantages tels que leur emploi s'impose, et le système de position à base 10, imaginé par les savants hindous, est en train de conquérir tout le globe, bien que (soit dit en passant) cette base $b=10$ soit loin d'être la meilleure; les bases $b=6$ ou $b=4$ seraient de beaucoup préférables‡.

*A ce congrès, dans l'assemblée générale du 2 août 1924, l'Association Française pour l'Avancement des Sciences a adopté un vœu analogue proposé par les Sections de Mathématiques, d'Astronomie et de Géodésie.

†Pour plus de détails, voir L. Gustave Du Pasquier, *Le développement de la notion de nombre.* Mém. de l'Université de Neuchâtel (Suisse), t. III, Neuchâtel et Paris 1921, VIII+191 p. L. Gustave Du Pasquier, *Étude comparative des systèmes de numération parlée.* Bulletin Soc. Neuchâteloise de Géographie, t. XXX, 1921, p. 19-49.

‡v. chap. VIII du livre ci-dessus.

Si l'on fait abstraction des irrégularités plus ou moins nombreuses dans la formation des noms numéraux, on peut classifier les systèmes multiplicatifs en trois catégories, d'importance d'ailleurs très inégale. Ce sont:

(*a*) Les systèmes *multiplicatifs complets*. On doit avoir à sa disposition des noms individuels pour désigner 1° chacune des unités simples, 1, 2, 3, ..., $b-1$; 2° chacune des puissances de la base, b, b^2, b^3, ..., b^n, aussi loin qu'on veut pousser la numération. A l'aide de ces éléments constitutifs, les noms de nombre, en prenant la base $b=10$, sont formés suivant le schéma

$$a_n . 10^n + a_{n-1} . 10^{n-1} + \ldots + a_3 . 10^3 + a_2 . 10^2 + a_1 . 10 + a_0 \tag{1}$$

avec les conditions $0 \leqq a_\lambda < 10$ ($\lambda = 0, 1, 2, \ldots, n$) résultant de ce que les a_λ représentent des unités simples. Tout nombre naturel est représentable, et d'une seule manière, sous cette forme, même si l'on prend comme base b un nombre naturel quelconque supérieur à 1. Le sanscrit fournit l'exemple le plus célèbre d'un tel système de numération.

(*b*) Les systèmes *multiplicatifs à progression arithmétique*. Les éléments constitutifs en sont (en prenant de nouveau la base $b=10$): 1° les noms des neuf premiers nombres; 2° les noms des r premières puissances de la base, $10, 10^2, 10^3, \ldots, 10^r$; pour fixer les idées, prenons $r=3$; 3° les noms de certaines autres puissances de la base, savoir celles dont les exposants forment la progression arithmétique en question; dans notre cas particulier

$$10^3,\ 10^6,\ 10^9,\ 10^{12},\ \ldots,\ 10^{3n}.$$

De ces $11+n$ éléments constitutifs, on peut déduire par composition additive et multiplicative le nom d'un nombre entier quelconque inférieur à 10^{3n+3}, en suivant la formule

$$(\alpha_n . 10^2 + \beta_n . 10 + \gamma_n) 10^{3n} + \ldots + (\alpha_3 . 10^2 + \beta_3 . 10 + \gamma_3) 10^9$$

$$+ (\alpha_2 . 10^2 + \beta_2 . 10 + \gamma_2) 10^6 + (\alpha_1 . 10^2 + \beta_1 . 10 + \gamma_1) 10^3 + \alpha_0 . 10^2 + \beta_0 . 10 + \gamma_0,$$

où les α_λ, les β_λ et les γ_λ, ou bien sont nuls, ou bien représentent des unités simples, donc des nombres naturels inférieurs à 10.

Exemple: La numération française actuelle (abstraction faite des irrégularités), puisqu'elle a pour les surunités les noms *dix* $=10$, *cent* $=10^2$, *mille* $=10^3$, puis

10^6 = *million*,	10^{15} = *quadrillion*,	10^{24} = *septillion*,
10^9 = *milliard* (ou *billion*),	10^{18} = *quintillion*,	10^{27} = *octillion* ,
10^{12} = *trillion*,	10^{21} = *sextillion*,	10^{30} = *nonillion*.

