QUARTERLY

OF

APPLIED MATHEMATICS

VOLUME IV JANUARY · 1947 NUMBER 4

QUARTERLY

OF

APPLIED MATHEMATICS

This periodical is published quarterly under the sponsorship of Brown University, Providence, R.I. For its support, an operational fund is being set up to which industrial organizations may contribute. To date, contributions of the following industrial companies are gratefully acknowledged:

BELL TELEPHONE LABORATORIES, INC.; NEW YORK, N. Y.,
THE BRISTOL COMPANY; WATERBURY, CONN.,
CURTISS WRIGHT CORPORATION; AIRPLANE DIVISION; BUFFALO, N. Y.,
EASTMAN KODAK COMPANY; ROCHESTER, N. Y.,
GENERAL ELECTRIC COMPANY; SCHENECTADY, N. Y.,
GULF RESEARCH AND DEVELOPMENT COMPANY; PITTSBURGH, PA.,
LEEDS & NORTHRUP COMPANY; PHILADELPHIA, PA.,
PRATT & WHITNEY, DIVISION NILES-BEMENT-POND COMPANY; WEST HARTFORD, CONN.,
REPUBLIC AVIATION CORPORATION; FARMINGDALE, LONG ISLAND, N. Y.,
UNITED AIRCRAFT CORPORATION; EAST HARTFORD, CONN.,
WESTINGHOUSE ELECTRIC AND MANUFACTURING COMPANY; PITTSBURGH, PA.

The QUARTERLY prints original papers in applied mathematics which have an intimate connection with application in industry or practical science. It is expected that each paper will be of a high scientific standard; that the presentation will be of such character that the paper can be easily read by those to whom it would be of interest; and that the mathematical argument, judged by the standard of the field of application, will be of an advanced character.

Manuscripts submitted for publication in the QUARTERLY OF APPLIED MATHEMATICS should be sent to the Managing Editor, Professor W. Prager, Quarterly of Applied Mathematics, Brown University, Providence 12, R. I., either directly or through any one of the Editors or Collaborators. In accordance with their general policy, the Editors welcome particularly contributions which will be of interest both to mathematicians and to engineers. Authors will receive galley proofs only. Seventy-five reprints without covers will be furnished free; additional reprints and covers will be supplied at cost.

The subscription price for the QUARTERLY is $6.00 per volume (April-January), single copies $2.00. Subscriptions and orders for single copies may be addressed to: Quarterly of Applied Mathematics, Brown University, Providence 12, R.I., or to 450 Ahnaip St., Menasha, Wisconsin.

Entered as second class matter March 14, 1944, at the post office at Providence, Rhode Island, under the act of March 3, 1879. Additional entry at Menasha, Wisconsin.

GEORGE BANTA PUBLISHING COMPANY, MENASHA, WISCONSIN

QUARTERLY

OF

APPLIED MATHEMATICS

VOLUME IV 1946

Printed by the
George Banta Publishing Company
Menasha, Wisconsin

CONTENTS

QUARTERLY OF APPLIED MATHEMATHICS

Vol. IV JANUARY, 1947 No. 4

THE REFLECTION OF AN ELECTROMAGNETIC PLANE WAVE BY AN INFINITE SET OF PLATES, I*

BY

J. F. CARLSON[1] AND A. E. HEINS[2]

Radiation Laboratory,[3] Massachusetts Institute of Technology

1. Introduction. It has been shown by J. Schwinger that a special class of boundary value problems in electrodynamics can be formulated mathematically as Wiener-Hopf[4] integral equations. These problems may be described as follows. A plane wave is incident upon a number of semi-infinite parallel metallic structures of zero thickness and perfect conductivity. By parallel structures we mean parallel planes or cylinders with parallel axes. It is then possible to express the electric or magnetic field at all points in space in terms of the surface current density on the metal with the aid of an appropriate Green's function. The vanishing of the components of the electric field which are tangential to the semi-infinite cylindrical metallic surfaces, leads to a system of inhomogeneous integral equations for the various surface current densities. This system of integral equations assumes the general form

$$g_i(x) = \sum_{j=1}^{n} \int_0^\infty K_{ij}(x-y) f_j(y) dy, \qquad x > 0, \quad i = 1, \cdots, n,$$

where the $f_j(y)$ are unknown functions, while the $K_{ij}(x)$ and $g_i(x)$ are known. The particular problem which we shall discuss below possesses certain periodicities, and for this case we find it possible to reduce the system to a single integral equation of the form

$$g(x) = \int_0^\infty K(x-y) f(y) dy, \qquad x > 0, \tag{1.1}$$

that is, an inhomogeneous Wiener-Hopf integral equation. Here $f(y)$ is unknown, while $K(x)$ and $g(x)$ are known functions.

The advantage of formulating this particular class of boundary value problems as Wiener-Hopf integral equations is that such equations are susceptible to a rigorous

* Received April 3, 1946.

[1] Now at Iowa State College, Ames, Iowa.

[2] Now at the Carnegie Institute of Technology, Pittsburgh, Pa.

[3] This paper is based on work done for the Office of Scientific Research and Development under contract OEMsr-262 with the Massachusetts Institute of Technology.

[4] R. E. A. C. Paley and N. Wiener, *The Fourier transform in the complex domain*, Am. Math. Soc. Colloquium Publication, 1934, Ch. IV.

E. C. Titchmarsh, *Theory of the Fourier integral*, Oxford University Press, Ch. XI, 1937.

J. S. Schwinger, *The theory of guided waves*, Radiation Laboratory Publication. To be published.

solution. We may thus find the functional form of the various surface current densities as well as the electric field. However, in such problems as we have described above, the physically interesting quantities may be calculated from the far field and these quantities in turn are closely related to the Fourier transform of the surface current densities. Since Eq. (1.1) is solved by transform techniques, these quantities can be obtained immediately.

The problem which we treat here is the following. A plane monochromatic electromagnetic wave whose direction of propagation lies in the plane of the paper, is incident upon an infinite set of staggered, equally spaced, semi-infinite metallic plates of zero thickness and perfect conductivity. These plates extend indefinitely in a direction perpendicular to the plane of the paper. (See Fig. 1 for a side view.) The angle of stagger with respect to a fixed direction (that of the cross section of the plates in Fig. 1) is α, while the direction of propagation with respect to this fixed line is θ, where $\alpha-\pi<\theta<\alpha$ and $0<\alpha\leqq\pi/2$. This structure has some properties which are analo-

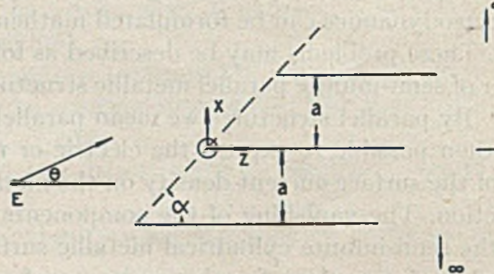

FIG. 1.

gous to those of metal mirrors and gratings. Thus when it is excited by a plane wave with arbitrary direction of propagation, there will be reflected plane waves in certain directions depending on the relative dimensions, the wave length and the direction of incidence.

2. Formulation of the problem. We assume that the electric field of the incident wave has only one component, namely, the component which is perpendicular to the plane of the paper. Since the incident electric field is independent of y and the boundary conditions on the plates must be fulfilled independently of y, no other components of the electric field will be excited. Thus all components of the magnetic field can be derived from this single component of the electric field $E_y(x, z)=\phi(x, z)$. For this case both of the components of the magnetic field lie in the plane of the paper and we shall refer to this problem as an "*H* plane" problem.

If we now write the Maxwell equations[5] in the form

$$\nabla \times \mathrm{E} = ik\mathrm{H}$$

and

$$\nabla \times \mathrm{H} = -ik\mathrm{E},$$

where $k=2\pi/\lambda$, and λ is the free space wave-length, we see immediately that the only components of the magnetic field are

[5] The time dependence of all field quantities is taken to be e^{-ikct} and may therefore be suppressed. c is the velocity of light. In the engineering literature, the time dependence is written as $\exp(ikct)$. In order to convert our final results to standard engineering form, one merely replaces i by $-j$.

$$ikH_x = -\frac{\partial\phi}{\partial z}$$

and

$$ikH_z = \frac{\partial\phi}{\partial x}\cdot$$

Upon eliminating H_x and H_z from the above equations we obtain the two dimensional wave equation,

$$\frac{\partial^2\phi}{\partial x^2}+\frac{\partial^2\phi}{\partial z^2}+k^2\phi = 0$$

which is to be solved subject to the boundary condition, $\phi=0$ on the metal plates since ϕ is the tangential component of the electric field. There are also conditions at infinity on the function $\phi(x, z)$ which we shall discuss later when we have need of them.

We now formulate the equation which expresses the electric field in terms of the surface current density on the metal plates. To this end, we start by modifying the structure in Fig. 1, so that there are now only a finite number of parallel plates, each of which is taken to be finite in length. The length of each plate is such that the amplitudes of the attenuated modes are negligibly small relative to the amplitude of the propagated mode in the parallel plate region before the end of the structure is

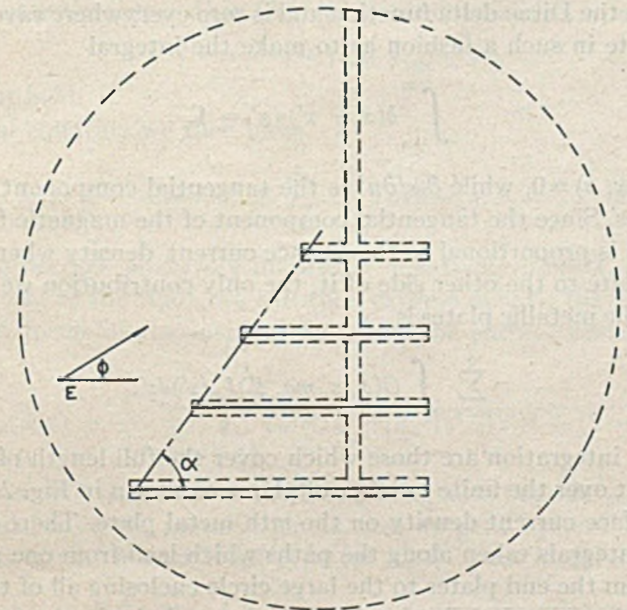

FIG. 2.

reached. (See Fig. 2 for a side view.) If we employ the free space Green's function, we may express $\phi(x, z)$ in terms of $\partial\phi/\partial n$, the normal derivative on the metallic plates. We have from Green's theorem

$$\phi(x, z) = \int_C\left(G\frac{\partial\phi}{\partial n'} - \phi\frac{\partial G}{\partial n'}\right)ds',$$

where the contour C is the one indicated by the dotted line in Fig. 2, ds' is the element of arc length along it and $G(x, z, x', z')$ is the free space Green's function. The outer boundary of the contour C is taken to be a circle of large radius. This is merely for convenience and the outer boundary might have been any other closed curve. $G(x, z, x', z')$ satisfies the homogeneous wave equation

$$\frac{\partial^2 G}{\partial x^2} + \frac{\partial^2 G}{\partial z^2} + k^2 G = 0$$

save for the point $x = x'$, $z = z'$. At this point

$$\int_{-\infty}^{\infty} \frac{\partial G}{\partial x}\Bigg|_{x=x'-0}^{x=x'+0} dz' = -1$$

and

$$\int_{-\infty}^{\infty} \frac{\partial G}{\partial z}\Bigg|_{z=z'-0}^{z=z'+0} dx' = -1.$$

This may be expressed symbolically by saying that $G(x, z, x', z')$ satisfies the inhomogeneous wave equation

$$\frac{\partial^2 G}{\partial x^2} + \frac{\partial^2 G}{\partial z^2} + k^2 G = -\delta(x - x')\delta(z - z'),$$

where $\delta(x - x')$ is the Dirac delta function and is zero everywhere save at $x = x'$, where it becomes infinite in such a fashion as to make the integral

$$\int_{-\infty}^{\infty} \delta(x - x')dx' = 1.$$

On the plates $\phi(x, z) = 0$, while $\partial\phi/\partial n'$ is the tangential component of the magnetic field on the plates. Since the tangential component of the magnetic field suffers a discontinuity which is proportional to the surface current density when we go from one side of a given plate to the other side of it, the only contribution we get from the integration along the metallic plates is

$$\sum_{m=p}^{q} \int G(x, z, ma, z')I_m(z')dz',$$

and the limits of integration are those which cover the full length of each plate. The sum is carried out over the finite number of plates as shown in Fig. 2. $I_m(z)$ is proportional to the surface current density on the mth metal plate. There is complete cancellation of the integrals taken along the paths which lead from one plate to the next or which lead from the end plates to the large circle enclosing all of the plates.

We now calculate the contribution from the large circle. In the first place, the free space Green's function which represents an outgoing wave for $\sqrt{x^2+z^2} \gg \sqrt{x'^2+z'^2}$ is $G(x, z, x', z') = (i/4)H_0^{(1)}[k\sqrt{(x-x')^2+(z-z')^2}]$ where $H_0^{(1)}$ is the Hankel function of the first kind. The contribution from the large circle is

$$\int_0^{2\pi} \left[G(r, r', \beta, \beta')\frac{\partial\phi(r', \beta')}{\partial r'} - \phi(r', \beta')\frac{\partial G(r, r', \beta, \beta')}{\partial r'}\right] r'd\beta', \qquad (2.1)$$

where

$$G(r, r', \beta, \beta') = \frac{i}{4} H_0^{(1)}[k\sqrt{r^2 + r'^2 - 2rr' \cos(\beta - \beta')}]$$

and $x = r \sin \beta$, $z = r \cos \beta$. If we now expand $G(r, r', \beta, \beta')$ in terms of cylindrical waves we have

$$G(r, r', \beta, \beta') = \frac{i}{4} \sum_{m=-\infty}^{\infty} H_m^{(1)}(kr')J_m(kr)e^{im(\beta-\beta')}, \qquad r < r'.$$

Furthermore, for any point outside of the region of the plates

$$\phi(x, z) = e^{ik(z\cos\theta + x \sin\theta)} + \sum_{m=-\infty}^{\infty} \alpha_n H_n^{(1)}(kr)e^{in\beta}, \tag{2.2}$$

where the first term represents the incident plane wave whose direction of propagation is θ, while the second term represents the scattered wave. We shall not be interested in the explicit form of the α_n's and indeed, we shall show that they do not enter explicitly into the formulation of the integral equation. The expression for the plane wave, $e^{ik[z\cos\theta + x\sin\theta]}$ may be expanded in terms of cylindrical waves by noting that

$$e^{ik(z\cos\theta + x\sin\theta)} = e^{ikr\cos(\theta-\beta)} = \sum_{m=-\infty}^{\infty} e^{+im\pi/2}J_m(kr)e^{im(\theta-\beta)}.$$

If we now evaluate the integrals in (2.1) we get immediately

$$e^{ik(z\cos\theta + x\sin\theta)} = \phi_{\text{inc}}(x, z)$$

i.e., the incident field.

For our final equation we then have

$$\phi(x, z) = \phi_{\text{inc}}(x, z) + \sum_{m=p}^{q} \int I_m(z')G(x, z, ma, z')dz'.$$

If we now let q become positively infinite, p negatively infinite, and let each plate extend indefinitely to the right, we can then express $\phi(x, z)$, the y component of the electric field, in terms of the incident field and the surface current density on the plates, that is,

$$\phi(x, z) = \phi_{\text{inc}}(x, z) + \frac{i}{4} \sum_{m=-\infty}^{\infty} \int_{mb}^{\infty} I_m(z')H_0^{(1)}[k\sqrt{(z - z')^2 + (x - ma)^2}]dz', \tag{2.3}$$

where $a = b \tan \alpha$. We now impose the electromagnetic boundary condition, namely that $\phi(x, z)$ vanishes on the metallic plates, and we get a system of simultaneous integral equations of the Wiener-Hopf type for $I_m(z)$. That is, for $x = na$

$$0 = \phi_{\text{inc}}(na, z) + \frac{i}{4} \sum_{m=-\infty}^{\infty} \int_{mb}^{\infty} I_m(z')H_0^{(1)}[k\sqrt{(z - z')^2 + (n - m)^2a^2}]dz' \tag{2.4}$$

for all n with $z > nb$, $n = 0, \pm 1, \pm 2, \cdots$.[5] Due to the periodic nature of the structure, the infinite set of simultaneous integral equations can be cast into the form (1.1).

[5] It is possible to obtain the integral equation (2.3) directly from the infinite structure indicated in Fig. 1. We have intentionally avoided this because it requires a more detailed knowledge of the field at infinity.

We close our discussion of the formulation of the integral equation (2.3) with some remarks about the range of values of a/λ which is allowed. In the parallel plate regions for z large and positive, $\phi(x, z)$ is asymptotic to $\sin(\pi x/a)e^{i\kappa z}$ where $\kappa=\sqrt{k^2-(\pi/a)^2}$. If now $k<\pi/a$, i.e., $\lambda>2a$, κ will be pure imaginary and hence for z sufficiently large and positive, $\phi(x, z)$ will vanish exponentially. In this case, the parallel plate regions cannot sustain a propagating mode. If $k>\pi/a$, i.e., $2a<\lambda$, then κ is real and the parallel plate region can sustain at least one mode consistent with the polarization which we have employed. In order that a second mode not propagate in this parallel plate region, we must further assume that $a<\lambda$. We also assume that there is a single reflected wave. Such a restriction puts further limitations on a/λ as well as on θ. These restrictions will appear when we have obtained the solution of the problem.

3. Fourier transform solution of the integral equation. Before we turn to the Fourier transform solution of the integral equation (2.4) we shall first convert it into one of the Wiener-Hopf type. We note that the surface current density of the mth plate has the same magnitude as that of the zeroth plate provided we measure the distance along the mth plate from its edge. Hence, the surface current density on the mth plate differs from that of the zeroth plate only by a phase factor. This phase factor arises because the amplitude of the incident wave differs from plate edge to plate edge by the factor

$$e^{ik(b\cos\theta+a\sin\theta)}.$$

Thus

$$I_m(z-mb)=I_0(z)e^{ikm(b\cos\theta+a\sin\theta)},$$

where $I_0(z)$ is the surface current density on the zeroth plate. Equation (2.4) may then be rewritten as

$$0=\phi_{\mathrm{inc}}(z, na)+\frac{i}{4}\sum_{m=-\infty}^{\infty}\int_0^{\infty}I_0(z)e^{ik\rho m}H_0^{(1)}[k\sqrt{(z-z'-mb)^2+(n-m)^2a^2}]dz', \quad (3.1)$$

where $\rho=b\cos\theta+a\sin\theta$. If we replace z by $z+nb$, Eq. (3.1) will read

$$0=e^{ik[(z+nb)\cos\theta+na\sin\theta]}$$
$$+\frac{i}{4}\sum_{m=-\infty}^{\infty}\int_0^{\infty}I_0(z')e^{ik\rho m}H_0^{(1)}[k\sqrt{\{(n-m)b+(z-z')\}^2+(n-m)^2a^2}]dz', z>0.$$

Finally, when we divide the last equation by $e^{ik\rho n}$ and put $m-n=q$, we get

$$0=e^{ikz\cos\theta}+\frac{i}{4}\sum_{q=-\infty}^{\infty}\int_0^{\infty}I_0(z')e^{ik\rho q}H_0^{(1)}[k\sqrt{q^2a^2+(qb+z-z')^2}]dz' \quad (3.2)$$

and this equation is of the Wiener-Hopf type.

In order to put this equation into a form which amenable to solution by Fourier transform methods, we extend it for negative z to be

$$\phi_1(z)=\frac{i}{4}\sum_{q=-\infty}^{\infty}\int_0^{\infty}I_0(z')e^{ik\rho q}H_0^{(1)}[k\sqrt{q^2a^2+(qb+z-z')^2}]dz', \quad z<0, \quad (3.3)$$

where $\phi_1(z)$ is an unknown function which is, save for a phase factor, the tangential component of the scattered electric field at $x=na$. In view of the periodic nature of the structure, the dependence of the integral equation on n is not explicit. We may now replace Eqs. (3.2) and (3.3) by the equation

$$\phi_1(z) = \phi_0(z) + \frac{i}{4}\sum_{q=-\infty}^{\infty}\int_{-\infty}^{\infty} I_0(z')e^{ik\rho q}H_0^{(1)}[k\sqrt{q^2a^2+(qb+z-z')^2}]dz', \qquad (3.4)$$

where now

$$\phi_1(z) \equiv 0 \quad \text{for} \quad z > 0,$$
$$I_0(z) \equiv 0 \quad \text{for} \quad z < 0,$$
$$\phi_0(z) \equiv \begin{cases} 0 & \text{for} \quad z < 0, \\ e^{ikz}\cos\theta & \text{for} \quad z > 0. \end{cases}$$

For analytical convenience, it is now assumed that k has a small positive imaginary part. This is tantamount to assuming that the medium is slightly absorbing.

Before we can apply the Fourier transform in the complex plane to the solution of Eq. (3.4) it is necessary to study the growth order of the functions $\phi_1(z)$, $I_0(z)$ and $\phi_0(z)$. It is clear from a direct study of the integral Eqs. (3.2) and (3.3) that these functions are integrable for all finite z. The half planes of regularity of the Fourier transforms of $\phi_0(z)$, $\phi_1(z)$ and $I_0(z)$ are, of course, determined from their growth orders at infinity and we now proceed to determine these orders. Since we know $\phi_0(z)$ explicitly, it is clear that its Fourier transform is

$$\int_0^{\infty} e^{-iwz}\phi_0(z')dz' = \frac{1}{i[w - k\cos\theta]}$$

and is regular in a lower half of the w plane defined by the inequality $\mathfrak{Im}w < \mathfrak{Im}(k\cos\theta)$. Save for a translation on the z variable and a phase factor which is independent of z, $I_0(z)$ is, in certain units, the surface current density on any metallic plate. For z sufficiently large and positive, $I_0(z)$ is asymptotic to the surface current density in any of the parallel plate regions, that is, it is asymptotic to $e^{i\kappa z}$. Since $I_0(z)$ is integrable at the origin, the Fourier transform of $I_0(z)$, that is

$$\int_0^{\infty} I_0(z')e^{-iwz'}dz',$$

is regular in some half plane defined by

$$\mathfrak{Im}w < \mathfrak{Im}(\kappa) \sim \frac{\mathfrak{Re}k\mathfrak{Im}k}{|\kappa|} > \mathfrak{Im}k,$$

since $\mathfrak{Re}(k)/|\kappa| > 1$.