C'est de là que vient *la règle latine:* pour énoncer un nombre écrit dans le système de position, il faut au préalable grouper ses chiffres par tranches de trois, en commençant par les unités simples.

(*c*) Les systèmes *multiplicatifs à progression géométrique.* Il faut dénommer individuellement, par des noms indépendants, les surunités suivantes:

$$b,\ b^2,\ b^4,\ b^8,\ b^{16},\ b^{32},\ \ldots,\ b^{2^n}.$$

Les éléments constitutifs d'un tel système, si l'on prend $b=10$, sont alors (1) les noms des unités simples, 1, 2, 3, ..., 9; puis (2) les noms des surunités 10, 10^2, 10^4, 10^8, 10^{16}, 10^{32}, ... On n'arrive guère à un tel système que par spéculation mathématique, en cherchant la solution du problème suivant: Quel est *le minimum* d'éléments constitutifs permettant de compter jusqu'à N, si l'on exclut des expressions telles que cent cents ou mille fois mille ou d'autres analogues avec réduplication? Par exemple, à l'aide de 20 noms indépendants, on pourrait dénommer univoquement les nombres jusqu'à ceux qui s'écrivent avec 2048 chiffres.

3. Les mots de *million* pour 10^6 et de *milliard* pour 10^9 sont actuellement adoptés par tous les peuples civilisés et, chose plus remarquable encore, ils ont actuellement partout le même sens. Grâce aux innombrables relations commerciales et scientifiques, ces deux noms numéraux sont devenus des termes internationaux à signification précise et univoque. On est tenté de croire que billion, trillion, etc., eurent la même bonne fortune. Mais il n'en est rien. On doit au contraire distinguer nettement entre deux groupes de langues.

(α) Chez quelques peuples dont la langue se rattache au groupe italique, c'est 10^3, *mille,* dernier nombre anciennement dénommé par un nom primaire, qui devint une nouvelle espèce de surunité. Elle servit par la suite à former un système décimal multiplicatif à progression arithmétique de raison 3. Il trouve son expression dans la règle latine susmentionnée.

(β) La plupart des peuples, en particulier ceux dont la langue appartient au groupe germanique, ont au contraire formé un système décimal multiplicatif à progression arithmétique de raison 6. Chez eux, c'est le million, 10^6, qui est la nouvelle espèce de surunité, de sorte que les puissances supérieures de la base dénommées par des noms primaires sont

$$10^6,\ 10^{12},\ 10^{18},\ 10^{24},\ \text{en général } 10^{6n}.$$

En voici le tableau:

10^6 = *million,*	10^{24} = *quadrillion,*	10^{42} = *septillion,*
10^{12} = *billion,*	10^{30} = *quintillion,*	10^{48} = *octillion,*
10^{18} = *trillion,*	10^{36} = *sextillion,*	10^{54} = *nonillion.*

Il en est résulté *la règle germanique:* pour énoncer un nombre écrit dans le système de position, on groupe préalablement ses chiffres par tranches de six, en commençant par les unités simples.

4. Actuellement, les mots de billion, de trillion, de quintillion, etc., ont chacun deux significations bien différentes, suivant que ces termes se trouvent dans un livre anglais, allemand, espagnol, danois, hollandais, suédois, etc., ou dans un écrit français, italien, etc.

Dans le premier cas,

$$1 \text{ billion} = 10^{12},\ 1 \text{ trillion} = 10^{18},\ 1 \text{ quintillion} = 10^{30}, \text{ etc.}$$

Dans le deuxième cas,

$$1 \text{ billion} = 10^{9},\ 1 \text{ trillion} = 10^{12},\ 1 \text{ quintillion} = 10^{18}, \text{ etc.}$$

Quand les journaux allemands disaient que les dettes de l'empire dépassaient 1 trillion de marks, ils entendaient par là un tout autre nombre que le lecteur français, savoir $10^{18} = 1\ 000\ 000\ 000\ 000\ 000\ 000$, chiffre que le Français énonce actuellement «un quintillion».

Or, chose remarquable, il n'en fut pas toujours ainsi. Pendant plus de deux siècles, la règle dite germanique était suivie en France également. Elle y était encore en usage à la fin du XV^e^ et au XVI^e^ siècle. Preuve en soit entre autres le fameux *Triparty en la science des nombres*, écrit à Lyon par Nicolas Chuquet en 1484, ou encore *l'Arithmétique* de J. Peletier, livre paru à Poitiers en 1551–52. Mais à partir du milieu du XVII^e^ siècle, la règle dite latine est adoptée à peu près dans toute la France.

Si la diversité est parfois utile, indispensable même, il est des domaines où elle ne peut que prêter à confusion et là, il vaut mieux tendre vers l'uniformité. La numération arithmétique est certainement l'un de ces domaines. N'est-il pas regrettable que dans un même pays, par exemple en Suisse, on enseigne aux enfants d'une partie de la nation que 1 billion = mille millions, et aux enfants de l'autre partie que 1 billion = 1 million de millions? Ou encore, aux uns, que les mots «un trillion, onze billions six cénts mille» signifient le nombre 1 011 000 600 000, et aux autres, que les mêmes mots signifient, au contraire, le nombre 1 000 011 000 000 600 000?

Il est temps qu'une autorité compétente crée, dans le domaine restreint qui nous occupe, l'uniformité que tout le monde désire.

5. Le système de numération à choisir pour être adopté internationalement devrait satisfaire aux postulats suivants.

(1) *Simplicité.* Les noms des surunités des divers rangs doivent se former suivant une règle simple.

(2) *Régularité.* Cette règle doit s'appliquer d'une manière conséquente, sans exception.

(3) *Correspondance.* On doit pouvoir déduire facilement du nom d'une surunité le nombre des chiffres qui figurent dans son expression écrite.

(4) *Million.* Le terme de million, étant déjà universellement adopté et d'ailleurs commode, doit figurer dans la liste avec le sens de 10^6.

(5) *Milliard.* Le terme de milliard ayant également acquis droit de cité, et partout avec le sens de mille millions, doit être si possible conservé pour désigner 10^9.

Ces conditions peuvent toutes être remplies par un système décimal multiplicatif à progression arithmétique de raison 3(v. § 2, *b*), c'est-à-dire ayant des noms primaires pour 10^3, 10^6, 10^9, 10^{12}, ..., 10^{3n}.

Tout est déjà fixé jusqu'à 10^9 par les postulats (4) et (5). Quant à 10^{12}, Anglais, Allemands, Danois, Espagnols, Hollandais, Suèdois, etc., l'appellent déjà billion, et les Français l'ont appelé ainsi pendant plus de deux siècles. La synonymie actuelle: 1 milliard = 1 billion, n'ajoute rien à la clarté de la langue française. Cette synonymie est superflue. Il serait donc rationnel d'y renoncer, en gardant le terme de milliard pour 10^9 et réservant celui de billion pour 10^{12}, comme on le faisait au XV^e^, au XVI^e^ et au XVII^e^ siècle.

Dès lors, ayant million pour 10^6, billion pour 10^{12}, il serait logique de continuer par trillion pour 10^{18}, quadrillion pour 10^{24}, etc., comme cela se fait déjà dans la plupart des langues, et dans les plus répandues. On satisferait ainsi au postulat (3). Pour remplir également le postulat (2) et tenir compte de l'apport du génie français, je proposerais de créer par analogie, sur le modèle de milliard et de billiard, les néologismes de *trilliard*, *quadrilliard*, etc. On aurait de cette façon l'échelle numérale suivante, apte à devenir internationale.:

10^6 = *million*,	10^{24} = *quadrillion*,	10^{42} = *septillion*,
10^9 = *milliard*,	10^{27} = *quadrilliard*,	10^{45} = *septilliard*,
10^{12} = *billion*,	10^{30} = *quintillion*,	10^{48} = *octillion*,
10^{15} = *billiard*,	10^{33} = *quintilliard*,	10^{51} = *octilliard*,
10^{18} = *trillion*,	10^{36} = *sextillion*,	10^{54} = *nonillion*,
10^{21} = *trilliard*,	10^{39} = *sextilliard*,	10^{57} = *nonilliard*.