We now investigate the asymptotic form of $\phi_1(z)$ for z large and negative. Before doing this, however, it is convenient to give another representation of the kernel of the integral equation (3.4). The kernel

$$\frac{i}{4}\sum_{q=-\infty}^{\infty} e^{ik\rho q}H_0^{(1)}[k\sqrt{q^2a^2+(qb+z)^2}]$$

has the Fourier integral representation

$$\frac{i}{4\pi}\int_C e^{iwz}\sum_{q=-\infty}^{\infty}\frac{e^{ik\rho q+i|q|a\sqrt{k^2-w^2}-iwqb}}{\sqrt{k^2-w^2}}\,dw, \tag{3.5}$$

where C is a contour which lies in the strip of regularity of the sum in (3.5). It is closed in the upper or lower half planes by a large semi-circle which passes between the poles of this sum depending upon whether $z>0$ or $z<0$. The strip of regularity is, of course, determined by the region in which the infinite series in (3.5) converges. A direct study of this series will reveal that the ordinates of convergence are given by the inequality, $\mathfrak{Im}k\cos(2\alpha-\theta)<\mathfrak{Im}w<\mathfrak{Im}k\cos\theta$. This now clarifies the reason why we imposed a small but positive imaginary part on k. Had we not done this, the series would only converge on the real axis of the w plane and as we shall see in the actual solution of the Wiener-Hopf equation, this situation would have presented us with some analytical difficulties.

We may now write the sum in the integral (3.5) in closed form as

$$\frac{i}{4\pi}\int_C\frac{e^{iwz}\sin a\sqrt{k^2-w^2}\,dw}{\sqrt{k^2-w^2}[\cos a\sqrt{k^2-w^2}-\cos(k\rho-wb)]}.$$

For $z<0$, we close the path C in the lower half of the w plane. The poles in the lower half plane are $w=k\cos(2\alpha-\theta)$ and two infinite sequences of poles both of which have negative imaginary parts. We shall have more to say about this double set of poles presently. Suffice it to be noted at this point, that the kernel has a second representation which for $z<0$ may now be written as

$$\frac{e^{ikz\cos(2\alpha-\theta)}}{2ak\sin(\alpha-\theta)}+\text{terms which attenuate exponentially for } z \text{ large and negative.}$$

It is clear then, that for z large and negative, $\phi_1(z)$ is asymptotic to

$$\int_0^\infty\frac{e^{ik(z-z')\cos(2\alpha-\theta)}}{2ak\sin(\alpha-\theta)}I_0(z')dz',$$

and thus, the Fourier transform of $\phi_1(z)$, i.e.,

$$\int_{-\infty}^0 e^{-iwz}\phi_1(z)dz,$$

is regular in the upper half of the w plane $\mathfrak{Im}w>\mathfrak{Im}k\cos(2\alpha-\theta)$.

The Fourier transforms involved in this problem then have a common strip of regularity, $\mathfrak{Im}((k\cos(\theta-2\alpha))<\mathfrak{Im}w<\mathfrak{Im}(k\cos\theta)$ and it is thus permissible to apply the Fourier transform to the integral equation (3.4) within this strip.

Let $\Phi_1(w)$ be the Fourier transform of $\phi_1(z)$ and $J(w)$ the Fourier transform of $I_0(z)$. The Fourier transform of the integral equation (3.4) is then

$$\Phi_1(w)=\frac{1}{i(w-k\cos\theta)}+\frac{J(w)\sin a\sqrt{k^2-w^2}}{2\sqrt{k^2-w^2}\,[\cos a\sqrt{k^2-w^2}-\cos(k\rho-wb)]}. \tag{3.6}$$

The Wiener-Hopf theory now tells us that we can split this transform equation into two parts. One part will be regular in an upper half plane, $\mathfrak{J}m w > \mathfrak{J}m k \cos(\theta - 2\alpha)$, the other in a lower half plane $\mathfrak{J}m w < \mathfrak{J}m k \cos\theta$ and both of these half planes have a common region of regularity. It is well to note here that we use the term regularity in a slightly extended sense. We imply by regularity that the function has neither zeros, branch points nor poles in the region of regularity. That is, the function as well as its reciprocal is "regular" in the conventional sense of the term. Suppose we assume that we can write

$$\frac{K_-(w)}{K_+(w)} = \frac{\sin a\sqrt{k^2 - w^2}}{\sqrt{k^2 - w^2}\,[\cos a\sqrt{k^2 - w^2} - \cos(k\rho - wb)]},$$

where $K_-(w)$ is regular in the proper lower half plane and $K_+(w)$ is regular in the proper upper half plane and that there is a common strip of regularity for both $K_-(w)$ and $K_+(w)$. Then

$$\Phi_1(w)K_+(w) = \frac{K_+(w)}{i(w - k\cos\theta)} + \frac{J(w)K_-(w)}{2}\cdot \tag{3.7}$$

The left side of Eq. (3.7) is regular in an upper half plane while the second term on the right side is regular in a lower half plane. The term

$$\frac{K_+(w)}{i(w - k\cos\theta)}$$

is only regular in the strip of regularity. This function may be decomposed into two functions in such a manner that one function is regular in the appropriate upper and the other in the appropriate lower half plane, since

$$\frac{K_+(w)}{i(w - k\cos\theta)} = \frac{K_+(w) - K_+(k\cos\theta)}{i(w - k\cos\theta)} + \frac{K_+(k\cos\theta)}{i(w - k\cos\theta)}\cdot$$

The first term on the right no longer has a singularity at $w = k\cos\theta$, but is regular in the upper half plane and the second term is regular in the lower half plane. Thus Eq. (3.7) can be rewritten in the form

$$\Phi_1(w)K_+(w) - \frac{K_+(w) - K_+(k\cos\theta)}{i(w - k\cos\theta)} = \frac{J(w)K_-(w)}{2} + \frac{K_+(k\cos\theta)}{i(w - k\cos\theta)}\cdot \tag{3.8}$$

The right side of the equation is regular in the lower half plane $\mathfrak{J}m w < \mathfrak{J}m k\cos\theta$ while the left side is regular in the upper half plane $\mathfrak{J}m w > \mathfrak{J}m k\cos(\theta - 2\alpha)$. Both sides have a common strip of regularity and hence the left side of (3.8) is the analytical continuation of the right side. Such an equality can only hold if both sides of Eq. (3.8) are equal to an integral function, that is, a function regular everywhere in the complex w plane. We have then

$$\frac{J(w)K_-(w)}{2} + \frac{K_+(k\cos\theta)}{i(w - k\cos\theta)} = \text{integral function} \tag{3.9}$$

and also

$$\Phi_1(w)K_+(w) - \frac{K_+(w) - K_+(k\cos\theta)}{i(w - k\cos\theta)} = \text{integral function.} \tag{3.10}$$

We shall now show that it is possible to decompose the function

$$\frac{\sin a\sqrt{k^2 - w^2}}{\sqrt{k^2 - w^2}\,[\cos a\sqrt{k^2 - w^2} - \cos (k\rho - wb)]}$$

into two functions, one of which is regular in the lower half plane $\mathfrak{Im}\,w < \mathfrak{Im}\,k \cos \theta$, while the other is regular in the upper half plane $\mathfrak{Im}\,w > \mathfrak{Im}\,k \cos (\theta - 2\alpha)$. The denominator of the fraction may be written as

$$\begin{aligned}
&\cos a\sqrt{k^2 - w^2} - \cos (k\rho - wb)\\
&= 2 \sin \frac{[a\sqrt{k^2 - w^2} + k\rho - wb]}{2} \sin \frac{[k\rho - wb - a\sqrt{k^2 - w^2}]}{2}\\
&= \tfrac{1}{2}[a\sqrt{k^2 - w^2} + k\rho - wb][k\rho - wb - a\sqrt{k^2 - w^2}]\\
&\quad\times \prod_{n=1}^{\infty}\left[1 - \frac{(a\sqrt{k^2 - w^2} + k\rho - wb)^2}{4n^2\pi^2}\right]\prod_{n=1}^{\infty}\left[1 - \frac{(k\rho - wb - a\sqrt{k^2 - w^2})^2}{4n^2\pi^2}\right]\\
&= \tfrac{1}{2}[(k\rho - wb)^2 - a^2(k^2 - w^2)]\prod_{n=1}^{\infty}\left[1 - \frac{a\sqrt{k^2 - w^2} + k\rho - wb}{2n\pi}\right]e^{(a\sqrt{k^2-w^2}+k\rho-wb)/2n\pi}\\
&\quad\times \prod_{n=-\infty}^{-1}\left[1 - \frac{a\sqrt{k^2 - w^2} + k\rho - wb}{2n\pi}\right]e^{(a\sqrt{k^2-w^2}+k\rho-wb)/2n\pi}\\
&\quad\times \prod_{n=1}^{\infty}\left[1 - \frac{k\rho - wb - a\sqrt{k^2 - w^2}}{2n\pi}\right]e^{(k\rho-wb-a\sqrt{k^2-w^2})/2n\pi}\\
&\quad\times \prod_{n=-\infty}^{-1}\left[1 - \frac{k\rho - wb - a\sqrt{k^2 - w^2}}{2n\pi}\right]e^{(k\rho-wb-a\sqrt{k^2-w^2})/2n\pi}.
\end{aligned}$$

The exponential factors in each of these products has been inserted to render the products absolutely convergent. The above expression may now be rewritten to read

$$\tfrac{1}{2}(a^2 + b^2)(w - \sigma_1)(w - \sigma_2) \prod_{n=-\infty}^{\infty}{}' \left[\left\{1 - \frac{k\rho - wb}{2n\pi}\right\}^2 - \frac{a^2(k^2 - w^2)}{4n^2\pi^2}\right]e^{(k\rho-wb)/n\pi} \qquad (3.11)$$

where the prime on the products denotes the absence of the term $n = 0$ in the product. The infinite product in the last expression may now be expressed in a manner such that it puts into evidence the portion which is regular in the correct upper half and lower half planes. Indeed we may express (3.11) as

$$\begin{aligned}
&\tfrac{1}{2}(a^2 + b^2)(w - \sigma_1)(w - \sigma_2) \prod_{n=-\infty}^{\infty}{}' [\Delta_n - i\Psi_n]e^{[(k\rho-wb+wai)/2\pi n]+i(\pi/2-\alpha)}\\
&\qquad\qquad\times \prod_{n=-\infty}^{\infty}{}' [\Delta_n + i\Psi_n)e^{[(k\rho-wb-wai)/2n\pi]-i(\pi/2-\alpha)},
\end{aligned}$$

where now

$$\sigma_1 = k \cos \theta, \qquad \sigma_2 = k \cos (2\alpha - \theta),$$

and

$$\Delta_n = \sqrt{\sin^2 \alpha\left(1 - \frac{k\rho}{2\pi n}\right)^2 - \left(\frac{ak}{2\pi n}\right)^2}, \qquad \Psi_n = \cos \alpha\left(1 - \frac{k\rho}{2\pi n}\right) + \frac{wa \csc \alpha}{2\pi n},$$

where again, the exponential factors following the infinite products have been chosen to insure the absolute convergence of the product. One should note at this point that the choice of these exponential factors is not unique and indeed need only be asymptotic to the factors which we have chosen. However, we shall see that a second integral function $\chi(w)$, introduced into the decomposition of $K(w)$, is determined in terms of the factors which we have chosen. We have finally that the factor

$$(w-\sigma_1)\prod_{n=-\infty}^{-1}[\Delta_n - i\Psi_n]e^{[(k\rho-wb+wai)/2\pi n]+i(\pi/2-\alpha)}\prod_{n=1}^{\infty}[\Delta_n + i\Psi_n]e^{[(k\rho-wb-wai)/2\pi n]-i(\pi/2-\alpha)}$$

has no zeros in the lower half plane $\mathfrak{Im}\,w<\mathfrak{Im}\,k\cos\theta$, while the factor

$$(w-\sigma_2)\prod_{n=1}^{\infty}[\Delta_n - i\Psi_n]e^{[(k\rho-wb-wai)/2\pi n]+i(\pi/2-\alpha)}\prod_{n=-\infty}^{-1}[\Delta_n + i\Psi_n]e^{[(k\rho-wb-wai)/2\pi n]-i(\pi/2-\alpha)}$$

has no zeros in the upper half plane $\mathfrak{Im}\,w>\mathfrak{Im}[k\cos(\theta-2\alpha)]$. The factorization of

$$\frac{\sin a\sqrt{k^2-w^2}}{\sqrt{k^2-w^2}}$$

is more direct, for

$$\frac{\sin a\sqrt{k^2-w^2}}{\sqrt{k^2-w^2}} = a\prod_{n=1}^{\infty}\left[1-\frac{a^2(k^2-w^2)}{n^2\pi^2}\right]$$

$$= \frac{a^3}{\pi}(w-\kappa)e^{-iaw/\pi}(w+\kappa)e^{iaw/\pi}\prod_{n=2}^{\infty}\left[\sqrt{1-\left(\frac{ak}{\pi n}\right)^2}+\frac{iaw}{\pi n}\right]e^{-iaw/\pi n}$$

$$\times\prod_{n=2}^{\infty}\left[\sqrt{1-\left(\frac{ak}{\pi n}\right)^2}-\frac{iaw}{\pi n}\right]e^{iaw/\pi n}.$$

The factor

$$\frac{a}{\pi}(w-\kappa)e^{-iaw/\pi}\prod_{n=2}^{\infty}\left[\sqrt{1-\left(\frac{ak}{\pi n}\right)^2}+\frac{iaw}{\pi n}\right]e^{-iaw/\pi n}$$

has no zeros in the lower half plane $\mathfrak{Im}\,w<\mathfrak{Im}\,\kappa$, while the factor

$$\frac{a}{\pi}(w+\kappa)e^{iaw/\pi}\prod_{n=2}^{\infty}\left[\sqrt{1-\left(\frac{ak}{\pi n}\right)^2}-\frac{iaw}{\pi n}\right]e^{iaw/\pi n}$$

has no zeros in the upper half plane $\mathfrak{Im}\,w>\mathfrak{Im}(-\kappa)$. We thus find that

$$K_-(w) = \frac{\prod_{n=2}^{\infty}\left[\sqrt{1-\left(\frac{ak}{\pi n}\right)^2}+\frac{iaw}{\pi n}\right]e^{-iaw/\pi n}\frac{a}{\pi}(w-\kappa)e^{-iaw/\pi}e^{\chi(w)}}{(w-\sigma_1)\prod_{n=-\infty}^{-1}[\Delta_n - i\Psi_n)e^{[(k\rho-wb+wai)/2\pi n]+i(\pi/2-\alpha)}\prod_{n=1}^{\infty}[\Delta_n + i\Psi_n]e^{[(k\rho-wb-wai)/2\pi n]-i(\pi/2-\alpha)}}$$

is free of zeros and poles in the lower half plane $\mathfrak{Im}\,w<\mathfrak{Im}\,k\cos\theta$. The factor $e^{\chi(w)}$ will be determined so as to make $K_-(w)$ have algebraic growth as $|w|\to\infty$ for $\mathfrak{Im}\,w<0$. With $\chi(w)$ so chosen, the integral function sought can only be of algebraic growth for $|w|\to\infty$. $K_-(w)$ is regular in the lower half plane $\mathfrak{Im}\,w<\mathfrak{Im}\,k\cos\theta$. Finally,

$$K_+(w) = \frac{(a^2+b^2)(w-\sigma_2)e^{\chi(w)}\prod_{n=1}^{\infty}[\Delta_n - i\Psi_n]e^{[(k\rho-wb+wai)/2\pi n]+i(\pi/2-\alpha)}\prod_{n=-\infty}^{-1}[\Delta_n + i\Psi_n]e^{[(k\rho-wb+wai)/2\pi n]-i(\pi/2-\alpha)}}{2a\prod_{n=2}^{\infty}\left[\sqrt{1-\left(\frac{ak}{\pi n}\right)^2}-\frac{iaw}{\pi n}\right]e^{iaw/\pi n}\frac{a}{\pi}(w+\kappa)e^{iaw/\pi}}$$

has no zeros or poles in the upper half plane $\Im m w > \Im m k \cos(\theta - 2\alpha)$.

We shall now discuss the asymptotic form of $K_-(w)$ as $|w| \to \infty$, $\Im m w < 0$. This procedure will enable us to determine the unknown integral function $\chi(w)$. It has been shown by Schwinger[7] that functions of the form of $K_-(w)$ are independent of ka for $|w| \to \infty$, $\Im m w < 0$, and $\pi < ak < 2\pi$. Thus

$$K_-(w) = \frac{e^{\chi(w)} \prod_{n=2}^{\infty} \left[1 + \frac{iaw}{\pi n}\right] e^{-iaw/\pi n} \frac{a}{\pi} e^{-iaw/\pi}}{\prod_{n=-\infty}^{-1} \left[\sin\alpha - i\left(\cos\alpha + \frac{wa\csc\alpha}{2\pi n}\right)\right] e^{[(wai-wb)/2\pi n]+i(\pi/2-\alpha)}} \times \frac{1}{\prod_{n=1}^{\infty} \left[\sin\alpha + i\left(\cos\alpha + \frac{wa\csc\alpha}{2\pi n}\right)\right] e^{-[(wai+wb[/2\pi n]-i(\pi/2-\alpha)}}. \tag{3.12}$$

The products in (3.12) are now in the form of gamma functions and

$$K_-(w) \sim \frac{e^{\chi(w)} \left(\frac{wa\csc\alpha}{2\pi}\right)^2 e^{iwa\gamma/\pi} \Gamma\left(\frac{-wa\csc\alpha e^{-i\alpha}}{2\pi}\right) \Gamma\left(\frac{wa\csc\alpha e^{i\alpha}}{2\pi}\right)}{\frac{iaw}{\pi}\left(1 + \frac{iaw}{\pi}\right) e^{iaw\gamma/\pi} \Gamma\left(\frac{iaw}{\pi}\right)},$$

where γ is the Euler-Mascheroni constant. Using the Stirling expansion theorem for $|w| \to \infty$, $\Im m w < 0$ we get

$$K_-(w) \sim \frac{e^{\chi(w)} a \csc^2\alpha \left(-\frac{aw\csc\alpha}{2\pi} e^{-i\alpha}\right)^{-[(wa\csc\alpha)/2\pi]e^{-i\alpha}-1/2} \left(\frac{aw\csc\alpha}{2\pi} e^{i\alpha}\right) e^{[(wa\csc\alpha)/2\pi]e^{i\alpha}-1/2}}{4\pi^2 i \left(\frac{iaw}{\pi}\right)^{(iaw/\pi)-1/2}}$$

$$\sim \frac{Ce^{\chi(w)+iaw/\pi[(\alpha-\pi/2)\cot\alpha+\ln(\csc\alpha)/2]}}{w^{1/2}},$$

where C is a constant. Thus if we choose

$$\chi(w) = \frac{-iaw}{\pi}\left[\left(\alpha - \frac{\pi}{2}\right)\cot\alpha - \ln 2\sin\alpha\right],$$

$K_-(w)$ will have algebraic growth for $|w|$ large, $\Im m w < 0$.

Now $J(w)$, which is proportional to the Fourier transform of the surface current density on the various plates, approaches zero for $|w|$ large, $\Im m w < 0$. This assumes, of course, that $I_0(z)$ can at most be of exponential growth for z large and positive and is integrable for z finite. Thus $K_-(w)J(w)$ approaches zero for $|w|$ large and $\Im m w < 0$. If we now return to Eq. (3.12) we see that as $|w|$ becomes large, $\Im m w < 0$, the integral function in (3.9) is asymptotic to zero. We may now apply the same argument to Eq. (3.10) and find that the integral function is again asymptotic to zero. But by a theorem of Liouville, and analytic function which is bounded in the entire complex plane is constant and in this case the constant must be zero. We thus have

$$J(w) = \frac{2iK_+(k\cos\theta)}{K_-(w)(w - k\cos\theta)}.$$

If we were interested in the explicit form of the surface current density, we could obtain it from $J(w)$ by evaluating the Fourier inversion integral

[7] J. S. Schwinger, loc. cit.

$$\frac{i}{\pi}\int_C \frac{K_+(k\cos\theta)e^{iwz}dw}{K_-(w)(w-k\cos\theta)},$$

where C is a contour which may be taken as a straight line within the strip of regularity of the Fourier transforms of $I(z)$, $\phi_1(z)$, $\phi_0(z)$ and $K(z)$. The contour is closed above by a semi-circle, which by familiar arguments in contour integration may be shown to make no contribution to the value of the integral. In the next section we shall show that it is possible to find the reflection and transmission coefficients without evaluating this integral in detail.

4. Investigation of the far fields. In order to find the reflection and transmission coefficients, we now investigate the asymptotic form of $\phi(x, z)$ for $|z|$ large. To this end we note that Eq. (2.3) can be written in Fourier integral representation as

$$\phi(x, z) = \phi_{\text{inc}}(x, z) + \frac{i}{4\pi}\int_C \sum_{m=-\infty}^{\infty} \frac{e^{iwz+ikm\rho-iwmb+|x-ma|\sqrt{k^2-w^2}}J(w)dw}{\sqrt{k^2-w^2}},$$

where C is the contour which we described at the end of Section 3. This in turn, may simplified to

$$\phi(x, z) = \phi_{\text{inc}}(x, z)$$
$$-\frac{i}{4\pi}\int_C J(w)e^{iwz+i(k\rho-wb)}\frac{[\sin\sqrt{k^2-w^2}(x-an-a)+e^{i(k\rho-wb)}\sin\sqrt{k^2-w^2}(an-x]dw}{\sqrt{k^2-w^2}\,[\cos a\sqrt{k^2-w^2}-\cos(k\rho-wb)]}, \quad (4.1)$$

where n is the greatest integer contained in x/a. From (4.1) one can get the asymptotic form of $\phi(x, z)$ as z becomes large and positive. Since $J(w)$ is regular in the lower half of the w plane $\Im mw<\Im mk\cos\theta$, we can close the contour C by a large semicircle which passes between the poles in the upper half plane. For $na<x<(n+1)a$ it can be seen that due to the form of the integrand, there is no contribution from this circular arc as its radius becomes infinite. In the upper half plane $\Im mw>\Im mk\cos(2\alpha-\theta)$, there are two poles which correspond to propagating modes, namely $w=k\cos\theta$ and $w=\kappa$. All other modes are attenuated modes in the sense that they have large positive imaginary parts compared to the imaginary parts of $k\cos\theta$ and κ. If we now express $J(w)$ as a function of w and use the above described contour in the evaluation of the integral in (4.1) we have then to consider the asymptotic form of

$$\frac{i}{2\pi}\int_C e^{i[k\rho-wb]n+iwz}\frac{[\sin(x-an-a)\sqrt{k^2-w^2}+e^{i(k\rho-wb)}\sin\sqrt{k^2-w^2}(an-x)]K_+(k\cos\theta)dw}{(w-k\cos\theta)K_+(w)\sin a\sqrt{k^2-w^2}}.$$

This in turn is equal to $[e^{ik(x\sin\theta+z\cos\theta)}-Te^{i\kappa z}\sin\pi x/a$ +terms which approach zero for $z\gg 0]$. For z large and positive, this is asymptotic to

$$\phi_{\text{inc}}(x, z) - Te^{i\kappa z}\sin\frac{\pi x}{a}\cdot$$

Hence, save for a numerical factor, the functional form of $\phi(x, z)$ as z becomes infinite is $e^{i\kappa z}\sin\pi x/a$, that is, it represents a travelling wave in the parallel plate region with propagation constant κ, as it should. The amplitude of this wave is

$$T = |T|e^{i\Theta} = \frac{\pi e^{in(k\rho-\kappa b)}(-)^n[1+e^{i(k\rho-\kappa b)}]K_+(k\cos\theta)}{(\kappa-k\cos\theta)a^2\kappa K_+(\kappa)}$$

and depends of course on the particular parallel plate region for which it has been computed. Since T is the amplitude of the wave transmitted in the parallel plate region it is the transmission coefficient because the amplitude of the incident wave has been taken to be unity. If we now assume that k is real, the magnitude of T is

$$| T | = \frac{2^{3/2} k \sin (\alpha - \theta)}{\sqrt{(k \cos \theta + \kappa)(\kappa - k \cos (2\alpha - \theta)}},$$

a quantity independent of the particular parallel plate region considered. Its phase angle depends, of course, on the particular parallel plate region. We shall not give the phase angle explicitly since we shall not use it in our later discussions.

For z large and negative we close the contour in the lower half of the w plane. There is again no contribution from the circular arc which is drawn between the poles in the lower half plane and so we need only evaluate the residues from the poles in the lower half plane. The dominant contribution now arises from the pole $w = k \cos (\theta - 2\alpha)$ and in this case the dominant term is

$$\frac{K_+(k \cos \theta) e^{ik[x \sin (2\alpha - \theta) + z \cos (2\alpha - \theta)]}}{k[\cos (2\alpha - \theta) - \cos \theta] K_+' [k \cos (2\alpha - \theta)]},$$

all other terms in the integrand approaching zero for z large and negative. Here $K_+' [k \cos (2\alpha - \theta)]$ means, as usual, the derivative of $K_+(w)$ with respect to w evaluated at $w = k \cos (2\alpha - \theta)$. The amplitude of the reflected plane wave is the reflection coefficient R if the amplitude of the incident wave is taken as unity, so that we now have

$$R = \frac{K_+(k \cos \theta)}{k[\cos (2\alpha - \theta) - \cos \theta] K_+' [k \cos (2\alpha - \theta)]}.$$

Assuming, once again that k is real, the reflection coefficient may then be rewritten in complex polar form as follows:

$$R = - e^{i(\Theta_1 - \Theta_2)} \sqrt{\frac{(k \cos \theta - \kappa)(k \cos (2\alpha - \theta) + \kappa)}{(k \cos \theta + \kappa)(k \cos (2\alpha - \theta) - \kappa)}},$$

where now

$$\Theta_1 = - \sum_{n=1}^{\infty} \left\{ \arc\sin \frac{\cos \alpha + \dfrac{ka}{2\pi n} \sin (\alpha - \theta)}{\sqrt{1 - \dfrac{ka}{\pi n} \sin \theta}} - \frac{ka \cos \theta}{2\pi n} - \left(\frac{\pi}{2} - \alpha \right) \right\}$$

$$+ \sum_{n=-\infty}^{-1} \left\{ \arc\sin \frac{\cos \alpha + \dfrac{ka}{2\pi n} \sin (\alpha - \theta)}{\sqrt{1 - \dfrac{ka}{\pi n} \sin \theta}} - \frac{ka \cos \theta}{2\pi n} - \left(\frac{\pi}{2} - \alpha \right) \right\}$$

$$+\sum_{n=2}^{\infty}\left\{\arc\sin\frac{ka\cos\theta}{\pi n\sqrt{1-\left(\frac{ka}{\pi n}\right)^2\sin^2\theta}}-\frac{ka}{\pi n}\cos\theta\right\}$$

$$-\frac{ak}{\pi}\cos\theta\left\{1+\left(\alpha-\frac{\pi}{2}\right)\cos\alpha-\ln 2\sin\alpha\right\},$$

and

$$\Theta_2=-\sum_{n=1}^{\infty}\left\{\arc\sin\frac{\cos\alpha+\frac{ka}{2\pi n}\sin(\theta-\alpha)}{\sqrt{1-\frac{ka}{\pi n}\sin(2\alpha-\theta)}}-\frac{ka}{2n\pi}\cos(\theta-2\alpha)-\left(\frac{\pi}{2}-\alpha\right)\right\}$$

$$+\sum_{n=-\infty}^{-1}\left\{\arc\sin\frac{\cos\alpha+\frac{ka}{2\pi n}\sin(\theta-\alpha)}{\sqrt{1-\frac{ka}{\pi n}\sin(2\alpha-\theta)}}-\frac{ka}{2n\pi}\cos(\theta-2\alpha)-\left(\frac{\pi}{2}-\alpha\right)\right\}$$

$$+\sum_{n=2}^{\infty}\left\{\arc\sin\frac{\frac{ka}{\pi n}\cos(2\alpha-\theta)}{\sqrt{1-\left(\frac{ka}{\pi n}\right)^2\sin^2(2\alpha-\theta)}}-\frac{ak}{n\pi}(\cos(2\alpha-\theta)\right\}$$

$$-\frac{ak}{\pi}\cos(2\alpha-\theta)\left\{1+\left(\alpha-\frac{\pi}{2}\right)\cos\alpha-\ln 2\sin\alpha\right\}.$$

It is evident that there will be restrictions on a/λ and θ if we are to have a single reflected plane wave. These restrictions become evident when we study the arc sin sums and observe that conceivably the first term in the sums beginning with index unity can exceed unity. We have tacitly assumed that they do not, for otherwise they would appear in the amplitude factor as real terms. Thus we must see what is implied by the condition that all factors in the infinite products be complex, or equivalently $\Delta_n^2>0$. If we demand that $\Delta_1^2>0$ it is clear that all other $\Delta_n s$, $n=1, 2, \cdots$ will also be >0. The condition $\Delta_1>0$ is equivalent to

(i) $$\frac{ak}{\pi}=\frac{2a}{\lambda}<\frac{\sin\alpha}{\cos^2\frac{1}{2}(\theta-\alpha)}$$

and

(ii) $$\frac{ak}{\pi}>-\frac{\sin\alpha}{\text{sn}^2\frac{1}{2}(\theta-\alpha)}.$$

Condition (ii) is always satisfied since a/λ is always positive. Condition (i) can be more restrictive than the condition $1/2<a/\lambda<1$. For example, if $\theta=\pi/12$, $\alpha=5\pi/12$, then condition (i) implies

$$a/\lambda<.65.$$

5. No propagation in the parallel plate regions. In the Fourier transform solution of the integral equation (3.4) we have assumed that there was only one propagating mode in the parallel plate region, i.e.,

$$I_m(z) \sim e^{i\kappa z} \sin \frac{\pi x}{a}, \qquad \kappa > 0$$

for z large and positive and z in the parallel plate region. Suppose now we dispense with this assumption and ask what form the reflection coefficient takes if we now assume that $\kappa^2 < 0$, i.e., $0 < a/\lambda < 1/2$. In this case κ is, of course, imaginary and

$$I_m(z) \sim e^{-\sqrt{(\pi^2/a^2)-k^2}\, z} \sin \frac{\pi x}{a}$$

for z large and positive and z in the parallel plate region. The result we desire can be obtained most easily by studying the result which we have obtained in Section 3.

We note that if κ is purely imaginary and k is real, the amplitude of the reflection coefficient becomes complex of magnitude unity. Indeed for $\kappa^2 < 0$

$$\sqrt{\frac{(k \cos \theta - \kappa)(k \cos (2\alpha - \theta) + \kappa)}{(k \cos \theta + \kappa)(k \cos (2\alpha - \theta) - \kappa)}}$$

$$= \exp\left\{ i \left[\arcsin \frac{ka \cos \theta}{\pi \sqrt{1 - \left(\frac{ka}{\pi}\right)^2 \sin^2 \theta}} - \arcsin \frac{ka \cos (2\alpha - \theta)}{\pi \sqrt{1 - \left(\frac{ka}{\pi}\right)^2 \sin^2 (2\alpha - \theta)}} \right] \right\}.$$

Thus for this situation, the amplitude of the reflection coefficient is -1. The phase angle Θ_1' is given by

$$\Theta_1' = \Theta_1 + \arcsin \frac{ka \cos \theta}{\pi \sqrt{1 - \left(\frac{ka}{\pi}\right)^2 \sin^2 \theta}},$$

while Θ_2' is now given by

$$\Theta_2' = \Theta_2 + \arcsin \frac{ka}{\pi} \frac{\cos (2\alpha - \theta)}{\sqrt{1 - \left(\frac{ak}{\pi}\right)^2 \sin^2 (2\alpha - \theta)}}.$$

Hence the reflection coefficient for $0 < a/\lambda < 1/2$ is now $-e^{i(\Theta_1' - \Theta_2')}$. For a single reflected wave, the inequality (i) in Section 4 must still be satisfied, although now it is not as severe.

6. Discussion of results. It should be pointed out that some of the results obtained from our calculations can be interpreted in a simple physical manner. For convenience, in this discussion, instead of the angle θ we use the angle i, which the incident wave makes with the normal to the trace of the edges of the plates. It is readily verified that

$$i = \theta - \alpha + \frac{\pi}{2}$$

and also that the angle r which the reflected wave makes with the normal is also equal to i. The condition that there be only one reflected wave

$$\frac{2a}{\lambda} < \frac{\sin \alpha}{\cos^2 \frac{1}{2}(\theta - \alpha)}$$

is seen to be a result of simple grating theory. If the waves scattered by a uniform grating are not to interfere constructively in the region from which the waves are incident (except for the specular case $r = i$) the condition $a'(1 + \sin i) < \lambda$ must be satisfied where a' is the distance between neighboring scatterers. In our case $a' = a \csc \alpha$. Expressing θ in terms of i, the relation

$$\frac{2a}{\lambda} < \frac{\sin \alpha}{\cos^2 \frac{1}{2}(\theta - \alpha)}$$

is seen to be equivalent to $a \csc \alpha(1 + \sin i) < \lambda$. If this condition is satisfied and the condition for no propagation, $\lambda > 2a$, is also satisfied, the plates act as a perfect plane mirror. However, while the magnitude of the reflected wave is unity, its phase is not π but $\Theta_1' - \Theta_2'$. It is easily shown that it will be π on any plane parallel to that of the trace at a distance d given by

$$(4\pi d/\lambda) \cos i + 2m\pi = \Theta_2' - \Theta_1' \qquad m = 0, \pm 1, \pm 2, \cdots .$$

Therefore, as far as all distant fields are concerned, the plates behave in this case like a perfect plane mirror whose surface coincides with any of the planes given by the above equation.

When transmission is possible in the parallel plate region the wavelength in this region differs from that in free space. One would, therefore, expect to find some analogy with the phenomena associated with a plane interface between two dielectric media. This can be shown for the case $\alpha = \pi/2$. In this case the magnitude of the reflection coefficient is

$$|R| = \frac{k \cos i - \kappa}{k \cos i + \kappa}.$$

This expression is identical with that obtained for the reflection at a dielectric interface of a wave with the electric vector parallel to the interface. The phases are different in the two cases and one can again find a set of planes at a distance d from the trace given by

$$(4\pi d/\lambda) \cos i + 2m\pi = \Theta_2 - \Theta_1, \qquad m = 0, \pm 1, \cdots$$

such that the distant fields are identical in the two cases if we regard any one of the planes as the interface.

The expression for the magnitude of the reflection coefficient should be of use in estimating the reflection of waves incident on a metal lens provided that the radius of curvature of the lens (i.e., the angle α) does not vary too rapidly.

A PROBLEM IN THE PROPAGATION OF SHOCK*

BY

MONROE H. MARTIN

University of Maryland

Introduction. This paper deals with a single problem in the rectilinear motion of a gas, namely, *what is the subsequent behavior of a gas initially at rest if its initial density is a constant* ρ_0 *in the region* $|x|<1$ *and a constant* $\rho_2<\rho_0$ *in the region* $|x|>1$?

The behavior of the gas is an idealization of the behavior of the atmosphere in an infinitely long right circular cylinder after an explosion within the cylinder.

It is assumed that the pressure p and density ρ of the gas are related by the isentropic law $p=k^2\rho^\gamma$ where k^2 is constant for all x and all t. Under the law of conservation of energy[1] (Rankine-Hugoniot equation) there is a change[2] in entropy across a shock and the results in the paper may be regarded as an approximation to the actual state of affairs only in the case where the change in density across the shock is very small with a correspondingly small change in entropy.

At times the author has not hesitated to restrict attention to a monatomic gas ($\gamma=5/3$) in order to avoid formal mathematical difficulties.[3] The behavior of the gas undergoes marked changes as the difference $\rho_0-\rho_2$ between the initial densities is permitted to vary.[4]

If $\rho_0-\rho_2$ is sufficiently small the two initial shocks give rise to shocks traveling in opposite directions towards infinity as t increases indefinitely. Up to a certain instant the shocks travel with constant velocity greater than the velocity of sound in the undisturbed gas. After this instant their velocity of propagation decreases monotonically with time, to approach the velocity of sound in the undisturbed gas as the shocks recede to infinity.[5] The behavior of the gas between the two shocks is followed up to a stage when the mapping[6] of Riemann's (r, s)-plane upon the (x, t)-plane loses its one-to-one character. The further behavior of the gas still awaits determination.

Plates 1 and 2 at the end of the paper present qualitatively the variation of density (or pressure), over the gas for $\rho_0-\rho_2$ sufficiently small.

1. Fundamental principles. Assuming that the *pressure* p is a monotonic increasing function of the *density* ρ and denoting the *velocity* by u, the partial differential equations

---

* Received March 8, 1946.

[1] See Riemann-Weber, *Die partiellen Differentialgleichungen der Mathematischen Physik*, 6th ed., Friedr. Vieweg & Sohn Brawnschweig, 1919, vol. 2, pp. 549–550.

[2] Indeed under the Rankine-Hugoniot hypothesis it follows from formula (10) on p. 513 of Riemann-Weber, op. cit. that the entropy of the gas in back of the shock depends upon the ratio of the densities of the gas on the two sides of the shock. This ratio changes as the shock propagates and consequently the the entropy of the gas in back of the shock is not constant.

[3] The examination of other values of γ has been begun by R. C. Rand in his doctorate thesis entitled *The rectilinear motion of a gas subsequent to an internal explosion*. A copy of this thesis is on file in the library of the University of Maryland.

[4] R. C. Rand, loc. cit.

[5] See, however, the last sentence in §6 of the present paper.

[6] For a discussion of this mapping see Riemann-Weber, loc. cit., pp. 533–536 or *The rectilinear motion of a gas*, Amer. J. Math. 65, 391–401 (1943). This paper will be cited as I.

$$\rho(u_t + u_x u) + G^2\rho_x = 0, \qquad \rho_t + (\rho u)_x = 0, \qquad G^2 = G^2(\rho) = p',$$

for u, ρ become

$$r_t + \alpha r_x = 0, \qquad s_t + \beta s_x = 0, \qquad \alpha = u + G, \qquad \beta = u - G, \tag{1}$$

if we set

$$u = r + s, \qquad \int_0^\rho (G/\rho)d\rho = r - s > 0. \tag{2}$$

Clearly u, ρ are monotonic increasing functions of $r+s$, $r-s$ respectively and

$$\alpha = \alpha(r, s) = r + s + G(\rho(r - s)), \quad \beta = \beta(r, s) = r + s - G(\rho(r - s)) \tag{3}$$

satisfy

$$\alpha(-s, -r) = -\beta(r, s), \qquad \beta(-s, -r) = -\alpha(r, s). \tag{4}$$

A point of the (u, ρ)-plane, or its correspondent by (2) in the (r, s)-plane, is said to *represent* or *be* a *state of the gas*. The points of the (r, s)-plane representing states of the gas comprise a half-plane $r \geqq s$ termed the *state plane*. Points representing states having the same velocity (density) lie on the lines $r+s=$const. ($r-s=$const.) and the velocity (density) of a state increases with the distance of the point (r, s) from the line of zero velocity $r=-s$ (the line of zero density $r=s$). The velocity is positive or negative according as (r, s) lies above or below the line $r=-s$.

In general a solution $r=r(x, t)$, $s=s(x, t)$ of (1) transforms a region of the (x, t)-plane into a region in the state plane and is single valued. The inverse transformation T: $x=x(r, s)$, $t=t(r, s)$ is not necessarily single-valued and is regarded as assigning the state (r, s) to its transform (x, t). Corresponding to (1) there is the system

$$x_r - \beta t_r = 0, \qquad x_s - \alpha t_s = 0, \tag{5}$$

of partial differential equations for $x(r, s)$, $t(r, s)$ in T. The Jacobian J of T is

$$J = -(\alpha - \beta)t_r t_s = -2Gt_r t_s. \tag{6}$$

If x, t are solutions of (5), the system of Pfaff

$$dw = (x - \alpha t)dr + (x - \beta t)ds, \qquad dv = 2\frac{x - \alpha t - v}{\alpha - \beta}dr - 2\frac{x - \beta t - v}{\alpha - \beta}ds,$$

is completely integrable, and conversely. When we write

$$x - \alpha t = w_r, \qquad x - \beta t = w_s, \tag{7}$$

the integrability condition for the second equation becomes[7]

$$(\alpha - \beta)w_{rs} - \beta_r(w_r - w_s) = 0. \tag{8}$$

Taking w a solution of (8) it follows from (7) that a transformation T is

$$T_w: \qquad x = -\frac{\beta w_r - \alpha w_s}{\alpha - \beta}, \qquad t = -\frac{w_r - w_s}{\alpha - \beta}.$$

The following theorem is a direct consequence of (4).

[7] Cf. Riemann-Weber, loc. cit., pp. 536–538 or pp. 393–394 of I.

THEOREM 1. *Given $w=w(r, s)$ a solution of* (8), *another solution is $\bar{w}=w(-s, -r)$ and T_w, $T_{\bar{w}}$ map points which are reflections of each other in the line of zero velocity $r=-s$, into points which are reflections of each other in the line $x=0$.*

As a corollary, we see that if $w(r, s)=w(-s, -r)$ points which are reflections of each other in the line $r=-s$, are carried by T_w into points which are reflections of each other in $x=0$.

The theorem is obvious *a priori* on physical grounds. Given any motion of the gas, its particles may be reflected in the plane $x=0$ to gain another motion.

Taking $r=r_0=$const., the second equation in (1) upon multiplication by $d\beta/ds$ becomes

$$\beta_t+\beta\beta_x=\frac{\partial(x-\beta t, \beta)}{\partial(x, t)}=0,$$

and therefore a solution of (1) is given implicitly by[8]

$$r=r_0, \qquad x-\beta t=\Psi(\beta), \tag{9}$$

$\Psi(\beta)$ denoting an arbitrary function of β. Corresponding to $s=s_0=$const., a solution of (1) is obtained from

$$x-\alpha t=\Phi(\alpha), \qquad s=s_0. \tag{10}$$

For a fixed s in (9) the state (r_0, s) is assigned[9] to all points of the straight line $x-\beta t=\Psi(\beta)$. This line is termed a *propagation line* and the state (r_0, s) is said to be *propagated* along it. Physically the state (r_0, s) is propagated through the gas with a velocity β with respect to a fixed plane.

Let us assume that T_w puts the states of a region R of the state plane in (one-to-one correspondence with the points of a region X of the (x, t)-plane. The transform by T_w of a segment of $r=$const. ($s=$const.) in R is a curve in X termed an *r-curve* (*s-curve*). The r and s-curves provide a curvilinear coordinate system on X from which the state of the gas may be read off at any point of X.

From (5) the slope of an $r(s)$-curve[10] is $1/\alpha$ $(1/\beta)$; from (9), (10) the propagation lines drawn from the points of an $r(s)$-curve have slope $1/\beta$ $(1/\alpha)$. Therefore *the tangents drawn to $s(r)$-curves at the points of an $r(s)$-curve are propagation lines and, in so far as they do not intersect, may be used to assigned the states on the $r(s)$-curve to the points of the region covered by them.*

Two $r(s)$-curves C, $\bar{C}$ transforms of $r=r_0$ $(s=s_0)$ under T_w, $T_{\bar{w}}$ respectively are said to be *propagated* from each other if the propagation lines drawn from points of C, $\bar{C}$ which are transforms of the same state are identical.

LEMMA 1. *Two $r[s]$-curves C, $\bar{C}$ transforms by T_w, $T_{\bar{w}}$ of $r=r_0$ $[s=s_0]$ are propagated from each other if, and only if $w_s(r_0, s)=\bar{w}_s(r_0, s)$ $[w_r(r, s_0)=\bar{w}_r(r, s_0)]$.*

From (7) parametric equations of C, $\bar{C}$ are

$$C: \quad x-\alpha(r_0, s)t=w_r(r_0, s), \qquad x-\beta(r_0, s)t=w_s(r_0, s),$$

$$\bar{C}: \quad x-\alpha(r_0, s)t=\bar{w}_r(r_0, s), \qquad x-\beta(r_0, s)t=\bar{w}_s(r_0, s).$$

---

[8] Cf. Riemann-Weber, loc. cit., p. 518.

[9] Cf. Riemann-Weber, loc. cit., pp. 516–520.

[10] First noted by R. C. Rand.

Along a propagation line propagating the state (r_0, s) we have $x-\beta(r_0, s)t=\text{const.}$ Hence propagation lines drawn from a point on C and a point on $\overline{C}$, both transforms of the same state (r_0, s) will be identical if, and only if, $w_s(r_0, s)=\bar{w}_s(r_0, s)$.

It is interesting to note that tangents drawn to C, $\overline{C}$ at points which are transforms of the same state are parallel.

LEMMA 2. *Given two $r[s]$-curves C, $\overline{C}$ which are propagated from each other, curve $\overline{C}$ will pass through a point $(\bar{x}, \bar{t})$ on the propagation line propagating the state $(r_0, \bar{s})$, $[(\bar{r}, s_0)]$ from C if and only if $\bar{w}_r(r_0, \bar{s})=\bar{x}-\alpha(r_0, \bar{s})\bar{t}\,[\bar{w}_s(\bar{r}, s_0)=\bar{x}-\beta(\bar{r}, s_0)\bar{t}]$. This condition determines $\bar{w}_r(r_0, s)$ $[\bar{w}_s(r, s_0)]$ uniquely.*

The first part of the lemma follows from the parametric equations of $\overline{C}$. To prove the second part we set $r=r_0$, $w=\bar{w}$ in (8) to obtain an ordinary, linear differential equation for $\bar{w}_r$ since $\bar{w}_s=w_s$ is a known function of s. This determines $\bar{w}_r$ uniquely, for $\bar{w}_r$ is known when $s=\bar{s}$.

2. Shocks and buffer waves. Under the assumption that G increases with ρ it follows that $G=G(\rho(r-s))$ is an increasing [decreasing] function of $r[s]$ for fixed $s[r]$; from (3), one concludes that $\alpha[\beta]$ is an increasing function of $r[s]$ for fixed $s[r]$.