Pour 10^{60}, je proposerais *millionarde*, terme qui devrait être considéré comme une nouvelle espèce de surunité. Pour énoncer des nombres très grands, on grouperait au préalable leurs chiffres par tranches de soixante, en commençant par les unités simples. A partir de là, le système de numération parlée devrait être à progression géométrique (v. § 2, *c*). C'est dire qu'on devrait compter des millions de millionardes, des billions de millionardes etc. jusqu'à 10^{120}, et que les surunités à dénommer spécialement seraient

$$10^{60},\ 10^{60\cdot 2},\ 10^{60\cdot 4},\ 10^{60\cdot 8},\ 10^{60\cdot 16},\ 10^{60\cdot 32},\ \ldots$$

Avec les termes de *bimillionarde* pour 10^{120}, de *trimillionarde* pour 10^{240}, *quadrimillionarde* pour 10^{480} et *quintimillionarde* pour 10^{960}, on pourrait déjà compter jusqu'aux nombres de 1919 chiffres. D'ailleurs, le terme de «*n*-imillionarde» pour $10^{60\cdot 2^{n-1}}$, où l'on pourra substituer à *n* un nombre aussi grand que l'on voudra, ouvre à la numération parlée le champ de l'infini.

6. Les considérations qui précèdent m'ont amené aux conclusions suivantes:

(1) Les progrès scientifiques, le développement des affaires commerciales et en général le degré de civilisation ont eu pour conséquence que le besoin de très grands nombres s'est fait sentir. Dans la numération parlée, il y a actuellement deux systèmes différents qui se font concurrence. Leur coexistence,

explicable au point de vue historique, n'a plus aucune raison d'être et ne peut que prêter à confusion. Une unification dans ce domaine est désirable et serait dans l'intérêt de tout le monde.

(2) Les mathématiciens devraient donner le bon exemple en tombant d'accord d'adopter un système unique.

(3) Ce système, pour ne froisser aucun amour-propre national, doit tenir compte des deux courants en présence. Le système que je propose paraît apte à devenir international, parcequ'il remplit toutes les conditions requises de simplicité, de régularité et d'adaptation à l'état de fait actuel.

Je me permets donc de proposer au Congrès International de Toronto d'accorder à ces conclusions l'appui de son autorité, éventuellement de créer un organe chargé d'étudier la question de l'unification de la terminologie dans la numération parlée, puis d'inviter les autorités compétentes des divers pays à faire introduire dans l'enseignement la solution adoptée par le Congrès. J'espère que mes confrères s'intéresseront à la question et voudront aider à réaliser l'accord.

ABSTRACTS OF COMMUNICATIONS

SECTION VI

COLONIAL AMERICAN ARITHMETICS

By Professor L. C. Karpinski,
University of Michigan, Ann Arbor, Michigan, U.S.A.

The English arithmetics published in America before the Revolutionary War were largely reprints of texts popular in England. In this paper the author lists all the arithmetics known which were published in the new world before 1775 and discusses the sources of the popular texts in the English language.

THE MATHEMATICAL WORKS OF SIR W. R. HAMILTON

By Professor A. W. Conway,
University College, Dublin, Irish Free State.

A short account of the published and unpublished works of Sir W. R. Hamilton.

THE DOCTRINAL FUNCTION: ITS RÔLE IN MATHEMATICS AND GENERAL THOUGHT

BY PROFESSOR C. J. KEYSER
Columbia University, New York, N.Y., U.S.A.

The concept of a doctrinal function defined. The function's characteristic properties: autonomy, compatibility, compendence, fertility. The genesis of doctrinal functions. Their character as forms devoid of content. Their relation to doctrines (true or false). Simple and compound functions. Their service in economising intellectual energy. Their rôle in "the logic of discovery", in criticism, in the rationalisation of empirical science whether physical or moral. Question respecting the possibility of one supreme function embracing all others.