LEMMA 3. *If initially $r=r_0$ for $-\infty<x<+\infty$ and $s=s_1$ or s_2 as $x<0$ or $x>0$ with $s_1<s_2$, subsequently the state of the gas is unchanged exterior to the "buffer region" between the lines $x=\beta(r_0, s_1)t$, $x=\beta(r_0, s_2)t$. Within this region the state (r_0, s) with $s_1<s<s_2$ is propagated*[11] *along the propagation line $x=\beta(r_0, s)t$.*

Initial states are propagated along the propagation lines

$$x-\beta(r_0, s_1)t = k_1 < 0, \qquad x-\beta(r_0, s_2)t = k_2 > 0, \qquad \beta(r_0, s_1) < \beta(r_0, s_2),$$

which diverge as shown in Figure 1 to assign the state (r_0, s_1) to the region on the left of OA_1 and the state (r_0, s_2) to the region on the right of OA_2. To obtain the states in

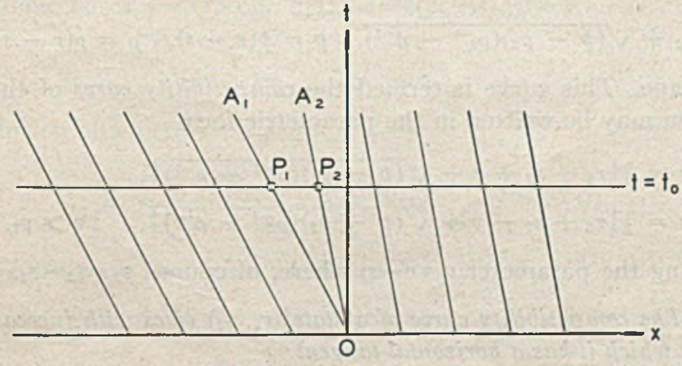

FIG. 1.

the buffer region A_1OA_2 one sets $\Psi(\beta)\equiv 0$ in (9) and draws the propagation line $x=\beta(r_0, s)t$ from 0. Along this propagation line the state is (r_0, s) and as s ranges from s_1 to s_2 the propagation line turns from OA_1 to OA_2 to assign states to all points of the

[11] Cf. Riemann-Weber, loc. cit., pp. 520–521.

buffer region. It will be observed that the states vary continuously along a line $t=t_0>0$.

In this solution of (1), the inverse transformation T is not single-valued the, segment $s_1 \leqq s \leqq s_2$ of $r=r_0$ being carried by T into the half-plane $t \geqq 0$.

Physically the buffer region corresponds to a disturbance P_1P_2 affecting two bodies of gas of different uniform states in contact with one another initially, the end points of the disturbance traveling with the local velocity of sound in the two bodies of gas. The passage of this disturbance through the gas is termed a *buffer wave.*

A shock exists at $x=\xi$ if $\rho_1 \neq \rho_2$ and is propagated with a velocity[12]

$$\dot{\xi} = u_1 \pm \sqrt{\frac{\rho_2}{\rho_1} \frac{p_1 - p_2}{\rho_1 - \rho_2}} = u_2 \pm \sqrt{\frac{\rho_1}{\rho_2} \frac{p_1 - p_2}{\rho_1 - \rho_2}}, \tag{11}$$

where

$$u_1 = u(\xi - 0), \qquad \rho_1 = \rho(\xi - 0), \qquad p_1 = p(\xi - 0),$$
$$u_2 = u(\xi + 0), \qquad \rho_2 = \rho(\xi + 0), \qquad p_2 = p(\xi + 0).$$

The curve $x=\xi(t)$ in the (x, t)-plane is termed a *shock curve.* It will be sufficient for the purposes of this investigation to consider progressive condensation shocks arising when $\rho_1 > \rho_2$ and the positive sign is taken in (11). For a shock of this type one has the condition

$$u_1 - u_2 = \sqrt{(p_1 - p_2)(\rho_2^{-1} - \rho_1^{-1})}, \tag{12}$$

with

$$\dot{\xi} = (u_1\rho_1 - u_2\rho_2)(\rho_1 - \rho_2)^{-1}. \tag{13}$$

If (r_1, s_1), (r_2, s_2) denote the correspondents of (u_1, ρ_1), (u_2, ρ_2) by (2) and the state (r_2, s_2) on the right of the shock is given, the state (r_1, s_1) on the left of the shock is not uniquely determined but, by (12), may be any point of the curve.

$$r + s = r_2 + s_2 + \sqrt{(p - p_2)(\rho_2^{-1} - \rho^{-1})}, \quad p = p(r - s), \quad \rho = \rho(r - s) > \rho_2 \tag{14}$$

in the state plane. This curve is termed the *compatibility curve* of the state (r_2, s_2) and its equation may be written in the parametric form

$$r = \tfrac{1}{2}\{r_2 + s_2 + v + \sqrt{(p - p_2)(\rho_2^{-1} - \rho^{-1})}\},$$
$$s = \tfrac{1}{2}\{r_2 + s_2 - v + \sqrt{(p - p_2)(\rho_2^{-1} - \rho^{-1})}\}, \qquad v > v_2, \tag{14'}$$

upon introducing the parameter $v=r-s$, where, of course, $v_2=r_2-s_2$.

LEMMA 4. *The compatibility curve of a state (r_2, s_2) rises with increasing r from the point (r_2, s_2), at which it has a horizontal tangent.*

Both derivatives of r, s with respect to v will be positive provided

$$\frac{\rho}{\rho_2} G^2(\rho) > \frac{p - p_2}{\rho - \rho_2} = G^2(\bar{\rho}) \quad \text{where} \quad \rho_2 < \bar{\rho} < \rho,$$

---

[12] See, for example, Riemann-Weber, op. cit. p. 513.

which inequality holds for $\rho > \rho_2$ under the assumption that G increases with ρ.

We shall now consider the simultaneous generation of a shock and buffer wave[13] as pictured in Figure 2 where B is a buffer region between two regions R_0, R_1 of uniform state (r_0, s_0), (r_1, s_1) respectively and OS is the shock line separating R_1 from the region R_2 of uniform state (r_2, s_2).

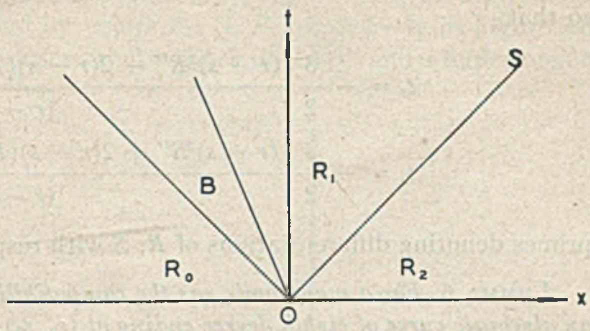

FIG. 2.

LEMMA 5. *A shock and buffer wave are generated simultaneously*[13] *at the contact of two bodies of gas of different uniform states* (r_0, s_0), (r_2, s_2) *provided the point* (r_0, s_0) *in the state plane lies directly underneath the compatibility curve of the state* (r_2, s_2).

Choosing the state (r_2, s_2) in R_2 arbitrarily, the state in R_1 must be represented by a point on the compatibility curve of the state (r_2, s_2); and if this point lies directly above (r_0, s_0) the existence of the buffer region B is assured by Lemma 3.

3. The isentropic case. Here $p = k^2\rho^\gamma$ with $k, \gamma > 1$ constants and G increases with ρ so that the results of §2 remain in force. Moreover

$$G = \frac{\gamma - 1}{2}(r - s), \qquad \alpha = \frac{1+\gamma}{2} r + \frac{3-\gamma}{2} s, \qquad \beta = \frac{3-\gamma}{2} r + \frac{1+\gamma}{2} s, \tag{15}$$

and (8) becomes

$$(r - s)w_{rs} - m(w_r - w_s) = 0, \qquad m = \frac{3-\gamma}{2(\gamma - 1)}. \tag{16}$$

For monatomic gases $\gamma = 5/3$ one finds $\alpha = \frac{2}{3}(2r+s)$, $\beta = \frac{2}{3}(r+2s)$. Also $m = 1$ and (16) becomes

$$(r - s)w_{rs} - (w_r - w_s) = 0, \tag{16'}$$

the general solution of which is

$$w = \frac{R - S}{r - s}, \qquad R = R(r), \qquad S = S(s), \tag{17}$$

$R(r)$, $S(s)$ being arbitrary functions. The transformation T_w is

$$x = -\frac{(r + 2s)w_r - (2r + s)w_s}{r - s}, \qquad t = -\frac{3}{2}\,\frac{w_r - w_s}{r - s}, \tag{18}$$

or, from (17)

$$x = 3\,\frac{r + s}{(r - s)^3}(R - S) - \frac{(r + 2s)R' + (2r + s)S'}{(r - s)^2},$$
$$t = \frac{3(R - S)}{(r - s)^3} - \frac{3}{2}\,\frac{R' + S'}{(r - s)^2}, \tag{18'}$$

[13] Cf. Riemann-Weber, loc. cit., pp. 527–529.

so that

$$t_r = -\frac{3}{2}\,\frac{(r-s)^2R'' - 2(r-s)(2R'+S') + 6(R-S)}{(r-s)^4},$$
$$t_s = -\frac{3}{2}\,\frac{(r-s)^2S'' + 2(r-s)(R'+2S') - 6(R-S)}{(r-s)^4}, \qquad (19)$$

primes denoting differentiations of R, S with respect to their arguments.

LEMMA 6. *For a monatomic gas the compatibility curve of a state* (r_2, s_2) *is an arc of an algebraic curve of eighth degree ending at* (r_2, s_2), *about which the points of the arc permit the expansions*

$$r = r_2 + v - v_2 + \kappa(v - v_2)^3 + \cdots, \quad s = s_2 + \kappa(v - v_2)^3 + \cdots, \quad \kappa = 1/3v_2^2. \qquad (20)$$

At all other points r is a regular analytic function of s with a positive derivative and r, s are regular analytic functions (14′) *of the uniformizing parameter v, with respect to which they possess positive derivatives.*

For monatomic gases equations (14), (14′) become

$$r + s = r_2 + s_2 + \sqrt{\tfrac{1}{15}[(r-s)^5 - (r_2-s_2)^5][(r_2-s_2)^{-3} - (r-s)^{-3}]}, \qquad (21)$$

$$r = \tfrac{1}{2}\left[r_2 + s_2 + v + \sqrt{\tfrac{1}{15}(v^5 - v_2^5)(v_2^{-3} - v^{-3})}\right],$$
$$s = \tfrac{1}{2}\left(r_2 + s_2 - v + \sqrt{\tfrac{1}{15}(v^5 - v_2^5)(v_2^{-3} - v^{-3})}\right], \qquad (21')$$

from which the statements in the lemma follow straight forwardly.

We return to the general adiabatic case. A point in the state plane represents a state for which the velocity is subsonic, sonic or supersonic according as the point lies in, on the boundary of, or exterior to the region $\alpha > 0$, $\beta < 0$ between the straight lines $\alpha = 0$, $\beta = 0$.

LEMMA 7. *The angle of inclination* θ *of the tangent at a point of an* $r[s]$*-curve is less* [*greater*] *than the angle of inclination* ϕ *of the propagation line drawn from this point. Both angles lie between* 0 *and* π *and are decreasing functions of* $s[r]$.

The lemma is obvious in view of (15) and previous results in §1 on the slopes of r, s-curves and propagation lines.

LEMMA 8. *If* T_w *puts a region R of the state plane in* (1-1) *correspondence with a region X of the* (x, t)*-plane and if the Jacobian J of* T_w *never vanishes in R, the curvature of an* $r[s]$*-curve in X has a fixed sign and the parts of the propagation lines drawn on the convex side do not intersect.*

4. The first initial value problem. Returning to the problem formulated in the introduction, the correspondents of the initial states $(0, \rho_0)$, $(0, \rho_2)$ of the gas are represented by the points $P_0(r_0, s_0)$, $P_2(r_2, s_2)$ of the state plane in Figure 3a. Both P_0, P_2 lie on the line of zero velocity $r+s=0$, with $r_0 > r_2$, since $\rho_0 > \rho_2$.

From Lemma 4 we observe that P_0 lies directly underneath a point $Q(r_0, s_1)$ of the compatibility curve of P_2 and therefore, according to Lemma 5, a shock and buffer wave are generated simultaneously in the gas at $x=1$. In Figure 3b the shock line from $A(1, 0)$ is AQ'' and the buffer region is $P_0'AQ'$. States in the regions OAP_0', $Q'AQ''$, $Q''Ax$ are represented by points P_0, Q, P_2 respectively in Figure 3a. It is obvious from symmetry considerations that a shock line $\overline{A}\overline{Q}''$ and a buffer region

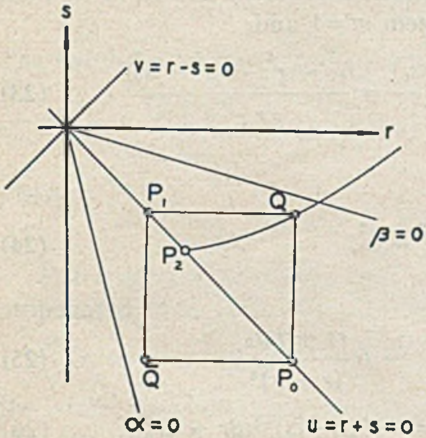

FIG. 3a.

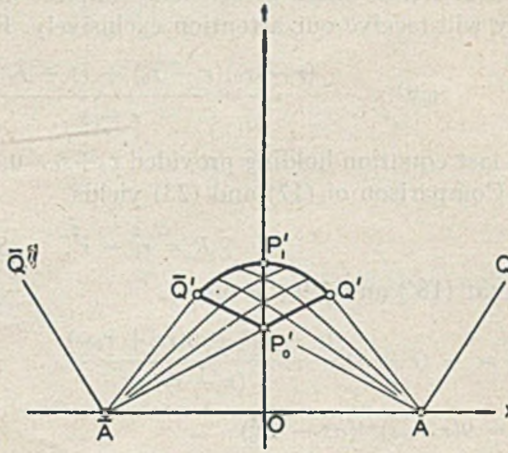

FIG. 3b.

$P_0'\overline{A}\overline{Q}'$ emanate from $\overline{A}(-1, 0)$ and that states in the regions $O\overline{A}P_0'$, $\overline{Q}'\overline{A}\overline{Q}''$, $\overline{Q}''\overline{A}x$ are represented by points P_0, $\overline{Q}$, P_2 in Figure 3a with $\overline{Q}$ the reflection of Q in the line $r+s=0$. In the buffer region emanating from $A[\overline{A}]$, we have $r=r_0[s=s_0]$ and the equations of the propagation line are

$$x - \beta(r_0, s)t = +1, \qquad s_0 \leqq s \leqq s_1, \qquad [x - \alpha(r, s_0)t = -1, \qquad r_1 \leqq r \leqq r_0] \qquad (r_1 = -s_1). \quad (22)$$

As $s[r]$ ranges from $s_0[r_0]$ to $s_1[r_1]$ the propagation line from $A[\overline{A}]$ turns from $AP_0'[\overline{A}P_0']$ to $AQ'[\overline{A}\overline{Q}']$, with $t=G_0^{-1}$ where $G_0=G(\rho_0)$ at P_0'. These propagation lines intersect on the t-axis above P_0' to assign different states to their intersection points. We avoid such a physical impossibility by terminating them on the arcs $P_0'Q'$, $P_0'\overline{Q}'$ in Figure 3b. The propagation lines assign the states on $\overline{Q}P_0Q$ to the points of $\overline{Q}'P_0'Q'$ and we seek a T_w which carries $\overline{Q}P_0Q$ into $\overline{Q}'P_0'Q'$ and assigns the same states to the same points of the latter arc. A comparison of (7) with (22) leads to the following initial value problem.

THE FIRST INITIAL VALUE PROBLEM. *Given two constants r_0, s_0, find a solution $w^{(1)}$ of* (16) *for which $w_r^{(1)}(r, s_0) = -1$, $w_s^{(1)}(r_0, s) = +1$.*

Before giving the solution for the general adiabatic case, we recall a few facts concerning the resolvent[14] of (16). This resolvent is a two parameter family of solutions $v=v(r, s; r_0, s_0)$ of the conjugate equation $(r-s)v_{rs}+m(v_r-v_s)=0$ meeting the initial conditions $v_r(r, s_0; r_0, s_0)=+1$, $v_s(r_0, s; r_0, s_0)=-1$, and is given by

$$v = (r - r_0)\left(\frac{r_0 - s}{r_0 - s_0}\right)^m F_1\left(1 - m; m; -m; 2; \frac{r - r_0}{s_0 - r_0}; \frac{r - r_0}{s - r_0}\right)$$
$$- (s - s_0)\left(\frac{r - s_0}{r_0 - s_0}\right)^m F_1\left(1 - m; m; -m; 2; \frac{s - s_0}{r_0 - s_0}; \frac{s - s_0}{r - s_0}\right),$$

[14] See I, in particular §3 and §5.

where F_1 is Appell's first hypergeometric function of two variables. *The solution* $w^{(1)}$ *of the first initial value problem is obtained by replacing* m *by* $-m$ *and changing the sign of the resolvent.*

Monatomic gases present the simplest mathematical problem and from now on they will receive our attention exclusively. For them $m=1$ and

$$w^{(1)} = -\frac{(r-r_0)(r-s_0)+(s-r_0)(s-s_0)}{r-s} = \frac{r_0^2-r^2-(s^2-s_0^2)}{r-s}, \tag{23}$$

the last equation holding provided $r_0+s_0=0$.

Comparison of (17) and (23) yields

$$R = r_0^2 - r^2, \qquad S = s^2 - s_0^2; \tag{24}$$

so that (18′) and (19) become

$$x^{(1)} = -(r+s)\frac{(r-s)^2+6(rs+r_0s_0)}{(r-s)^3}, \qquad t^{(1)} = -6\frac{rs+r_0s_0}{(r-s)^3}, \tag{25}$$

$$t_r^{(1)} = 9(r-s)^{-4}(\alpha s - 2r_0^2) \qquad t_s^{(1)} = -9(r-s)^{-4}(\beta r - 2r_0^2), \tag{26}$$

where the superscripts record that $w=w^{(1)}$ in T_w. In the subsonic region of the state plane $\alpha>0$, $\beta<0$ and, therefore in this region

$$t_r^{(1)} < 0, \qquad t_s^{(1)} > 0, \qquad J^{(1)} > 0. \tag{27}$$

The square $P_0QP_1\bar{Q}$ in Figure 3a is termed the *primary region.* As ρ_0 increases from ρ_2 the primary region expands from point P_2 till eventually P_1 leaves the state plane. We consider only values of ρ_0 for which the primary region lies entirely in the subsonic region and forego examination of the several interesting cases which arise[15] when this is not the case.

Arc $P_0'Q'$, the transform of P_0Q by $T_{w^{(1)}}$, is tangent to $\bar{A}P_0'$ at P_0' and has slope $\tan\theta = 1/\alpha > 0$. Since $t^{(1)}$ increases by (27) and the acute angle θ decreases by Lemma 7 with increasing s, arc $P_0'Q'$ is concave downwards. Likewise arc $Q'P_1'$ the transform of QP_1 by $T_{w^{(1)}}$, is concave downwards. From (23) and the corollary to Theorem 1 it follows that arcs $P_0'\bar{Q}'$, $\bar{Q}'P_1'$ are concave downwards.

The boundary $P_0QP_1\bar{Q}$ of the primary region and the boundary $P_0'Q'P_1'\bar{Q}'$ of its transform under $T_{w^{(1)}}$ are in one-to-one correspondence with $J^{(1)}>0$ holding in the interior of the primary region. It follows[16] that the interiors of the two regions are in one-to-one correspondence to assign a unique state to each point of the region $P_0'Q'P_1'\bar{Q}'$ in Figure 3b.

5. The second initial value problem. To extend our knowledge of the states of the gas we draw propagation lines from the points of the arc $Q'P_1'$. These propagation lines are tangent to r-curves on $Q'P_1'$ and according to Lemma 8 do not intersect on the convex side of $Q'P_1'$.

The equation of the propagation line from Q' is

$$x - \frac{2}{3}(2r_0+s_1)t = w_r^{(1)}(r_0, s_1) = -1 - 2\frac{r_0+s_1}{r_0-s_1},$$

[15] A beginning in this direction has been made by Rand, loc. cit.,

[16] See, for example, G. A. Bliss, *Fundamental existence theorems* vol. III, Amer. Math. Soc. Colloquium Publications, reprinted 1934, p. 42.

and the equation of the shock line from A is, from (13), since $r_2+s_2=0$,

$$x - \frac{(r_0+s_1)(r_0-s_1)^3}{(r_0-s_1)^3-(r_2-s_2)^3}\,t = 1.$$

The two lines intersect in a point Q'' with coordinates

$$x'' = 1 + \frac{12r_0}{\lambda(r_0, s_1)}(r_0+s_1)(r_0-s_1)^2, \quad t'' = \frac{12r_0}{\lambda(r_0, s_1)}\,\frac{(r_0-s_1)^3-(r_2-s_2)^3}{r_0-s_1}, \tag{28}$$

where

$$\lambda(r, s) = (r-s)^4 - 2(2r+s)(r_2-s_2)^3. \tag{29}$$

Referring to Lemmas 1 and 2 a solution $w^{(2)}$ of (16′) transforms QP_1 into an s-curve propagated from $Q'P_1'$ and containing Q'' if

$$w_r^{(2)}(r, s_1) = w_r^{(1)}(r, s_1), \quad r_1 \leqq r \leqq r_0, \qquad w_s^{(2)}(r_0, s_1) = x'' - \tfrac{2}{3}(r_0+2s_1)t'', \tag{30}$$

the latter condition determining $w_s^{(2)}$ uniquely on QP_1.

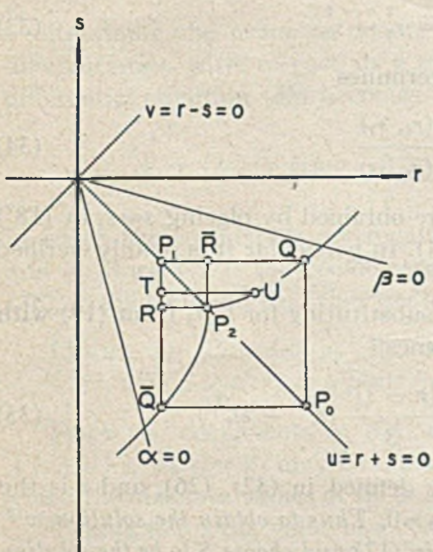

FIG. 4a.

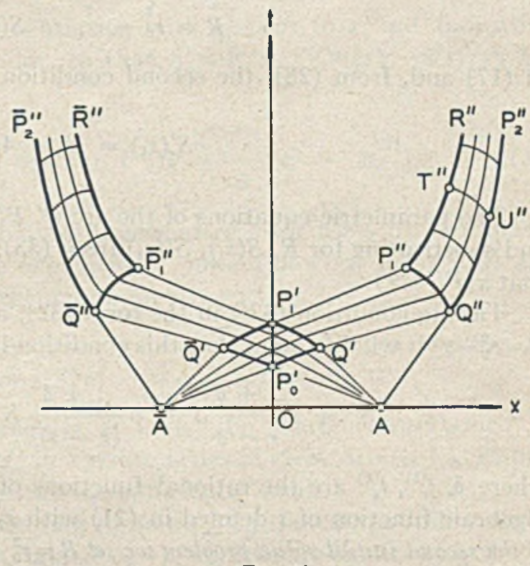

FIG. 4b.

Let the arc $Q''P_2''$ in Figure 4b indicate the prolongation of the shock line AQ''. On the right of $Q''P_2''$ the state of the gas is $P_2(r_2, s_2)$ and the states immediately on the left of $Q''P_2''$ are represented by points (r, s) on the compatibility curve (21). Thus $Q''P_2''$ is the transform by $T_{w^{(2)}}$ of the compatibility curve QP_2.

On the one hand the slope of $Q''P_2''$ is

$$\frac{dt}{dx} = \frac{t_r r' + t_s}{x_r r' + x_s} = \frac{3}{2}\,\frac{t_r r' + t_s}{(r+2s)t_r r' + (2r+s)t_s},$$

where $r=r(s)$ is defined implicitly by (21) and its derivative r' is

$$r' = \frac{8(r-s)^5 - 20(r_2-s_2)^3(r+2s)^2 - 3(r_2-s_2)^5}{8(r-s)^5 - 20(r_2-s_2)^3(2r+s)^2 - 3(r_2-s_2)^5}.$$

On the other hand, from (13)

$$\frac{dt}{dx} = \frac{(r-s)^3 - (r_2 - s_2)^3}{(r+s)(r-s)^3}, \tag{31}$$

and a comparison of the two results yields the condition

$$\frac{t_s}{t_r} = \frac{\mu(r,s)}{\lambda(r,s)} r' = \delta, \qquad \mu(r,s) = (r-s)^4 + 2(r+2s)(r_2-s_2)^3, \tag{32}$$

along QP_2, or

$$\delta w_{rr}^{(2)} - 2(1+\delta) w_{rs}^{(2)} + w_{ss}^{(2)} = 0. \tag{32'}$$

The Second Initial Value Problem. *To construct a solution $w^{(2)}$ of (16′) meeting the conditions (30) on the side QP_1 of the primary region and the condition (32′) along the arc QP_2 of the compatibility curve.*

From (24) the first condition in (30) is met by taking

$$R = r_0^2 - r^2, \qquad S(s_1) = s_1^2 - s_0^2, \tag{33}$$

in (17) and, from (28), the second condition determines

$$S'(s_1) = 2s_1 - 4r_0 \frac{\mu(r_0, s_1)}{\lambda(r_0, s_1)} \cdot \tag{34}$$

The parametric equations of the arc $Q''P_1''$ are obtained by placing $s = s_1$ in (18′) and substituting for R, $S(s_1)$, $S'(s_1)$ from (33), (34). In particular it is readily verified that $x(P_1'') > 1$.

Taking condition (32′) in the form (32), and substituting for t_r, t_s from (19) with $R = r_0^2 - r^2$ it will be found that this condition becomes[17]

$$S'' + 2\frac{2+\delta}{r-s} S' + 6\frac{1+\delta}{(r-s)^2} S = \frac{(r-s)^2}{3}\left(t_s^{(1)} - \delta t_r^{(1)}\right), \tag{35}$$

where δ, $t_r^{(1)}$, $t_s^{(1)}$ are the rational functions of r, s defined in (32), (26), and r is the algebraic function of s defined in (21) with $r_2 + s_2 = 0$. *Thus to obtain the solution $w^{(2)}$ of the second initial value problem we set $R = r_0^2 - r^2$ in* (17) *and choose S to be the solution of the ordinary differential equation of second order* (35), *subject to the initial conditions in* (33), (34).

Lemma 9. *The value of δ at a point P of the compatibility curve QP_2 tends to $+\infty$ as P tends to P_2. More precisely δ is a positive regular analytic function of the parameter v on QP_2, except at $v = v_2$, where it has a pole of the third order and a Laurent expansion of the form*

$$\delta = v_2^3 (v - v_2)^{-3} (1 + \cdots). \tag{36}$$

To prove $\delta > 0$ we have

$$\lambda = v(v^3 - v_2^3) - 3uv_2^3, \qquad \mu = v(v^3 - v_2^3) + 3uv_2^3,$$

[17] The form of the second member in (35) is due to Rand.

and the equation of the compatibility curve QP_2

$$15v_2^3u^2v^3 = (v^5 - v_2^5)(v^3 - v_2^3).$$

It is obvious that $\mu>0$ for $v>v_2$, and $\lambda>0$ follows from $\lambda\mu=v^2(v^3-v_2^3)^2-9u^2v_2^6>0$ for $v>v_2$. To establish this inequality, one multiplies the equation of the compatibility curve by $3v_2^3$ and observes that $3v_2^3(v^5-v_2^5)<5v^5(v^3-v_2^3)$ holds for $v>v_2$. Now $r'>0$ by Lemma 6 and therefore $\delta>0$ for $v>v_2$ by (32).

From (20) one has $u=r+s=v-v_2+2\kappa(v-v_2)^3+\cdots$, so that

$$\lambda = 6v_2^2(v - v_2)^2 + \cdots, \quad \mu = 6v_2^3(v - v_2) + \cdots, \quad r' = v_2^2(v - v_2)^{-2}(1 + \cdots), \tag{37}$$

hold along QP_2, and the Laurent expansion for δ then follows from (32).

It is apparent from (20) and Lemma 9 that the coefficients of the differential equation (35) present a singular point at $s=s_2$.

LEMMA 10. *The introduction of v as independent variable in the differential equation* (35) *leads to a differential equation for $V=S(s(v))$ in which the coefficients are regular analytic functions of v for $v\geqq v_2$.*

Retaining the prime to denote differentiation with respect to s and indicating differentiation with respect to v by a dot, so that $S'=\dot V/\dot s$, $S''=(\dot s\ddot V-\dot V\ddot s)/\dot s^3$, the differential equation (35) becomes

$$\ddot V + (2(2 + \delta)\dot s/v - \ddot s/\dot s)\dot V + 6\dot s^2v^{-2}(1 + \delta)V = \frac{v^2\dot s^2}{3}(t_s^{(1)} - \delta t_r^{(1)}), \tag{38}$$

in which the coefficients are regular analytic functions of v for $v>v_2$ by Lemmas 6 and 9. Moreover if the coefficients are expanded in powers of $v-v_2$ using (20) and (36), it will be found that they are also regular about v_2.

LEMMA 11. *Provided $\rho_0-\rho_2>0$ is sufficiently small, S and S' are negative for $s_2\leqq s\leqq s_1$ with S tending to a finite limit and S' to $-\infty$ as s approaches s_2.*

Since the coefficients in (38) are regular at $v=v_2$, the solution determined by $V(v_0)=V_0$, $\dot V(v_0)=\dot V_0$ may be expanded[18] in a power series in $v-v_2$, v_0-v_2, V_0, $\dot V_0$ provided the absolute values of these quantities are sufficiently small.

Taking v_0 for the value of v corresponding to point Q on the compatibility curve, v_0-v_2 can be made arbitrarily small by taking $\rho_0-\rho_2$ sufficiently small, with the coordinates of Q given by

$$r_0 = r_2 + v_0 - v_2 + \kappa(v_0 - v_2)^3 + \cdots, \quad s_1 = s_2 + \kappa(v_0 - v_2)^3 + \cdots. \tag{39}$$

The initial conditions for S in (33), (34) lead to the initial conditions

$$V(v_0) = s_1^2 - r_0^2, \qquad \dot V(v_0) = 2\dot s(v_0)\left[s_1 - 2r_0\frac{\mu(r_0, s_1)}{\lambda(r_0, s_1)}\right], \tag{40}$$

for V. From (37), (39) we obtain the expansions

$$V(v_0) = -v_2(v_0 - v_2) - (v_0 - v_2)^2 + \cdots, \quad \dot V(v_0) = -2(v_0 - v_2) + \cdots, \tag{41}$$

[18] J. Horn, *Gewöhnliche Differentialgleichungen beliebiger Ordnung*, Sammlung Schubert, vol. 50, Leipzig, 1905, pp. 27–28.

valid for sufficiently small $|v_0-v_2|$. It follows from (38) that the expansion of $\ddot{V}(v_0)$ in powers of v_0-v_2 begins with a term of at least first degree in v_0-v_2.

When the expansions (41) are substituted in the expansion of the solution V in powers of $v-v_2$, v_0-v_2, V_0, $\dot{V}_0$ it appears that V may be expanded in powers of $v-v_2$, v_0-v_2 provided $|v-v_2|$, $|v_0-v_2|$ are sufficiently small. To obtain the linear and quadratic terms of this expansion, we substitute from (41) in Taylor's series

$$V = V(v_0) + \dot{V}(v_0)(v - v_0) + \frac{\ddot{V}(v_0)}{2}(v - v_0)^2 + \cdots$$

to obtain

$$V = - v_2(v_0 - v_2) - 2(v_0 - v_2)(v - v_2) + (v_0 - v_2)^2 + \cdots ,$$

the third term in Taylor's series being neglected since $\ddot{V}(v_0)$ contains the factor v_0-v_2.

It follows that both V, $\dot{V}$ are negative for $v_2 \leqq v \leqq v_0$ for sufficiently small $v_0-v_2>0$. It is clear that S tends to a finite negative limit as s tends to s_2 and, since $\dot{s}$ is positive and tends to zero as v tends to v_2, one concludes that S' tends to $-\infty$ as s tends to s_2, provided, of course, that $\rho_0-\rho_2$ is sufficiently small.

The subregion $P_2QP_1RP_2$ in Figure 4a of the primary region is termed the *secondary region.*

LEMMA 12. *The partial derivatives $t_r^{(2)}$, $t_s^{(2)}$ and the Jacobian $J^{(2)}=-2Gt_r^{(2)}t_s^{(2)}$ of $T_{w^{(2)}}$ are negative in the secondary region for sufficiently small $\rho_0-\rho_2>0$.*

We take R, S in (19) as determined by the second initial value problem and find

$$t_r^{(2)} = \frac{1}{2} t_r^{(1)} + \frac{3S'}{(r - s)^3} + \frac{9S}{(r - s)^4}, \quad t_s^{(2)} = \frac{1}{2} t_s^{(1)} - \frac{3}{2}\frac{S''}{(r - s)^2} - \frac{6S'}{(r - s)^3} - \frac{9S}{(r - s)^4},$$

from which $t_r^{(2)}<0$ follows from (27) and Lemma 11 for sufficiently small $\rho_0-\rho_2>0$.

To prove $t_s^{(2)}<0$ we have

$$S'' + 2\frac{2 + \bar{\delta}}{\bar{r} - s} S' + 6 \frac{1 + \bar{\delta}}{(\bar{r} - s)^2} S = \frac{(\bar{r} - s)^2}{3} (\bar{t}_s^{(1)} - \bar{\delta}\bar{t}_r^{(1)}),$$

where $\bar{r}=r(s)$ is the function of s defined in Lemma 6 and $\bar{\delta}=\delta(\bar{r}, s)$, $\bar{t}_r^{(1)}=t_r^{(1)}(\bar{r}, s)$, $\bar{t}_s^{(1)}=t_s^{(1)}(\bar{r}, s)$. When S'' is eliminated from $t_s^{(2)}$ it is found that $t_s^{(2)}=AS'+BS+C$, where

$$A = \frac{6}{(\bar{r} - s)(r - s)^2}\left[1 + \frac{\bar{\delta}}{2} - \frac{\bar{r} - s}{r - s}\right] > 6(\bar{r} - s)^{-1}(r - s)^{-2}\left(\frac{\bar{\delta}}{2} - \frac{r_0}{r_1}\right),$$

$$B = \frac{9}{(r - s)^2(\bar{r} - s)^2}\left[1 + \bar{\delta} - \left(\frac{\bar{r} - s}{r - s}\right)^2\right] > 9(r - s)^{-2}(\bar{r} - s)^{-2}\left(\bar{\delta} - \frac{r_0^2}{r_1^2}\right),$$

$$C = \frac{1}{2} t_s^{(1)} + \frac{1}{2}\left(\frac{\bar{r} - s}{r - s}\right)^2(\bar{\delta}\bar{t}_r^{(1)} - \bar{t}_s^{(1)}) < \frac{1}{2}(\bar{\delta}\bar{t}_r^{(1)} + \bar{t}_s^{(1)}),$$

in view of the inequalities

$$2r_1 = r_1 - s_1 < r - s \leqq \bar{r} - s < r_0 - s_0 = 2r_0,$$

valid in the secondary region. $t_r^{(1)}$ has a negative upper bound and $t_s^{(1)}$ a positive lower bound in the secondary region independent of ρ_0. Moreover r_0, r_1 tend to r_2 as ρ_0 approaches ρ_2. In view of Lemma 9 we have $A>0$, $B>0$, $C<0$, and therefore $t_s^{(2)}<0$ by Lemma 11 for sufficiently small $\rho_0-\rho_2$.

We shall now investigate the mapping by $T_{w^{(2)}}$ of the secondary region upon the (x, t)-plane. Taking R, S in (18′) as determined above in the solution of the second initial value problem, $T_{w^{(2)}}$ is

$$x^{(2)} = \frac{(r+s)(3r_0^2 - r^2) - 4rs^2}{(r-s)^3} - 3\frac{r+s}{(r-s)^3}S - \frac{2r+s}{(r-s)^2}S',$$

$$t^{(2)} = 3\frac{r_0^2 - rs}{(r-s)^3} - \frac{3S}{(r-s)^3} - \frac{3}{2}\frac{S'}{(r-s)^2}.$$

From Lemma 11 it follows (at least for $\rho_0-\rho_2$ sufficiently small) that $x^{(2)}$, $t^{(2)}$ become infinite as the point (r, s) of the secondary region approaches the side RP_2. We shall accordingly consider first the mapping by $T_{w^{(2)}}$ of the subregion UQP_1TU, the line TU being parallel to RP_2.

Sides P_1Q, TU transform into s-curves $P_1''Q''$, $T''U''$. From Lemma 12 t decreases and from (5) x increases as r increases along P_1Q, TU. We conclude from Lemma 7 that $P_1''Q''$, $T''U''$ are concave downward as shown in Figure 4b.

Side P_1T transforms into an r-curve $P_1'\,T''$ which is concave upwards.

Arc QU of the compatibility curve transforms into the shock curve $Q''U''$. Along QU x and t are monotonic decreasing functions of v, as is the slope dt/dx of $Q''U''$, for, from (31)

$$\frac{d}{dv}\left(\frac{dt}{dx}\right) = -\frac{\lambda\dot{r} + \mu\dot{s}}{(r+s)^2(r-s)^4},$$

inasmuch as $\lambda>0$, $\mu>0$ hold on QU. The shock curve is accordingly concave upwards to imply that the velocity of propagation of the shock decreases as t increases.

Finally we let U approach P_2 along QP_2. The s-curve $T''U''$ recedes to infinity in the (x, t)-plane and the secondary region, exclusive of side RP_2, is accordingly mapped in (1-1) fashion by $T_{w^{(2)}}$ upon a region indicated by $P_2''Q''P_1'\,R''$ in the (x, t)-plane to determine the states of the gas in this region.

The slope of the r-curve $P_1''T''$ at T'' tends to $1/\alpha(r_1, s_2)$ as T'' recedes to infinity. The slope of the shock curve at U'' is, from (11),

$$\frac{dt}{dx} = \sqrt{\frac{\rho_2}{\rho}\,\frac{\rho-\rho_2}{p-p_2}},$$

where ρ denotes the density for the state U and $p=p(\rho)$. As $U\to P_2$ we have $\rho\to\rho_2$, which implies $dt/dx\to 1/G_2$. This means that the velocity of propagation of the shock tends toward the local velocity of sound in the exterior body of gas through which the shock travels as it recedes to infinity.

The determination of the states of the gas in the region $\overline{P}_2''\overline{Q}''\overline{P}_1''\,\overline{R}''$ may now be left to symmetry considerations or Theorem 1.

6. The third initial value problem. We take up the problem of determining the states of the gas in the region of the (x, t)-plane lying above the curve $\overline{R}''\overline{P}_1'\ P_1'P_1'\ R''$ in Figure 4b.

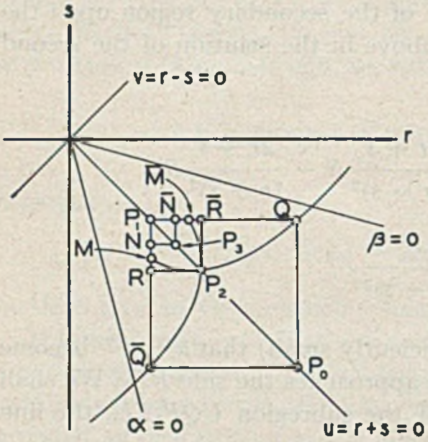

FIG. 5a.

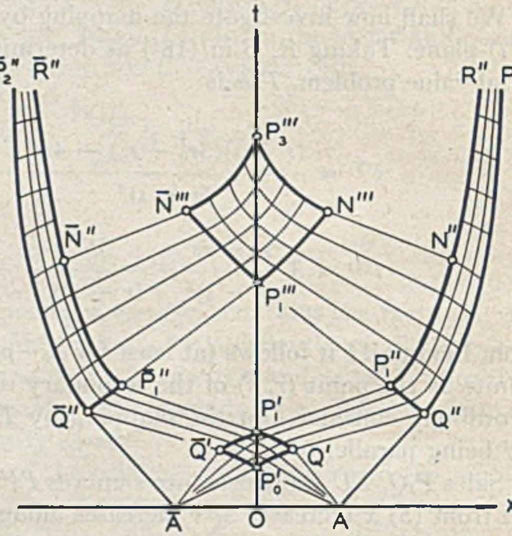

FIG. 5b.

Propagation lines drawn from $P_1''R''$ in Figure 5b have slope $1/\beta<0$ and, from Lemma 8 can intersect only on the concave side of $P_1''R''$. We shall prove that they do not meet in the region $x>0$ if $\rho_0-\rho_2$ is sufficiently small. Since ϕ is a monotonic decreasing function of s by Lemma 7, it will be sufficient to prove that the t-intercept T of a propagation line is a monotonic decreasing function of s.

In the equation of a propagation line $t=\beta^{-1}x+T$ we replace x, t by the coordinates of a point on $P_1''R''$ obtained from $T_{w^{(2)}}$ to obtain

$$T = -\frac{w^{(2)}(r_1, s)}{\beta(r_1, s)} = \frac{(r_1 - s)S' + S + r_1^2 - r_0^2}{\beta(r_1, s)(r_1 - s)^2};$$

from which

$$\frac{dT}{ds} = \frac{\beta(r_1, s)(r_1 - s)^2 S'' + 4s(r_1 - s)S' + 4sS + 4s(r_1^2 - r_0^2)}{\beta^2(r_1, s)(r_1 - s)^3}.$$

After S'' is eliminated by (35) it will be found that $dT/ds<0$ holds for sufficiently small $\rho_0-\rho_2$. The principle of the argument is essentially the same as the one employed to prove that $t_s^{(2)}<0$ in Lemma 12 and is omitted.

From symmetry considerations propagation lines drawn from $\overline{P}_1''\overline{R}''$ do not intersect in the region $x<0$. Propagation lines drawn from $P_1''R''$ and $\overline{P}_1''\overline{R}''$ symmetrically placed with respect to the t-axis intersect upon it and, excepting the two drawn from P_1'', $\overline{P}_1''$, assign different states to their points of intersection. This is avoided in Figure 5b by terminating the propagation lines on arcs $P_1'''N'''$, $P_1'''\overline{N}'''$, the coordinates of P_1''' being $x=0$, $t=-w_s^{(2)}(r_1, s_1)/\beta(r_1, s_1)$.

By Lemma 1 an r-curve, the transform of P_1R by $T_{w^{(3)}}$ will be propagated from $P_1''R''$ provided

$$w_s^{(3)}(r_1, s) = w_s^{(2)}(r_1, s) \quad \text{for} \quad s_2 \leqq s \leqq s_1, \tag{42}$$

and by Lemma 2 will contain P_1'', in case

$$w_r^{(3)}(r_1, s_1) = \frac{\alpha(r_1, s_1)}{\beta(r_1, s_1)} w_s^{(2)}(r_1, s_1) = - w_s^{(2)}(r_1, s_1). \tag{43}$$

At points symmetric to the t-axis states have the same density and opposite velocities. From the corollary to Theorem 1 this will be the case in the region above $\overline{N}'''P_1''N'''$ if this region is the transform by T_w of a region in the state plane symmetric to the line $r+s=0$, provided $w(r, s)=w(-s, -r)$ holds in this region.

THE THIRD INITIAL VALUE PROBLEM. *To construct a solution* $w^{(3)}$ *of* (16′) *meeting the symmetry condition* $w^{(3)}(r, s)=w^{(3)}(-s, -r)$ *and the initial conditions* (42), (43) *on the side* $P_1 R$ *of the secondary region.*

The solution of this initial value problem

$$w^{(3)}(r, s) = - \frac{S(-r) + S(s)}{r - s}, \tag{44}$$

is obtained by setting $R=-S(-r)$ in (17), where $S(s)$ is the function entering in the solution of the second initial value problem.

The symmetry condition is obviously fulfilled. From (33) we find $w^{(3)}(r_1, s) = w^{(2)}(r_1, s)$ and (42) follows by differentiation. Condition (43) is likewise a consequence of (33).

The subregion $P_1RP_2\overline{R}P_1$ of the secondary region in Figure 5a is termed the *tertiary region.*

The mapping by $T_{w^{(3)}}$ of the tertiary region upon the (x, t)-plane is not (1-1). If R is replaced by $-S(-r)$ in (19) one obtains

$$t_r^{(3)} = \frac{3}{2} \frac{(r-s)^2S''(-r) + 2(r-s)[2S'(-r) + S'(s)] + 6[S(-r) + S(s)]}{(r-s)^4}. \tag{45}$$

In particular on $r=-s$ (along P_1P_2)

$$t_r^{(3)} = \frac{3}{8s^4}[s^2S''(s) - 3sS'(s) + 3S(s)].$$

For $\rho_0-\rho_2$ sufficiently small $t_r^{(3)}$ is positive along P_1P_2 in view of (27), (35) and Lemmas 9 and 11. On the other hand, if we fix r in (45) and allow s to approach s_2 it appears that $t_r^{(3)}$ eventually becomes negative because of the behavior of S, S' as s tends to s_2. Hence there exists a subregion $P_1P_2MP_1$ of the tertiary region within which $t_r^{(3)}$ is positive, except along MP_2 where $t_r^{(3)}=0$.