TIME-BINDING: THE GENERAL THEORY

By Count Alfred Korzybski,
Fifth Avenue Bank, New York, N.Y., U.S.A.

Dependence of human knowledge on the properties of light and sound (speech). Importance of correct symbolism and its conditions. "Organism as a whole" and "Joint phenomenon"—two fundamental principles. Applications. The Anthropometer. The mechanism of time-binding. Confusion of types and orders. The problem of meaning, its solution. Geometrical structure of all human knowledge. Consequences. Theory of universal agreement. Its effect upon educational and scientific methods and the revision of doctrines in general. The connection between correct symbolism, postulational methods, "Doctrinal Function" (Keyser), and modern physico-mathematical developments. Deductive "natural" and "social" sciences. The deductive science of Man.

The Anthropometer was exhibited and demonstrated.

L'UNIVERSITÉ ET LA PRÉPARATION DES PROFESSEURS DE MATHÉMATIQUES

PAR M. HENRI FEHR,
Professeur à l'Université de Genève, Genève, Suisse.

L'auteur examine le rôle que doit jouer l'Université dans la préparation des professeurs de mathématiques de l'enseignement secondaire. Bien que la recherche scientifique doive rester au premier plan du but de l'enseignement supérieur, l'Université ne doit pas perdre de vue sa mission vis à vis de l'enseignement secondaire auquel elle doit fournir de bons maîtres.

L'auteur s'attache plus particulièrement à la partie scientifique de la préparation professionelle des maîtres. Elle doit comprendre notamment une étude approfondie des principes fondamentaux des mathématiques, ainsi que de la méthodologie et de la didactique mathématique. Cet enseignement ne doit pas être donné sous la forme d'un cours ayant un caractère dogmatique, mais plutôt sous la forme de conférences auxquelles les candidats eux-mêmes sont appelés à prendre une part active. C'est ici qu'il convient d'appliquer la devise americaine *learning by doing* (apprendre en agissant). Ces conférences, faites sous la direction d'un professeur, suivant un plan bien ordonné, comprendront, par exemple, l'étude des concepts fondamentaux, le rôle des définitions en mathématiques, l'examen de traités classiques en usage dans les principaux pays, etc.

Il y a aussi lieu de signaler l'œuvre de la Commission internationale de l'enseignement mathématique et de faire connaître les documents relatifs aux pays environnants.

DE AEQUALITATE*

Del Professore G. Peano,
Università di Torino, Torino, Italia.

Aequalitate es indicato per ae, initiale de "aequatur" deformato in ∞ ab Vieta (mortuo in 1603) ad Leibniz (m. 1716). Recorde, a. 1557, introduce signo $=$, adoptato ab Newton (m. 1727), et nunc de usu universale.

Relatione $=$ habe tres proprietate sequente:

1. $x = x$
2. si $x = y$, tunc $y = x$
3. si $x = y$, et $y = z$, tunc $x = z$.

Me adopta symbolo de logica-mathematica, $\supset$ pro indica deductione. Punctos divide propositione in partes, in loco de parenthesi. Tunc propositiones 2 et 3 sume forma:

2. $x = y . \supset . y = x$
3. $x = y . y = z . \supset . x = z$

Si nos indica per xRy uno relatione inter duo ente x et y que pertine ad uno campo (classe), dicto "campo de R", tunc:

1. Relatione R es *reflexivo* si, pro omne x in campo R, es xRx
2. Relatione R es *symmetrico*, si pro omne x et y in campo R, $xRy . \supset . yRx$.
3. Relatione R es *transitivo*, si $xRy . yRz . \supset . xRz$.

Vocabulo *transitivo* es introducto ab De Morgan, anno 1856.

Vocabulo *symmetrico*, in isto significatione, ab Schröder, anno 1890.

Vocabulo *reflexivo* ab Vailati, anno 1891.

Vide:

Whitehead and Russell, *Principia mathematica*, anno 1910, pag. 178.

Shearman, *The development of Symbolic Logic*, London 1906.

In praesente scripto, me demonstra quod proprietates 1 reflexivo, 2 symmetrico, 3 transitivo, es inter se independente; suffice de expone exemplo de relatione que habe duo proprietate et nos omne tres.

Relatione "distantia de duo puncto x et y es minore de uno metro", es relatione reflexivo, symmetrico, et non transitivo. Alio exemplo es "numeros (naturale) x et y habe divisore commune (maiore de uno)".

*L'articolo qui pubblicato è scritto in "Latino sine flexione" nella quale lingua tutte le parole sono latine, sotto forma del tema (ablativo o imperativo); non c' è grammatica.

Relatione inter numeros "$x \geqq y$" es reflexivo et transitivo, non symmetrico.

Alio exemplo "x es multiplo de y", "x es potestate de y".

Si campo de relatione R consta ex n elemento: 1, 2, . . . n, tunc relatione R consta de classe de dyades (x,y). Si nos repraesenta dyade (x,y) per puncto de abscissa horizontale x et de ordinata verticale y, et si in isto puncto nos scribe $+$ si dyade pertine ad relatione, et scribe $-$ si non, tunc omne figura de $+$ et de $-$ repraesenta relatione.

Pro $n=2$, figura $\begin{array}{c|cc} & 1 & 2 \\ \hline 1 & - & - \\ 2 & - & - \end{array}$ repraesenta relatione nullo, satisfacto per nullo dyade.

$\begin{array}{c|cc} & 1 & 2 \\ \hline 1 & + & + \\ 2 & + & + \end{array}$ repraesenta relatione satisfacto ab omne dyade.

$\begin{array}{c|cc} & 1 & 2 \\ \hline 1 & + & - \\ 2 & - & + \end{array}$ repraesenta $x=y$.

$\begin{array}{c|cc} & 1 & 2 \\ \hline 1 & - & + \\ 2 & + & - \end{array}$ repraesenta x non $=y$.

Relatione es reflexivo, si contine omne dyade $x=y$, que es repraesentato per punctos de diagonale. Pote contine etiam alios elemento.

Relatione es symmetrico si suo figura es symmetrico pro diagonale.

Figura $\left|\begin{array}{cc} + & - \\ - & - \end{array}\right|$ repraesenta relatione symmetrico et transitivo, non reflexivo; significa $x=1$ et $y=1$; es producto logico de conditione in x per identico conditione in y.

"x et y es numero primo" es relatione symmetrico et transitivo, non reflexivo; et es producto de conditione in x per conditione in y.

Figura $\left|\begin{array}{ccc} + & - & - \\ - & + & - \\ - & - & - \end{array}\right|$ repraesenta relatione symmetrico transitivo, non reflexivo, et non producto logico de conditiones in x et in y.

Relatione inter numeros x et y "residuo de x pro 3 es differente ab 0, et aequa residuo de y pro 3", es symmetrico, transitivo, non reflexivo, et non producto logico de conditiones in x et in y.

Figura $\left|\begin{array}{cccc} + & + & - & - \\ + & + & - & - \\ - & - & + & - \\ - & - & - & - \end{array}\right|$ repraesenta alio relatione symmetrico, transitivo, non reflexivo.