Differentiating equations (5) partially with respect to r and s and eliminating x_{rs}, we find that t satisfies the partial differential equation $(\alpha-\beta)t_{rs}=\beta_s t_r-\alpha_r t_s$ which reduces, in the adiabatic case, to

$$(r-s)t_{rs} = \frac{1}{2}\frac{\gamma+1}{\gamma-1}(t_r - t_s).$$

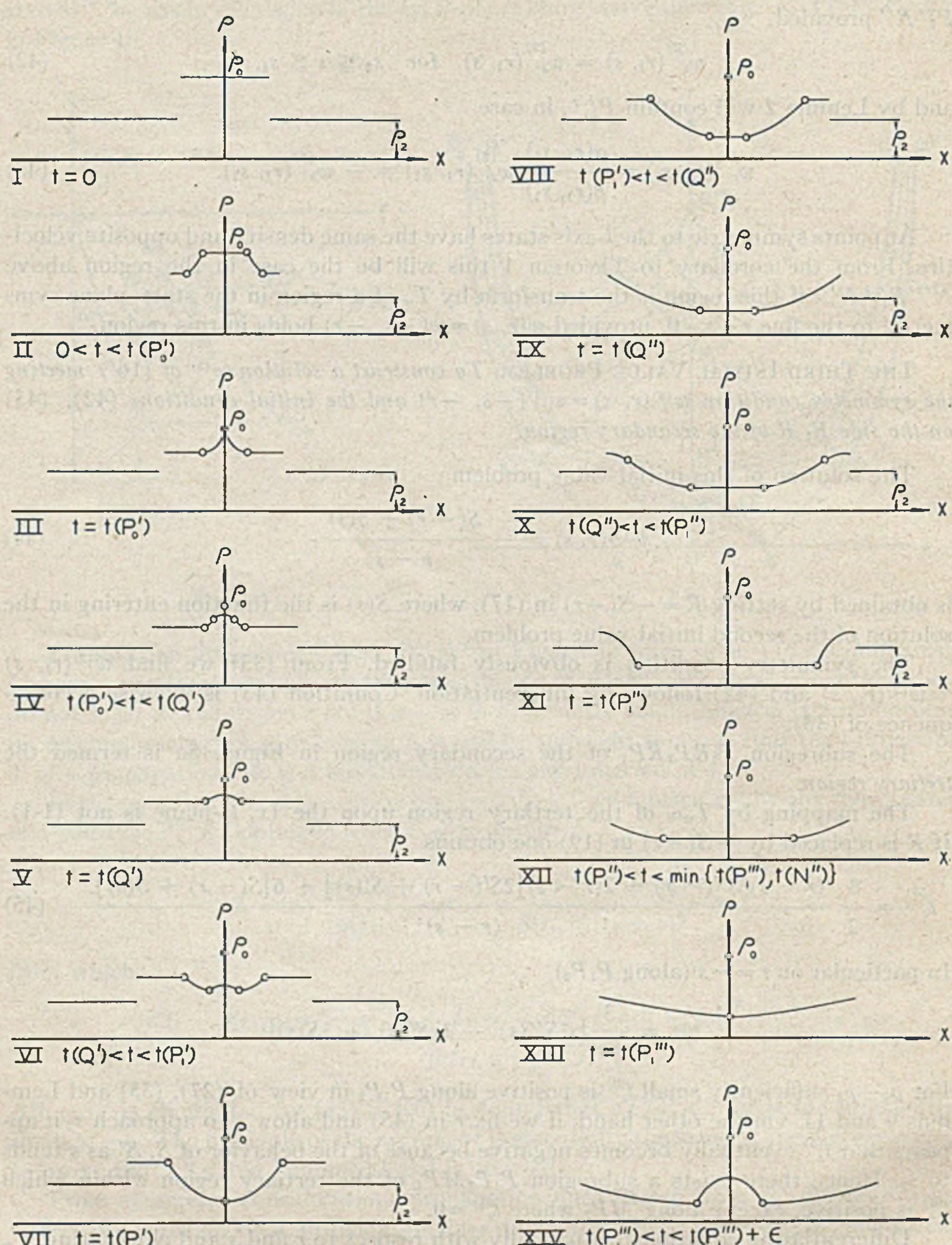

PLATE 1. Variation of density ρ with distance x for fixed time t.

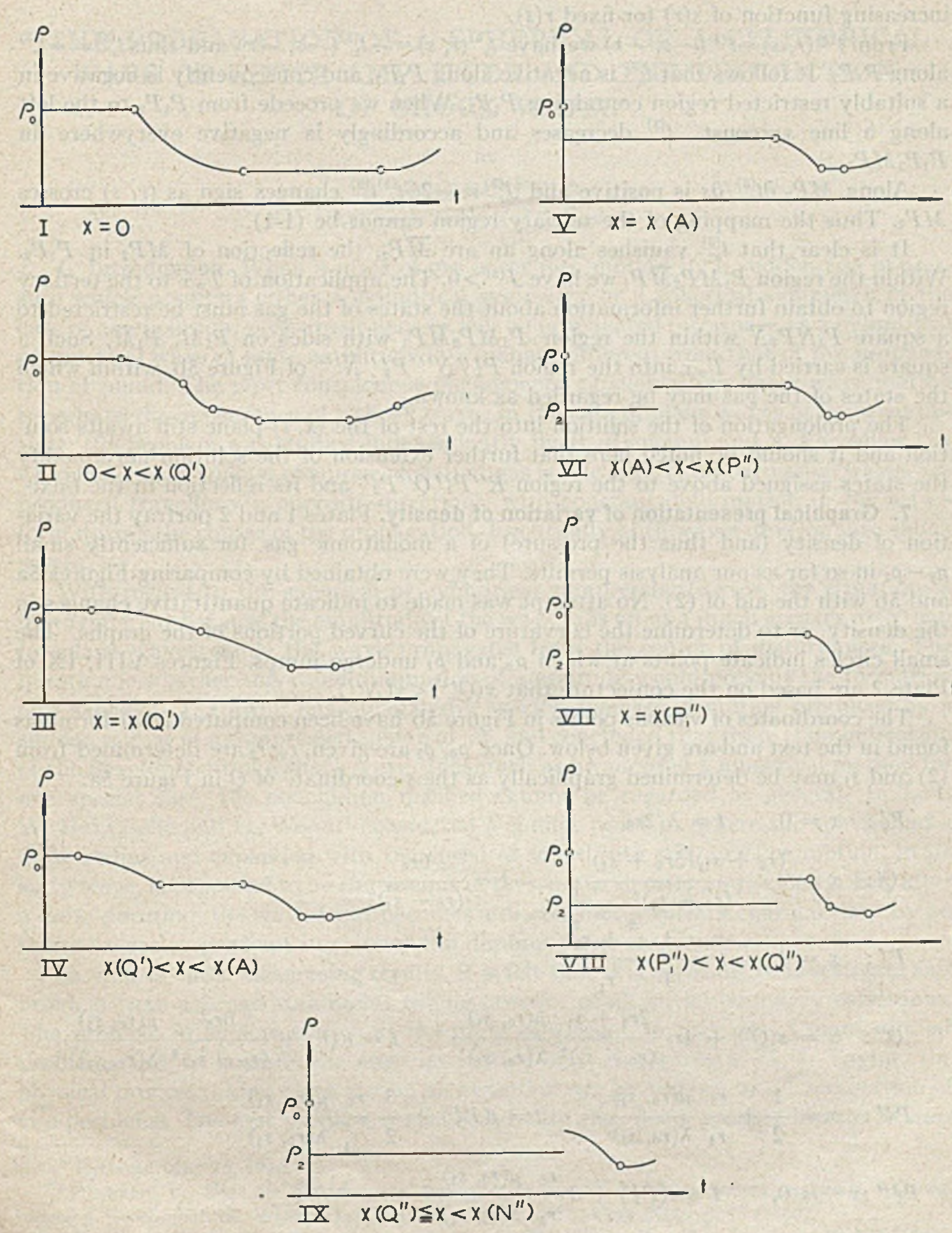

PLATE 2. Variation of density ρ with time t for fixed distance x.

It follows that in a region in which t_r is positive and t_s is negative, $t_r(t_s)$ is a monotonic increasing function of $s(r)$ for fixed $r(s)$.

From $t^{(3)}(r, s) = t^{(3)}(-s, -r)$ we have $t_s^{(3)}(r, s) = -t_r^{(3)}(-s, -r)$ and thus $t_s^{(3)} = -t_r^{(3)}$ along P_1P_2. It follows that $t_s^{(3)}$ is negative along P_1P_2, and consequently is negative in a suitably restricted region containing P_1P_2. When we procede from P_1P_2 to the left along a line $s=\text{const.}$, $t_s^{(3)}$ decreases and accordingly is negative everywhere in $P_1P_2MP_1$.

Along MP_2 $\partial t_r^{(3)}/\partial s$ is positive and $J_r^{(3)} = -2Gt_r^{(3)}t_s^{(3)}$ changes sign as (r, s) crosses MP_2. Thus the mapping of the tertiary region cannot be (1-1).

It is clear that $t_s^{(3)}$ vanishes along an arc $\overline{M}P_2$, the reflection of MP_2 in P_1P_2. Within the region $P_1MP_2\overline{M}P_1$ we have $J^{(3)} > 0$. The application of $T_{w^{(3)}}$ to the tertiary region to obtain further information about the states of the gas must be restricted to a square $P_1NP_3\overline{N}$ within the region $P_1MP_2\overline{M}P_1$ with sides on P_1M, $P_1\overline{M}$. Such a square is carried by $T_{w^{(3)}}$ into the region $P_1'''N'''P_3'''\overline{N}'''$ of Figure 5b, within which the states of the gas may be regarded as known.

The prolongation of the solution into the rest of the (x, t)-plane still awaits solution and it should be noted here that further extension of the solution may modify the states assigned above to the region $R''P_1''Q''P_2''$ and its reflection in the t-axis.

7. Graphical presentation of variation of density. Plates 1 and 2 portray the variation of density (and thus the pressure) of a monatomic gas for sufficiently small $\rho_0 - \rho_2$ in so far as our analysis permits. They were obtained by comparing Figures 5a and 5b with the aid of (2). No attempt was made to indicate quantitative changes in the density, or to determine the curvature of the curved portions of the graphs. The small circles indicate points at which ρ_x and ρ_t undergo jumps. Figures VIII, IX of Plate 2 are based on the conjecture that $x(Q'') < x(N'')$.

The coordinates of various points in Figure 5b have been computed from formulas found in the text and are given below. Once ρ_0, ρ_2 are given, r_0, r_2 are determined from (2) and s_1 may be determined graphically as the s-coordinate of Q in Figure 5a.

$$P_0':\quad x = 0, \qquad t = 3/2r_0,$$

$$Q':\quad x = \frac{(r_0 + s_1)(5r_0 + s_1)}{(r_0 - s_1)^2}, \qquad t = \frac{6r_0}{(r_0 - s_1)^2},$$

$$P_1':\quad x = 0, \qquad t = \frac{3}{4}\,\frac{r_0^2 + r_1^2}{r_1^3},$$

$$Q'':\quad x = x(Q') + 4r_0\frac{2r_0 + s_1}{(r_0 - s_1)^2}\,\frac{\mu(r_0, s_1)}{\lambda(r_0, s_1)}, \qquad t = t(Q') + \frac{6r_0}{(r_0 - s_1)^2}\,\frac{\mu(r_0, s_1)}{\lambda(r_0, s_1)},$$

$$P_1'':\quad x = \frac{1}{2} + \frac{r_0}{r_1}\,\frac{\mu(r_0, s_1)}{\lambda(r_0, s_1)}, \qquad t = t(P_1') + \frac{3}{2}\,\frac{r_0}{r_1^2}\,\frac{\mu(r_0, s_1)}{\lambda(r_0, s_1)},$$

$$P_1''':x = 0, \qquad t = t(P_1') + 3\frac{r_0}{r_1}\,\frac{\mu(r_0, s_1)}{\lambda(r_0, s_1)}\cdot$$

THE PROPAGATION OF A SPHERICAL OR A CYLINDRICAL WAVE OF FINITE AMPLITUDE AND THE PRODUCTION OF SHOCK WAVES*

BY

YUNG-HUAI KUO

California Institute of Technology

1. Introduction. When a mass of gas is set into motion by a sudden rise of pressure which possesses either a cylindrical symmetry or a spherical symmetry in the case of an explosion, pressure or density will be propagated into space as a cylindrical or spherical wave of finite amplitude in a manner different from that of the propagation of sound. The most conspicuous phenomenon of such a non-linear wave motion is perhaps the appearance of a shock wave. In the case of plane waves of finite amplitude, the problem was studied independently by B. Riemann[1] and S. Earnshaw.[2] It was shown that when a compressed slab of gas is released, two progressive waves are produced travelling in opposite directions, with constant deformation in the waveform during the course of the propagation. Eventually both waves develop into shock waves.

With regard to the spherical or cylindrical compression waves, the situation is quite different because the amplitude of the wave falls off at a much greater rate than for plane waves, while the wave propagates from the center of disturbances. The question is whether this rapid diminution of amplitude would prevent the formation of a shock. J. J. Unwin[3] has calculated a specific example of motion produced by a sudden release of a compressed sphere of air, and concluded that there is no indication of the development of a shock wave. Inasmuch as he adopted a numerical method for one special case, the concludion reached cannot be regarded as general. In fact, W. Hantzsche and H. Wendt[4] considered a similar problem, where the sphere had a finite radius and expanded with the speed of sound into still air. The motion, in its early stage, is supposed to be continuous in pressure or density and velocity. But after a finite duration, the wave-front becomes a discontinuity surface characterized by an infinite velocity gradient in spite of the diminution of amplitude.

In view of these disagreeing results, it is felt that it is desirable to investigate this problem from a broad standpoint taking account of all initial boundary conditions. The problem of explosion such as the burst of a bomb is only one of many similar problems and, to be sure, the most interesting one. According to G. I. Taylor, the physical process taking place during an explosion can be treated, as a combination of two problems. The first problem is concerned with the effects produced in the atmos-

* Received May 15, 1946.

[1] Riemann, B., *Über die Fortpflanzung ebener Luftwellen von endlicher Schwingungsweite*, Abhandlungen d. Gesellschaft der Wissenschaften zu Göttingen, Math.-Phys. Klase **8**, 43 (1860).

[2] Earnshaw, S., *On the mathematical theory of sound*, Phil. Trans. Roy. Soc. London, **150**, 133 (1860).

[3] Unwin, J. J., *The production of waves by a sudden release of a spherical distribution of compressed air in the atmosphere*. Proc. Roy. Soc. (A) **178**, 153 (1941).

[4] Hantzche, W. and Wendt, H., *Zum Verdichtungsstoss bei Zylinder- und Kugelwellen*, Jahrbuch 1940 der deutschen Luftfahrtforschung I, 536.

phere by a rapidly expanding spherical or cylindrical solid shell which compresses the surrounding air. In this case the motion of air in contact with the shell is completely prescribed by the motion of shell itself. The second problem deals with the motion produced by a compressed sphere or cylinder of air which is suddenly released. Each one of these constitutes a separate mathematical problem. To enlarge the scope of this discussion, the very meaning of the term explosion will be understood here as any process that is capable to create a pressure disturbance with spherical or cylindrical symmetry, propagating as a wave of finite amplitude.

An explosion is assumed to take place, during a short interval of time, in an infinite space which is filled only with air not abstructed by any solid bodies. Since the coefficients of viscosity and heat conduction for gases are generally very small, so long as the motion is continuous, the air may be regarded as non-viscous and non-conducting. The thermodynamic change of state of a fluid-particle along the path is then adiabatic; and if, initially, the entropy of the air is uniform throughout the space, the motion is isentropic. For the first problem stated above this condition is satisfied. Namely, at the moment the shell starts to expand, the outside air may certainly be assumed to be at the standard conditions. After the shell has started to expand, it compresses the air and, thereby, sets it into motion; but, during this process, no heat has been imparted to the air, its thermodynamic state must remain on the same adiabatic curve. In the case of a compressed sphere or cylinder of air, it is reasonable to assume that the pressure or density was built under adiabatic compression at all points. Hence as long as the motion is continuous, it will be isentropic.

The present study reveals that such a continuous and isentropic motion generally does not exist in the whole field. This type of motion breaks down when a "limiting line" appears, which would make the solution multi-valued. This would be impossible unless the motion is discontinuous. Hence, the appearance of a "limiting line" serves to indicate the necessity of presence of a shock wave in the actual motion. After the shock is formed, the Rankine-Hugoniot theory asserts that the process through which a fluid-particle has undergone by crossing the shock-front is irreversible and, consequently, the entropy increases in a discontinuous manner. The jump in entropy is not constant, however. It varies as the shock wave propagates, because the conditions at the shock change with time. As a result the motion behind such a non-uniform shock cannot be isentropic. Therefore once the "limiting line" appears, isentropic flow cannot be maintained and the resultant flow cannot be analyzed by the present method.

The mathematical condition for the appearance of a "limiting line" in the case of a spherical or cylindrical isentropic motion is that one of the two families of characteristics admits an envelope, just as in the case of a plane wave. Along this envelope the accelerations of the fluid-particles are infinite. In fact, a closer examination indicates that the motion generally must break down even before the "limiting line" is reached. It then seems that any motion of a compressible fluid has a tendency to develop a shock wave and that the effect of the "spreading" in the case of a non-linear spherical or cylindrical wave plays but a minor role.

2. Differential equations of motion. The motion under consideration is supposed to be axially or spherically symmetric, i.e., at any instant the velocity u, pressure p and density ρ depend on the time and the radial distance x only. If the effects of viscosity and of body force are neglected, the equations governing the motion are

$$u_t + uu_x + \frac{p_x}{\rho} = 0, \tag{2.1}$$

$$\rho_t + u\rho_x + \rho\left(u_x + \frac{\alpha u}{x}\right) = 0. \tag{2.2}$$

Here the subscripts denote the partial derivatives with respect to the variable indicated by the subscript; $\alpha = 1$ for a cylindrical and $\alpha = 2$ for a spherical wave. In each case, the variable x will be interpreted differently. Furthermore, it is assumed that the motion is continuous and that the effects of viscosity and heat-transfer in the fluid can be ignored. If initially constant, throughout the fluid, the entropy then remains constant. In other words, for an ideal gas the relation between the pressure and density is

$$p = K\rho^{\gamma}, \tag{2.3}$$

where γ stands for the ratio of the specific heats and K is a constant. With a set of appropriate initial conditions the mathematical problem can then be solved, at least theoretically. However, we may understand the singular behavior of such a solution and the conditions for its existence without actually solving the differential equations.

By eliminating the pressure with the aid of Eq. (2.3) and by introducing the square of the sonic speed as a variable in the place of the density, we reduce Eqs. (2.1) and (2.2) to

$$u_t + uu_x + v_x = 0, \tag{2.4}$$

$$v_t + uv_x + \beta v\left(u_x + \frac{\alpha u}{x}\right) = 0, \tag{2.5}$$

where

$$\beta v = c^2, \qquad \beta = \gamma - 1,$$

and c is the speed of sound defined by $\sqrt{\gamma(p/\rho)}$. This system of differential equations is of the hyperbolic type, the two families of real characteristics C being determined by

$$(dx - udt)^2 - \beta v dt^2 = 0, \tag{2.6}$$

where v is positive.

As it stands, this system of equations can reveal but little information concerning the behavior of the solution. To expose such properties, one has to transform the differential equations to a new coordinate-system and then study the condition under which the transformation would be valid. In the case of a steady irrotational motion, this is well-known as the hodograph method which has been effectively and successfully applied by W. Tollmien[5] and H. S. Tsien[6] in investigating the two dimensional and three dimensional isentropic motion respectively. By a slight modification, it can also be applied to the present problem. To this end, the following one-one point-transformation is introduced

[5] Tollmien, W., *Grenzlinien adiabatischer Potentialströmungen*, Z. angew. Math. Mech. **21**, 140 (1941).

[6] Tsien, H. S., *The "limiting line" in mixed subsonic and supersonic flows of compressible fluids*, N.A.C.A. Tech. Note 961 (1945).

$$u = u(t, x), \qquad v = v(t, x). \tag{2.7}$$

We have

$$u_t = \frac{x_v}{J}, \qquad u_x = -\frac{t_v}{J},$$

$$v_t = -\frac{x_u}{J}, \qquad v_x = \frac{t_u}{J},$$

provided the Jacobian $J(u, v) \equiv t_u x_v - t_v x_u \neq 0$. Equations (2.4) and (2.5) will then be transformed into

$$x_v - u t_v + t_u = 0, \tag{2.8}$$

$$x_u - u t_u + \beta v t_v - \frac{\alpha\beta uv}{x}(t_u x_v - t_v x_u) = 0. \tag{2.9}$$

This system of equations can be simplified considerably by introducing a function $\chi(u, v)$ defined by

$$x - ut = \chi_u, \qquad t = -\chi_v, \tag{2.10}$$

so that Eq. (2.8) is satisfied identically while Eq. (2.9) reduces to

$$\chi_{uu} - \beta v \chi_{vv} - \frac{\alpha\beta uv}{x}(\chi_{uu}\chi_{vv} - \chi_{uv}^2 - \chi_v \chi_{vv}) = \chi_v. \tag{2.11}$$

The corresponding characteristics Γ in the u, v-plane are determined by

$$\left(1 - \frac{\alpha\beta uv}{x}\chi_{vv}\right) dv^2 - \frac{2\alpha\beta uv}{x}\chi_{uv} du dv - \alpha\beta uv\left(\frac{1}{\alpha u} + \frac{\chi_{uu} - \chi_v}{x}\right) du^2 = 0. \tag{2.12}$$

3. Limiting line. The relationship between the characteristics C and Γ associated respectively with the differential equation in the t, x- and u, v-planes has an important bearing on the singular character of the solution and its elucidation often contributes much toward the understanding of the nature of the physical problem. For this purpose, we first transform the differential equation (2.6) by means of the following pair of relations:

$$dx = (\chi_{uu} - \chi_v - u\chi_{uv})du + (\chi_{uv} - u\chi_{vv})dv,$$
$$dt = -\chi_{uv} du - \chi_{vv} dv.$$

Substituting in Eq. (2.6) together with Eq. (2.11), we bring the equation of the characteristics C into the form

$$J\left[\left(1 - \frac{\alpha\beta uv}{x}\chi_{vv}\right) dv^2 - \frac{2\alpha\beta uv}{x}\chi_{uv} du dv - \alpha\beta uv\left(\frac{1}{\alpha u} + \frac{\chi_{uu} - \chi_v}{x}\right) du^2\right] = 0. \tag{3.1}$$

This shows that if $J \neq 0$, the characteristics C in the t, x-plane correspond to the characteristics Γ in the u, v-plane. However, circumstances may arise such that

$$J(u, v) \equiv \chi_{uu}\chi_{vv} - \chi_{uv}^2 - \chi_v \chi_{vv} = 0, \tag{3.2}$$

while

$$\left(1 - \frac{\alpha\beta uv}{x}\chi_{vv}\right)dv^2 - \frac{2\alpha\beta uv}{x}\chi_{uv}dudv - \alpha\beta uv\left(\frac{1}{\alpha u} + \frac{\chi_{uu} - \chi_v}{x}\right)du^2 \neq 0$$

and the characteristic equation (2.6) is again satisfied. This means that if a point moves along a line λ defined by Eq. (3.2), the corresponding point will describe a line l in the t, x-plane, having the same tantents as the characteristics C. It does not coincide, however, with any one of the characteristics C. This may be proved as follows.

The differential equation for the path s of a fluid-particle in the t, x-plane is

$$\left(\frac{dx}{dt}\right)_s = u. \tag{3.3}$$

The corresponding path σ in the u, v-plane is given by

$$\left(\frac{dv}{du}\right)_\sigma = -\frac{\chi_{uu} - \chi_v}{\chi_{uv}}. \tag{3.4}$$

Now the differential equation for one family of characteristics, say Γ_+, is

$$\left(\frac{dv}{du}\right)_{\Gamma_+} = -\frac{\chi_{uu} - \chi_v + \sqrt{\beta v}\,\chi_{uv}}{\chi_{uv} + \sqrt{\beta v}\,\chi_{vv}}. \tag{3.5}$$

On the other hand, the vanishing of the Jacobian, when combined with Eq. (2.11), can be written as

$$(\chi_{uv} - \sqrt{\beta v}\,\chi_{vv})(\chi_{uv} + \sqrt{\beta v}\,\chi_{vv}) = 0. \tag{3.6}$$

It is easy to see that

$$\left(\frac{dv}{du}\right)_\sigma = \left(\frac{dv}{du}\right)_{\Gamma_+}, \quad \text{if} \quad \chi_{uv} = \sqrt{\beta v}\,\chi_{vv} \tag{3.7}$$

or

$$\left(\frac{dv}{du}\right)_\sigma = \left(\frac{dv}{du}\right)_{\Gamma_-}, \quad \text{if} \quad \chi_{uv} = -\sqrt{\beta v}\,\chi_{vv}. \tag{3.8}$$

The condition under which this result holds is both necessary and sufficient. This shows that the lines λ_+ and λ_- are respectively the locus of the points of tangency of the path σ with Γ_+ and σ with Γ_-. Furthermore, the paths σ do not have an envelope and that of Γ is

$$\beta v = 0$$

which corresponds to $\rho = 0$ and is, of course, uninteresting. Hence, it cannot belong to either family of the characteristics Γ. The only alternative is that it is an envelope of one family of the characteristics C in the t, x-plane. By analogy with the steady irrotational motion it is again called "limiting line," the justification will be found in the following section.

4. The properties of the "limiting line." Being the envelope of one family of real characteristics in the t, x-plane, the "limiting line" will be entirely in the field of motion. It is, therefore, paramount to investigate the behavior of the solution along this line.

Consider first the line element of a path s of a fluid-particle at the "limiting line" l. Generally, for any line element one obtains from Eq. (2.10)

$$dx = (\chi_{uu} - u\chi_{uv} - \chi_v)du + (\chi_{uv} - u\chi_{vv})dv,$$

$$dt = -\chi_{uv}du - \chi_{vv}dv.$$

Along a path s given by $dx/dt = u$, we have

$$(\chi_{uu} - \chi_v)du + \chi_{uv}dv = 0.$$

Using this relation to eliminate dv from dx and dt and by regarding u as a parameter, we obtain the following parametric equations for the path s:

$$dx = u\,\frac{\chi_{uu}\chi_{vv} - \chi_{uv}^2 - \chi_v\chi_{vv}}{\chi_{uv}}\,du, \tag{4.1}$$

$$dt = \frac{\chi_{uu}\chi_{vv} - \chi_{uv}^2 - \chi_v\chi_{vv}}{\chi_{uv}}\,du. \tag{4.2}$$

According to our previous findings, $J = 0$ yields two lines λ_+ and λ_-, each of which associates with only one group of characteristics Γ in the u, v-plane. This shows that on the "limiting line" dx and dt both become differentials of higher order and will change sign on crossing the line λ. This agrees, of course, with the cuspidal nature of the singularity.

Dividing both sides by dx and dt respectively, we obtain the following expressions for the derivatives u_x and u_t along s:

$$(u_x)_s = \frac{\chi_{uv}}{u(\chi_{uu}\chi_{vv} - \chi_{uv}^2 - \chi_v\chi_{vv})}, \tag{4.3}$$

$$(u_t)_s = \frac{\chi_{uv}}{\chi_{uu}\chi_{vv} - \chi_{uv}^2 - \chi_v\chi_{vv}}. \tag{4.4}$$

Thus on the "limiting line" the acceleration of a fluid-particle becomes infinite as χ_{uv} is finite there. This implies also an infinite pressure gradient [see Eq. (2.1)].

The physical state to which $J(u, v) = 0$ corresponds can be readily deduced. It can be summarized in the statement that if the Jacobian vanishes, then the motion in the immediate neighborhood of the line $J = 0$ is a compressive one. To prove this, let us consider the ratios v_t/u_x, u_t/v_x, u_t/u_x and v_t/v_x which, according to the relations obtined in Section 2, equal

$$\frac{v_t}{u_x} = -\beta v - u\frac{\chi_{uv}}{\chi_{vv}} - \frac{\alpha\beta uv}{x}\cdot\frac{J}{\chi_{vv}}, \qquad \frac{u_t}{u_x} = -u + \frac{\chi_{uv}}{\chi_{vv}},$$

$$\frac{u_t}{v_x} = u\frac{\chi_{vv}}{\chi_{uv}} - 1, \qquad \frac{v_t}{v_x} = -u + \beta v\frac{\chi_{vv}}{\chi_{uv}} + \frac{\alpha\beta uv}{x}\,\frac{J}{\chi_{uv}}.$$

In the u, v-plane, the expressions on the right-hand side are everywhere continuous. At the line λ_+ corresponding to $\chi_{uv} = \sqrt{\beta v}\chi_{vv}$, they become

$$\frac{v_t}{u_x} = -c^2\left(1 - \frac{u}{c}\right) < 0, \qquad \frac{u_t}{u_x} = c\left(1 - \frac{u}{c}\right) > 0,$$

$$\frac{u_t}{v_x} = -\left(1 - \frac{u}{c}\right) < 0, \qquad \frac{v_t}{v_x} = c\left(1 - \frac{u}{c}\right) > 0.$$

By continuity, the relative signs of the differential quotients hold in the neighborhood of the "limiting line." Thus, we conclude that either $v_t>0$, $v_x>0$ and $u_t<0$, $u_x<0$ or $v_t<0$, $v_x<0$ and $u_t>0$, $u_x>0$. The first case is exactly the condition for a compressive motion. Whereas the second case may either correspond to a rarefaction or to a change of sign of the Jacobian $J(u, v)$. As the rarefaction does not conform to the geometric properties of $J=0$, the second case corresponds to the second branch of the solution and hence can be disregarded.

5. Lost solution. In the previous sections, we assume that the Jacobian $J(u, v)$ does not vanish. Thus the one-to-one correspondence between the t, x- and u, v-planes is assured and the condition $J=0$ is restricted to the singular line l. In a special case the Jacobian may vanish identically, however. This vanishing of the Jacobian establishes a relation between v and u in the u, v-plane and, as a result, yields a class of solution not contained in the transformation (2.7). To study this form of solution, let us first set

$$v = v(u). \tag{5.1}$$

The differential equations (2.4) and (2.5) can then be rewritten as

$$u_t + \left(u + \frac{dv}{du}\right)u_x = 0, \tag{5.2}$$

$$u_t\frac{dv}{du} + \left(u\frac{dv}{du} + \beta v\right)u_x = -\frac{\alpha\beta uv}{x}\cdot \tag{5.3}$$

This type of solution has been discussed by K. Bechert[7] whose main result was as follows. By eliminating x and t the system of Eqs. (5.2) and (5.3) can be reduced to a second order non-linear total differential equation, based on the existence of a linear relation between t and x. By a slightly different procedure it can be shown that instead of a second order differential equation one can obtain a first order one of Abel's type being amenable to numerical integration. The main feature of the solution, however, can be discussed in the following manner.

Along $u=\text{const.}$, i.e., along

$$du = u_x dx + u_t dt = 0,$$

the slope of the curve $u=\text{const.}$ equals

$$\left(\frac{dx}{dt}\right)_u = -\frac{u_t}{u_x} = u + \frac{dv}{du}, \tag{5.4}$$

on account of Eq. (5.2). Since dv/du is a function of u alone, on $u=\text{const.}$ $(dv/du)_u$ is

[7] Bechert, K., *Über die Ausbreitung von Zylinder- und Kugelwellen in reibungsfreien Gasen und Flüssigkeiten*, Ann. Phys. (5) 39, 169 (1941).

constant. Therefore, the curve $u=$const. is a straight line in the t, x-plane. In conformity to the assumption (5.1), there exists a parameter ξ defined by

$$\xi = \frac{x}{c_0(t + t_0)}, \tag{5.5}$$

where c_0 is the speed of sound at $u=0$, and t_0 a suitable constant. It is clear that $\xi=$const. corresponds to $u=$const. In other words, both v and u may be regarded as functions of ξ.

If the determinant $v'^2-\beta v\neq 0$, u_x and u_t can be expressed in terms of u. We have

$$u_x = \frac{\alpha\beta uv}{x}\frac{1}{v'^2 - \beta v}, \tag{5.6}$$

$$u_t = -\frac{\alpha\beta uv}{x}\frac{u + v'}{v'^2 - \beta v}, \tag{5.7}$$

where the prime denotes the total differentiation with respect to u. Like in the general case, here again the solution possesses a singular line on which the partial derivatives generally become infinite. Its other properties will be studied presently. From Eq. (5.4) it is found that

$$\left(\frac{dx}{dt}\right)_u = u + v',$$

while the characteristics are

$$\left(\frac{dx}{dt}\right)_C = u \pm \sqrt{\beta v}.$$

On the other hand, where the singular line λ, i.e. the line

$$v'^2 - \beta v = 0, \tag{5.8}$$

intersects the integral-curve $v(u)$, we have

$$\left(\frac{dx}{dt}\right)_u = u \pm \sqrt{\beta v} = \left(\frac{dx}{dt}\right)_C. \tag{5.9}$$

This shows that at the singular point of the solution $v(u)$, the $u=$const. line becomes the envelope of one family of characteristics C. Hence the envelope is a straight line. Furthermore, according to Eqs. (4.1) and (4.2) the parametric equations of the path s are

$$dx = -\frac{x}{\alpha\beta vv'}(v' + \sqrt{\beta v})(v' - \sqrt{\beta v})du, \tag{5.10}$$

$$dt = -\frac{x}{\alpha\beta uvv'}(v' + \sqrt{\beta v})(v' - \sqrt{\beta v})du. \tag{5.11}$$

Since each factor on the right-hand side corresponds to a group of the characteristics C, on crossing the line λ, where this factor vanishes, the elements dx and dt change their signs. This proves that the line l, the image of λ, possesses all the characteristics of a "limiting line."

It is interesting to note the difference between plane and spherical waves. In the former case, Eq. (5.8) would be satisfied identically. This lets the lines $u=$const. degenerate into the characteristics. Indeed, it is also possible for one family of the characteristics which are straight lines to have an envelope; the differential quotients u_x, u_t are finite, however. Consequently, we have no "limiting line," in the strict sense. This does not mean, of course, that the solution is regular. As a matter of fact, the solution already becomes many-valued before this line is reached.

6. Lost solution: a special problem. From the foregoing conclusions, a compressive spherical or cylindrical wave always becomes indeterminate when a singular line is reached. As an illustration the following special problem is considered.

Suppose there is a divergent spherical or cylindrical wave propagating with velocity c_0 into still air. On the wave-front, where the motion agrees with the outside conditions, the state-variables ρ, p become equal to those of the still air and the velocity is zero. The path of the wave-front is then described by

$$x = c_0(t + t_0). \tag{6.1}$$

The mathematical problem can thus be formulated in the following way:

$$\left.\begin{aligned} u &= 0, \quad \text{when} \quad x \geqq c_0(t + t_0), \\ u &\neq 0, \quad \text{when} \quad x < c_0(t + t_0). \end{aligned}\right\} \tag{6.2}$$

A particularly simple case will be the one where both the pressure and the velocity are propagated with constant speed. In other words, these quantities depend only on a common parameter.

To simplify the amount of mathematical work involved, the differential equations (2.4) and (2.5) will be put into the following equivalent form:

$$(c^2 - \phi_x^2)\phi_{xx} - 2\phi_x\phi_{xt} - \phi_{tt} + \frac{2c^2\phi_x}{x} = 0 \tag{6.3}$$

by introducing a potential-function $\phi(t, x)$:

$$u = \phi_x, \qquad c^2 - c_0^2 + \frac{\beta}{2}\phi_x^2 = -\beta\phi_t. \tag{6.4}$$

In the case of a lost solution, there exists a parameter ξ defined by (5.5) such that $\xi=1$ corresponds to the initial curve (6.1). Then,

$$\phi(t, x) = c_0^2(t + t_0)f(\xi) \tag{6.5}$$

and hence

$$u(t, x) = c_0 f'(\xi), \tag{6.6}$$

$$c^2 = c_0^2\left[1 - \frac{\beta}{2}f'^2 - \beta(f - \xi f')\right], \tag{6.7}$$

where the prime indicates the total differentiation with respect to ξ, and the function $f(\xi)$ satisfies

$$[c^2 - c_0^2(f' - \xi)^2]\xi f'' + 2c_0^2 f' = 0 \tag{6.8}$$

subject to the initial conditions

$$f(1) = 0, \qquad f'(1) = 0. \tag{6.9}$$

The first condition, namely $f(1)=0$, is necessary to make $c=c_0$ on $\xi=1$. When the conditions (6.9) are substituted in Eq. (6.8), it appears that $f''(1)$ is arbitrary. We need not be alarmed by this situation, but recall that in this particular type of initial value problem, the "support" is a characteristic. Physically, this means that the initial conditions prescribed in this manner do not "know" the internal structure of the motion, because they propagate ahead with larger speed. It is only natural, then, that such an arbitrariness should arise which enables us to fit properly the physical conditions specified. This arbitrariness is only a partial one, however, since for a compressive motion the sign of $f''(1)$ is necessarily negative; for on $\xi=1$

$$(\rho_t)_1 + \rho_0(u_x)_1 = 0,$$

according to Eq. (2.2). In a compressive motion $(\rho_t)_1>0$, it follows that

$$(u_x)_1 = \frac{c_0}{x} f''(1) < 0. \tag{6.10}$$

Thus, for any compressive motion the absolute value of $f''(1)$ is determined in consistence with the physical process.

The differential equation (6.8) which determines the interior motion of a mass of air, has two singular points in the ξ, f-plane given by the vanishing of the coefficient of $f''(\xi)$. The geometrical interpretation is evident, when (6.8) is written as

$$(c + u - c_0\xi)(c - u + c_0\xi) = -\left[\left(\frac{dx}{dt}\right)_{C_+} - \left(\frac{dx}{dt}\right)_\xi\right]\left[\left(\frac{dx}{dt}\right)_{C_-} - \left(\frac{dx}{dt}\right)_\xi\right] \tag{6.11}$$

that is, when one family of characteristics become tangent to a line $\xi=$const., an infinite curvature would occur if u is finite there. According to what has been said in the last section, this characterizes the "limiting line" of the solution.

Let us push the discussion a step further. For this purpose only the first order terms need be retained. Taking β as a small parameter, one has accordingly

$$f(\xi) = f_0(\xi) + \beta f_1(\xi) + \cdots. \tag{6.12}$$

Substituting in Eq. (6.8) we obtain

$$[1 - (f_0' - \xi)^2]\xi f_0'' + 2f_0' = 0. \tag{6.13}$$

This equation is free from f_0; letting $w=f'$ we find

$$\frac{dw}{d\xi} = \frac{w}{\xi}\,\frac{2}{(w - \xi)^2 - 1}, \qquad 0 \leqq \xi \leqq 1. \tag{6.14}$$

Aside from the two singular lines

$$w = \xi + 1, \tag{6.15}$$ $$w = \xi - 1, \tag{6.16}$$

where the slope of w is infinite, there are two additional singularities (1, 0) and (0, 0) where the slope is indeterminate. The point (1, 0) acts as a sort of nodal point which makes the initial condition insufficient. The point (0, 0) is a saddle point as locally the equation behaves like

$$\frac{dw}{d\xi} = -\frac{2w}{\xi}, \tag{6.17}$$

which form is obtained by neglecting $(w-\xi)^2$ as compared with 1.

The situation can now be summarized. The integral curve starting from (1, 0) rises as ξ decreases and eventually intersects with the line (6.15) where it will have a vertical tangent at $\xi<1$. After it crosses this line its slope changes sign. This causes the curve to bend backward again. Thus, ξ is seen to assume a minimum value. Owing to the fact that the origin is a saddle point, no integral curve could possibly cross the line $\xi=0$. This fact makes the continuation of the solution as far as $\xi=0$ impossible.

7. Continuation of the solution. The results obtained in the previous sections show that, in the case of the propagation of a spherical or cylindrical wave, a continuous solution does not exist throughout the domain considered and can be constructed, at most, as far as a singular line l in the t, x-plane from a suitably chosen initial data. The line l thus acts as a sort of "frontier" into which no solution can enter and at which the solution is turned back as a second branch. The domain then is doubly covered. Physically, this is impossible and hence must be rejected as a solution. The question is: is it possible to connect it with a different solution beyond this line?

First, consider the line λ as a "support" with a given set of initial data and then solve the initial value problem[8] for a Monge-Ampère equation. Regarding λ as a parameter, we have along the line λ

$$\frac{d}{d\lambda}\chi_u = \chi_{uu}\frac{du}{d\lambda} + \chi_{uv}\frac{dv}{d\lambda}, \tag{7.1}$$

$$\frac{d}{d\lambda}\chi_v = \chi_{uv}\frac{du}{d\lambda} + \chi_{vv}\frac{dv}{d\lambda}, \tag{7.2}$$

and hence

$$(\chi_{uu}\chi_{vv} - \chi_{uv}^2)\frac{du}{d\lambda} = \chi_{vv}\frac{d}{d\lambda}\chi_u - \chi_{uv}\frac{d}{d\lambda}\chi_v.$$

Substituting this into Eq. (2.11) we obtain a linear relation between the partial derivatives:

$$\chi_{uu} + \left[\frac{\alpha\beta uv}{x}\left\{\frac{d\chi_u/d\lambda}{du/d\lambda} - \chi_v\right\} - \beta v\right]\chi_{vv} - \frac{\alpha\beta uv}{x}\,\frac{d\chi_v/d\lambda}{du/d\lambda}\chi_{uv} = \chi_v. \tag{7.3}$$

Since λ is not a characteristic, Eqs. (7.1), (7.2) and (7.3) are sufficient for a unique determination of χ_{uu}, χ_{uv} and χ_{vv}; and consequently a unique integral surface. The uniqueness of the solution is sufficient to show that the solution, when transformed

[8] Courant, R. and Hilbert, D., *Methoden der math. Physik*, vol. 2, J. Springer, Berlin, 1937, p. 344.

back to the t, x-plane, will correspond to the very one that doubles back at the "limiting line." A continuous solution is thus out of the question.

The alternative procedure would be to continue it by joining it smoothly at the line λ to the lost solution. This is also impossible. Indeed, if this were possible, the line λ would have to coincide with the integral curve $v(u)$ in order to provide a continuous solution. This is contradictory, because it is easy to show that the line λ does not satisfy the differential equation for $v(u)$.

The other possibility which remains to be investigated is to identify the "limiting line" as a shock wave so as to construct a discontinuous solution. This would require the continued solution to satisfy the shock conditions. Since, in general, the "limiting line" l is curved, as a result there would be a non-uniform shock wave in the motion, for which both the speed and the strength are no longer constant and therefore the entropy would be constantly changing across the shock. This very fact makes the original assumption untenable. Hence to continue discontinuously a solution with entropy constant everywhere is also impossible.

The problem might be solved, however, if the original hypothesis of isentropic motion is abandoned. To include the possibility that a shock wave may exist within the motion, the continued solution must satisfy the following more general set of equations:

$$u_t + uu_x + \frac{p_x}{\rho} = 0, \tag{7.4}$$

$$\rho_t + u\rho_x + \rho\left(u_x + \frac{\alpha u}{x}\right) = 0, \tag{7.5}$$

$$(p\rho^{-\gamma})_t + u(p\rho^{-\gamma})_x = 0. \tag{7.6}$$

The task then is to construct a solution which should satisfy both the initial and the shock conditions in a region bounded by the initial curve, the shock line and a characteristic drawn to the initial curve through the point where the envelope first appears. The shock line, however, is not given, it should be chosen in such a way that it yields a solution fulfilling all the prescribed conditions. The mathematical problem thus turns out to be extremely difficult.

The author wishes to thank Dr. H. S. Tsien for his invaluable discussions and criticism.

ON PROJECTILES OF MINIMUM WAVE DRAG*

BY

WILLIAM R. SEARS

*Cornell University***

1. Introduction. The wave resistance of slender bodies of revolution in symmetrical supersonic flow was calculated approximately by von Kármán,[1] by means of a distribution of singularities along the axis of the projectile. The individual singularity is characterized by a potential of the form $\phi_i(x, r) = \{(x-\xi_i)^2 - \alpha^2 r^2\}^{-1/2}$, where x, r are cylindrical coordinates, x being measured downstream from the nose of the projectile and r radially from the axis, ξ_i is the value of x corresponding to the singularity, α is the cotangent of the Mach angle of the undisturbed flow, so that

$$\alpha = \sqrt{(U/a)^2 - 1},$$

U and a being the stream velocity and the velocity of sound in the undisturbed flow. It will readily be verified that $\phi_i(x, r)$ is a solution of the linearized potential equation for supersonic flow with axial symmetry

$$\left(\frac{U^2}{a^2} - 1\right)\frac{\partial^2\phi}{\partial x^2} = \frac{\partial^2\phi}{\partial r^2} + \frac{1}{r}\frac{\partial\phi}{\partial r}\,. \tag{1}$$

Von Kármán calculated the wave resistance by integrating the transport of momentum across a cylindrical surface enclosing the body. In his approximation, the integral is independent of r and can be evaluated in the limit $r \to 0$. The result is†

$$R = -\pi\rho \int_0^\infty \int_0^\infty f'(x)f'(\xi) \log |x - \xi| \, dx d\xi, \tag{2}$$

where R is the wave resistance and $f(x)$ is the function specifying the distribution of singularities along the x axis. For bodies of finite length l, $f(x)$ is found to be indentically zero for $x > l$; hence both integrals in (2) can be replaced by integrals from 0 to l.

For slender bodies, von Kármán showed that approximately

$$f(x) = \frac{U}{2\pi}\frac{dS}{dx}, \tag{3}$$

where S is the cross-sectional area of the body.

In the present paper we shall amplify the analogy, already mentioned by von Kármán, between the wave resistance of a slender projectile and the induced drag of a wing. It will be shown that this analogy suggests a useful form for the calculation of

* Received June 11, 1946.

** This work was undertaken while the author was employed by Northrop Aircraft, Inc.

[1] Th. de Kármán, *The problem of resistance in compressible fluids*, Atti del V Convegno della "Fondazione Alessandro Volta," Rome, 1935, pp. 222–276.

† Von Kármán[1], Eq. (9.12). It might be mentioned that this formula is most easily obtained from Eq. (9.11) of the same reference by first integrating by parts with respect to x in order to obtain a form symmetrical in x and ξ; it will then be found that a double integral carried over half the first quadrant of an x, ξ plane can be identified with half the same integral carried over the entire quadrant.

the wave drag. The properties of projectiles of minimum wave drag for given length and volume, and for given length and caliber, will then be investigated.