TABLE OF CONTENTS

VOLUME II

CONTENTS OF VOLUME II

ABSTRACTS OF COMMUNICATIONS

Section III

COMMUNICATIONS

Section IV

ABSTRACTS OF COMMUNICATIONS

SECTION IV

COMMUNICATIONS

Section V

ABSTRACTS OF COMMUNICATIONS

SECTION V

COMMUNICATIONS

SECTION VI

ABSTRACTS OF COMMUNICATIONS

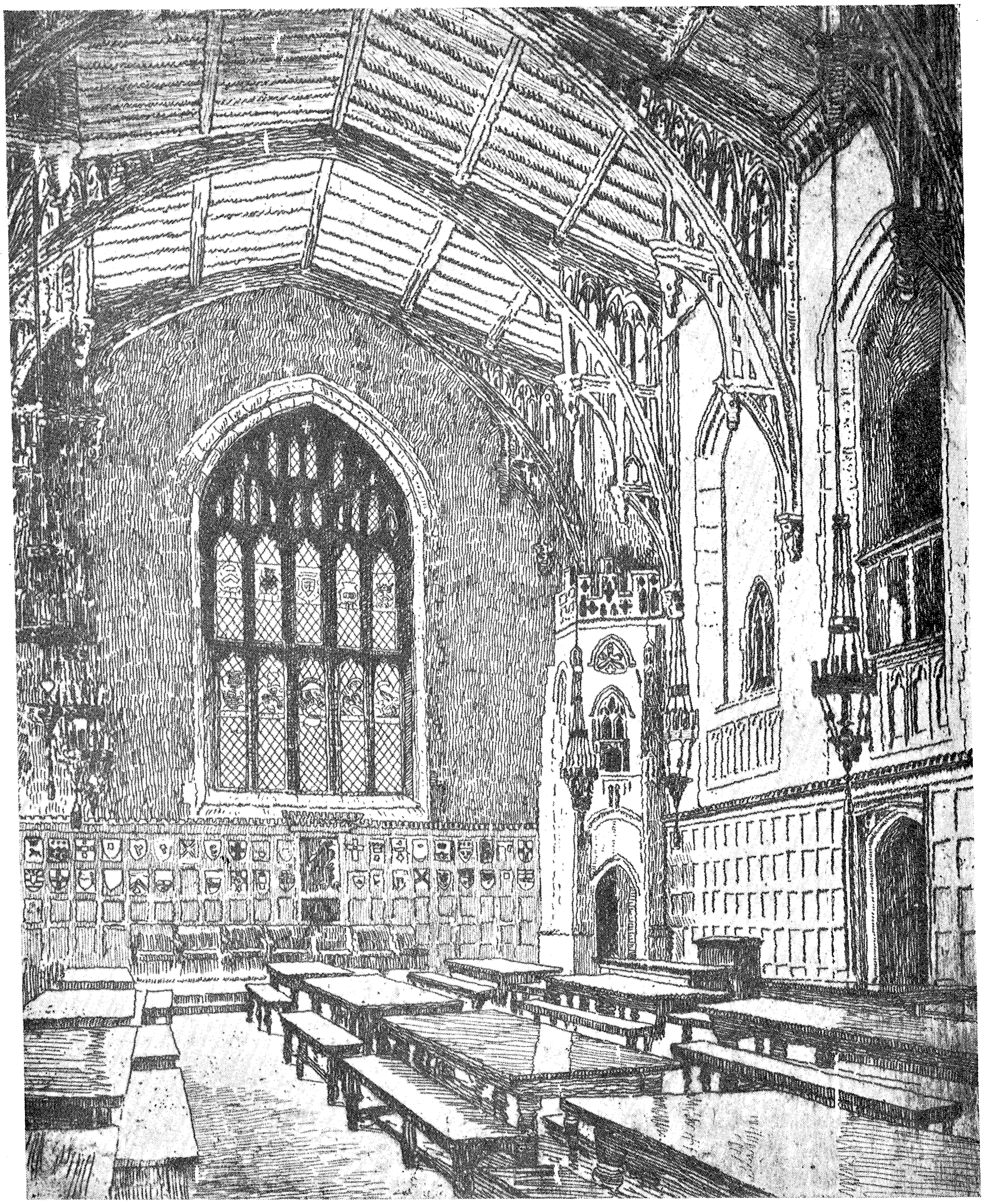
SECTION VI

INDEX OF NAMES

VOLUMES I AND II

INDEX OF NAMES

Roman numerals indicate volume; arabic numerals, page numbers. The reader will note that all page numbers after page 84 in Volume I refer to lectures, communications or abstracts of communications. All page numbers in Volume II refer to communications or abstracts of communications.



Great Hall, Hart House

Owen Staples