2. The induced-drag analogy. Formula (2) for the wave drag can be written in the form

$$R = -\pi\rho \int_0^l f'(x)F(x)dx \tag{2'}$$

or, after integration by parts, assuming $f(0) = f(l) = 0$; i.e. that the body has sharp points at front and rear,

$$R = \pi\rho \int_0^l f(x)F'(x)dx, \tag{2''}$$

where

$$F(x) = \int_0^l f'(\xi) \log | x - \xi | d\xi. \tag{4}$$

In the form (2″), von Kármán's analogy between the wave resistance and the induced drag of a finite wing in the Prandtl lifting-line theory[2] is evident: $f(x)$ is proportional to the circulation distribution over the span of the wing, $F'(x)$ is the corresponding downwash distribution, and R is the induced drag.

It is also useful to put (4) in another form, sometimes more convenient for calculation. Let us introduce the coordinates θ and ϑ defined by

$$\begin{aligned} x &= \frac{l}{2}(1 + \cos\theta), \qquad 0 \leqq \theta, \\ \xi &= \frac{l}{2}(1 + \cos\vartheta), \qquad \vartheta \leqq \pi. \end{aligned} \tag{5}$$

The expression for $F(x)$ then becomes

$$F(x) = \frac{l}{2}\int_0^\pi f'(\xi) \log | \cos\theta - \cos\vartheta | \sin\vartheta d\vartheta, \tag{6}$$

provided that $\int_0^l f(\xi)d\xi = 0$, as is always the case for closed bodies, in accordance with (3). Now the definitions in (5) can be taken to cover the range $-\pi \leqq \theta \leqq \pi$, $-\pi \leqq \vartheta \leqq \pi$, and $f'(\xi)$ can arbitrarily be defined to be an odd function of ϑ. Then (6) can easily be put into the form

$$F(x) = \frac{l}{2}\int_{-\pi}^\pi f'(\xi) \log \left| \sin \frac{\theta - \vartheta}{2} \right| \sin\vartheta d\vartheta$$

or, after integration by parts,

$$F(x) = -\frac{1}{2}\int_{-\pi}^\pi f(\xi) \cot \frac{\theta - \vartheta}{2} d\vartheta. \tag{7}$$

---

[2] L. Prandtl, *Tragflügeltheorie I*, from *Vier Abhandlungen zur Hydrodynamik and Aerodynamik*, Göttingen, 1927, pp. 9–35.

The induced-drag analogy pointed out above suggests that $f(\xi)$ be expanded in a sine series; this is the usual technique employed in the Prandtl wing theory:[3]

$$f\left(\frac{x}{\xi}\right) = \frac{lU}{2} \sum_{1}^{N} b_n \sin\left(n \frac{\theta}{\vartheta}\right). \tag{8}$$

Substituting in (7) and (2′), we obtain the following expression for the wave drag:

$$R = \frac{\pi^3}{4} \frac{\rho U^2}{2} l^2 \sum_{1}^{N} n b_n^2 \tag{9}$$

—again analogous to a well-known expression for the induced drag of a wing.[3]

3. Minimum wave drag for given volume and length. The expression for the cross-sectional area S corresponding to (8), in the approximation represented by (3), is

$$S = \frac{\pi l^2}{4}\left\{[\pi - \theta + \tfrac{1}{2}\sin 2\theta]b_1 - \sum_{2}^{N} b_n\left[\frac{\sin(n-1)\theta}{n-1} - \frac{\sin(n+1)\theta}{n+1}\right]\right\}. \tag{10}$$

It is clear that for closed projectiles, pointed front and rear, b_1 must vanish.* Also, the total volume occupied by the projectile is

$$\text{Vol.} = \int_0^l S dx = \pi^2\left(\frac{l}{2}\right)^3 (b_1 - \tfrac{1}{2}b_2) \tag{11}$$

or, for closed pointed bodies,

$$\text{Vol.} = -\tfrac{1}{2}\pi^2\left(\frac{l}{2}\right)^3 b_2. \tag{12}$$

Hence, for given length and volume, the minimum wave resistance is obtained when only b_2 is different from zero. The geometry of this body is given by

$$S = -\frac{\pi l^2 b_2}{4}(\sin\theta - \tfrac{1}{3}\sin 3\theta) = -\frac{\pi l^2 b_2}{3}\sin^3\theta \tag{13}$$

and its wave resistance is

$$\begin{aligned} R &= \frac{\pi^3}{2}\frac{\rho U^2}{2} l^2 b_2^2 \\ &= \frac{9}{8}\pi^2 \frac{\rho U^2}{2} S_{\max}\left(\frac{d_{\max}}{l}\right)^2 \\ &= \frac{8}{\pi}\frac{\rho U^2}{2}\left(\frac{l}{2}\right)^2\left[\frac{\text{Vol.}}{(l/2)^3}\right]^2. \end{aligned} \tag{14}$$

This is eight times the wave drag of von Kármán's ogive* of equal length and volume, or about 11.1 times that of von Kármán's ogive of equal length and caliber. (It

[3] Th. von Kármán and J. M. Burgers, *Aerodynamic Theory*, edited by W. F. Durand, vol. 2, J. Springer, Berlin, 1934, pp. 172–175.

* If $b_1 \neq 0$ while $b_2 = b_3 = \cdots = 0$, the ogive considered by von Kármán[1] is obtained. Its maximum cross-sectional area is $\pi^2 l^2 b_1/4$ and occurs at its stern, $x=l$. According to (9), its wave drag is $(\pi^3/4)(\rho U^2/2)l^2 b_1^2$ or $(\rho U^2/2)S_{\max}(d_{\max}/l)^2$, where $S_{\max}$ is its maximum cross-sectional area and $d_{\max}$ is its maximum diameter, or caliber.

should be mentioned that this comparison may be misleading in view of the fact that von Kármán's ogive has a blunt stern, so that its wave drag certainly does not represent its entire resistance, even in the absence of skin friction. Nevertheless, the wave resistance of that ogive may be taken as a convenient reference.)

The shape of the forward half of the symmetrical projectile represented in (13) is drawn in Fig. 1, for the case $l = 4d_{max}$. For comparison, there is also shown the shape of von Kármán's ogive having the same caliber and *one-half the length*.

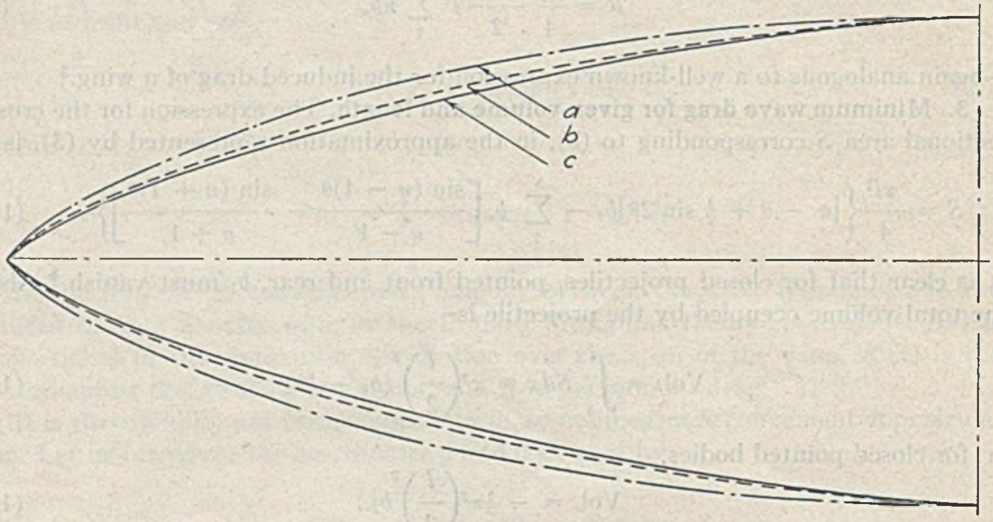

FIG. 1. Profiles of various projectiles of minimum wave drag: (a) volume and length given, (b) caliber and length given, (c) von Kármán's ogive of equal caliber and one-half the length. (Projectiles (a) and (b) are symmetrical fore-and-aft.)

4. Minimum wave drag for given caliber and length. To attack the problem of the body shape for minimum wave drag, caliber and length specified, we return to the expression for the wave drag given in (2′) and (4) and employ the methods of variation calculus. By virtue of the symmetry with respect to x and ξ, the variation of the resistance with varying body form assumes a simple form; viz.,

$$\delta R = -\pi\rho\left\{\int_0^l \delta f'(x)\int_0^l f'(\xi)\log|x-\xi|\,d\xi dx + \int_0^l f'(x)\int_0^l \delta f'(\xi)\log|x-\xi|\,d\xi dx\right\}$$

$$= -2\pi\rho\int_0^l \delta f'(x)F(x)dx. \tag{15}$$

In this section we shall provide for the possibility [excluded in obtaining (2″)] that dS/dx, and therefore $f(x)$, is discontinuous at the station where the maximum diameter occurs, $x = m$. Hence, integrating by parts in (15), and again assuming sharp points at bow and stern, we write

$$\delta R = -2\pi\rho\left\{F(m)\delta[f(x)]_m - \int_0^l \delta f(x)F'(x)dx\right\}$$

$$= -2\pi\rho\left\{F(m)\delta[f(x)]_m + \frac{U}{2\pi}\int_0^l \delta S(x)F''(x)dx\right\}, \tag{16}$$

where $[f(x)]_m$ denotes the value of the discontinuity in $f(x)$ at $x=m$, and the area function $S(x)$ has been assumed to be continuous.

In the form (16) it is clear that the shape of the part of the body forward of the maximum section at $x=m$ can be held fixed while $S(x)$ is varied over the rear part to achieve a minimum of R; then the rear shape can be fixed in this minimum-drag configuration while $S(x)$ is varied in front to minimize R; the result will be the minimum-drag shape for given maximum cross section at $x=m$. We shall also assume that the discontinuity of slope represented by $[f(x)]_m$ is not varied in the process; it will appear later that this is valid. The minimum-drag condition $\delta R=0$ is then obtained when

$$\left.\begin{aligned} F''(x) &= 0 \\ F(x) &= c_1x + c_2 \end{aligned}\right\} 0 \leqq x \leqq m,$$

$$\left.\begin{aligned} F''(x) &= 0 \\ F(x) &= c_3x + c_4 \end{aligned}\right\} m \leqq x \leqq l. \tag{17}$$

The analogy with the induced drag of a wing is again useful. The analogous problem is the following: to determine the spanwise circulation distribution $f(x)$ so as to obtain minimum induced drag, it being required that the total lift be zero, but that the lift carried on one side of a station $x=m$ have a given value, equal and opposite to that carried on the other side of that station. The result obtained in (17) states that the condition of minimum drag results when the downwash $F'(x)$ is constant in each of the two parts of the wing.

Fortunately, investigations have been made[4,5] of the behavior of the circulation distribution near a point on a lifting line where the downwash in discontinuous. It is found that the circulation function is continuous but has a vertical tangent and discontinuous curvature at such a point. Applying this result to our projectile problem, we can conclude that $f(x)$ will exhibit a singularity of this type at $x=m$. Moreover, since $F(x)$ can be interpreted as the downwash corresponding to the circulation distribution $S(x)$, we conclude that $F(x)$ cannot be discontinuous at $x=m$ if we exclude singularities of this type from the shape function $S(x)$. Accordingly, we write

$$(c_1 - c_3)m = c_4 - c_2. \tag{18}$$

The resistance of the minimum-wave-drag body is easily calculated from (2′); it is

---

[4] A. Betz and E. Petersohn, *Zur Theorie der Querruder*, Z. angew. Math. Mech. 8, 253–257 (1928); also Nat. Advis. Com. for Aeron. Tech. Memo. No. 542 (1929).

[5] H. Multhopp, *Die Berechnung der Auftriebsverteilung von Tragflügeln*, Luftfahrtforschung 15, 153–169 (1938).

$$R = -\pi\rho\left\{\frac{U}{2\pi}(c_3 - c_1)S_{\max} + (c_1 m + c_2)\left[f(x)\right]_m\right\}$$

$$= \rho\frac{U}{2}(c_1 - c_3)S_{\max}, \tag{19}$$

where $S_{\max}$ is the cross-sectional area at $x=m$.

The form of $f(x)$ corresponding to (17), and subsequently the shape of the body, can be determined by inverting (7) by means of the so-called "Reciprocity Theorem":[6]

$$f(x) = \frac{1}{2\pi^2}\int_0^{2\pi} F(\xi)\cot\frac{\theta - \vartheta}{2}\,d\vartheta. \tag{20}$$

The quadratures involved are rather tedious, but can be carried out. The result is

$$f(x) = \frac{1}{2\pi^2}\left\{(c_3 - c_1)(x - m)\log\frac{1-\cos(\theta+\mu)}{1-\cos(\theta-\mu)} + \left[(c_3 - c_1)\mu + \pi c_1\right]l\sin\theta\right\}, \tag{21}$$

where $m=(l/2)(1+\cos\mu)$. It can quickly be verified that this function has the type of singularity at $x=m$ that was predicted by the wing analogy.

This expression can be integrated again to evaluate the constants c_1 and c_3 and then to determine the function $S(x)$. By integrating from $x=0$ to $x=l$ it is determined that $(c_3-c_1)(\mu-\frac{1}{2}\sin 2\mu)+\pi c_1=0$. Finally, by carrying out the lengthy quadratures necessary to apply the condition $(U/2\pi)S_{\max}=\int_0^m f(x)dx$, it is found that

$$c_1 = 4US_{\max}\frac{\mu - \frac{1}{2}\sin 2\mu}{l^2\sin^4\mu}. \tag{22}$$

The wave resistance (19) then assumes the form

$$R = \frac{\rho U^2}{2}S_{\max}\left(\frac{d_{\max}}{l}\right)^2\frac{\pi^2}{\sin^4\mu}$$

or (23)

$$R = \frac{\rho U^2}{2}S_{\max}\left(\frac{d_{\max}}{l}\right)^2\left\{2\frac{m}{l/2} - \left(\frac{m}{l/2}\right)^2\right\}^{-2}.$$

Thus the wave drag varies symmetrically about $m=l/2$ or $\mu=\pi/2$, and is least if the maximum cross section is located at mid-length—i.e., for a symmetrical projectile. The wave drag of this projectile is π^2 times as great as that of von Kármán's ogive of equal length and caliber. Its shape is indicated in Figure 1.

5. Concluding remarks. A somewhat similar analysis of projectile shapes for minimum wave resistance has been made by Haack,[7] who considered only symmetrical projectiles. The results obtained here are in agreement with Haack's for such projectiles, except for the value of the drag of the minimum-wave-drag body for given length and volume, which seems to have been tabulated erroneously in the earlier paper.

---

[6] R. Courant and D. Hilbert, *Methoden der mathematischen Physik*, vol. 1, J. Springer, Berlin, 1931, p. 83.

[7] W. Haack, *Geschossformen kleinsten Wellenwiderstandes*, Bericht 139 der Lilienthal-Gesellschaft für Luftfahrt.

THE BOUNDARY LAYER IN A CORNER*

BY

G. F. CARRIER

Brown University

1. Introduction. The laminar flow of a relatively non-viscous fluid through a channel is characterized by the presence of a thin boundary layer along the walls. In straight channels, such boundary layers are usually assumed to have the velocity distribution determined by Blasius [1] for the flow past a flat plate, and the flow pattern in the neighborhood of any corner is not mentioned. It seems of interest to develop here the change in the Blasius flow implied by such a corner.

2. The boundary layer problem. We shall consider the laminar flow of an incompressible fluid which impinges with the uniform velocity V on the edges $x=0$ of the half planes $y=0$, $z=0$.

The Navier-Stokes equations and the continuity condition which govern such flows are

$$(\mathbf{v}\cdot\text{grad})\ \mathbf{v} + \rho^{-1}\ \text{grad}\ p = \nu\Delta\mathbf{v}, \tag{1}$$

$$\text{div}\ \mathbf{v} = 0. \tag{2}$$

Here $\mathbf{v}$ is the velocity with components u, v, w; p is the pressure, ν the kinematic viscosity, and ρ the density.

As v. Kármán has pointed out [2], the essence of the treatment of such equations in a boundary layer problem is to eliminate higher order terms (by a perturbation scheme or otherwise) in such a manner that the order of the equations is not decreased. In this way no boundary conditions need be relaxed. We may accomplish this by using what is essentially Prandtl's coordinate transformation [1], namely

$$\eta = y/(\nu x/V)^{1/2}, \qquad \zeta = z/(\nu x/V)^{1/2}. \tag{3}$$

We also define the parameter $\xi = (\nu/Vx)^{1/2}$.

Since the flow both within and outside the boundary layer may be expected to be essentially in the x direction and slowly varying in x, we may attempt to find a solution in the form

$$u = V[u_0(\eta, \zeta) + \xi u_1(\eta, \zeta) + \xi^2 u_2 + \cdots] \tag{4}$$

$$v = V(\xi v_1 + \xi^2 v_2 + \cdots) \tag{5}$$

$$w = V(\xi w_1 + \xi^2 w_2 + \cdots) \tag{6}$$

$$p = \rho V^2(p_0 + \xi p_1 + \cdots). \tag{7}$$

We commence the series for v and w with a term of order ξ, because we wish a solution for which v/V, w/V, are small. Furthermore, if we included terms v_0, w_0, the following set of equations would contain terms of order ξ^{-1} with no contribution from the viscous terms of Eqs. (1) and (2). Thus the solutions wherein v_0, w_0 were not identically zero would not provide results corresponding to the phenomenon under investigation.†

* Received Aug. 30, 1946.

† Actually, the fact that our results constitute a solution which obeys the differential equation and boundary conditions is sufficient justification for taking $v_0 \equiv w_0 \equiv 0$.

The substitution of Eqs. (4) to (7) into Eqs. (1) and (2) leads to the system

$$-\frac{u_0}{2}(\eta\partial u_0/\partial\eta + \zeta\partial u_0/\partial\zeta) + v_1\partial u_0/\partial\eta + w_1\partial u_0/\partial\beta - \frac{\eta}{2}\frac{\partial p_0}{\partial\eta} - \frac{\zeta}{2}\frac{\partial p_0}{\partial\zeta}$$
$$-\left(\frac{\partial^2}{\partial\eta^2} + \frac{\partial^2}{\partial\zeta^2}\right)u_0 + \xi(\cdots) + \cdots = 0 \tag{8}$$

$$\frac{\partial p_0}{\partial\eta} + \xi\frac{\partial p_1}{\partial\eta} - \xi^2\left(\frac{u_0}{2}\left[\eta\frac{\partial v_1}{\partial\eta} + \zeta\frac{\partial v_1}{\partial\eta}\right] + \frac{\partial^2 v_1}{\partial\eta^2} + \cdots\right) + \cdots = 0 \tag{9}$$

$$\frac{\partial p_0}{\partial\zeta} + \xi\frac{\partial p_1}{\partial\zeta} - \xi^2(\cdots) + \cdots = 0 \tag{10}$$

$$\frac{\eta}{2}\frac{\partial u_0}{\partial\eta} + \frac{\zeta}{2}\frac{\partial u_0}{\partial\zeta} - \frac{\partial v_1}{\partial\eta} - \frac{\partial w_1}{\partial\zeta} + \xi(\cdots) + \cdots = 0. \tag{11}$$

The solution of this system of equations requires that the coefficient of each power of ξ in each equation vanish. The first order approximation to the result is defined by the vanishing of the coefficients of ξ^0. The result can be expected to be valid only when the remaining terms of the series are negligible, that is when ξ is small. Thus the solution, like that for the flat plate, is valid only at sufficiently large distances from the leading edges of the planes.

We now note that the ξ^0 terms of Eqs. (9) and (10) vanish only if $p_0=\text{const}$; the ξ^0 term of Eq. (11) vanishes if we write

$$u_0 = g_{\eta\zeta}(\eta, \zeta), \qquad v_1 = \tfrac{1}{2}(\eta g_{\eta\zeta} - g_\zeta), \qquad w_1 = \tfrac{1}{2}(\zeta g_{\eta\zeta} - g_\eta).$$

Thus it remains to find $g(\eta, \zeta)$ such that,

$$g(0, \zeta) = g_\eta(0, \zeta) = g(\eta, 0) = g_\zeta(\eta, 0) = 0$$

and

$$\lim_{\eta,\zeta\to\infty} g_{\eta\zeta}(\eta, \zeta) = 1,$$

the implied symmetry condition

$$g(a, b) = g(b, a),$$

and the differential equation implied by Eq. (8)

$$g_{\eta\eta\eta\zeta} + g_{\eta\zeta\zeta\zeta} + \tfrac{1}{2}\{g_\zeta g_{\eta\eta\zeta} + g_\eta g_{\eta\zeta\zeta}\} = 0. \tag{12}$$

We may expect that far from the corner the solution will be essentially that for the flat plate. Hence, we write

$$g(\eta, \zeta) = f_0(\eta)f_0(\zeta) + h(\eta, \zeta) \tag{13}$$

where f_0 is that solution of

$$2f''' + ff'' = 0$$

such that $f(0)=f'(0)=0; f'_{\alpha\to\infty}(\alpha)=1$. This function is tabulated in [1].

Equations (12) and (13) lead to the equation

$$h_{\eta\eta\eta\zeta} + h_{\eta\zeta\zeta\zeta} + 2a(\eta, \zeta)h_{\eta\eta\zeta} + 2a(\zeta, \eta)h_{\eta\zeta\zeta} + b(\eta, \zeta)h_\eta + b(\zeta, \eta)h_\zeta + \tfrac{1}{2}(h_\zeta h_{\eta\eta\zeta} + h_\eta h_{\eta\zeta\zeta}) = \tfrac{1}{25}A(\eta, \zeta), \quad (14)$$

where

$$a(\eta, \zeta) = \tfrac{1}{4}f_0(\eta)f_0'(\zeta), \qquad b(\eta, \zeta) = \tfrac{1}{2}f_0''(\eta)f_0'(\zeta)$$

$$A(\eta, \zeta) = \tfrac{1}{2}\{f_0(\eta)f_0''(\eta)f_0'(\zeta)[1 - f_0'(\zeta)] + f_0(\zeta)f_0''(\zeta)f_0'(\eta)[1 - f_0'(\eta)]\}.$$

This equation may be integrated once each over η and ζ taking account of the boundary conditions to yield (when $\varphi = -25h_{\eta\zeta}$)

$$\Delta\varphi + 2a(\eta, \zeta)\partial\varphi/\partial\eta + 2a(\zeta, \eta)\partial\varphi/\partial\zeta + b(\eta, \zeta)\int_0^\eta \varphi d\eta + b(\zeta, \eta)\int_0^\zeta \varphi d\zeta + \frac{1}{50}\left[\varphi_\eta\int_0^\eta \varphi d\eta + \varphi_\zeta\int_0^\zeta \varphi d\zeta\right] = \tfrac{1}{25}A(\eta, \zeta). \quad (15)$$

The boundary conditions are

$$\varphi(0, \zeta) = \varphi(\eta, 0) = \lim_{\eta,\zeta\to\infty} \varphi(\eta, \zeta) = 0.$$

This last form of the equation seems best suited for numerical evaluation. The relaxation method [3] appears to be the most appropriate for the determination of φ so we form the difference equation derived from Eq. (15) by taking points spaced unity apart in η and ζ. The subscripts m and n are used to index these point positions. The difference equation is

$$\varphi_{m+1,n} + \varphi_{m-1,n} + \varphi_{m,n-1} + \varphi_{m,n+1} - 4\varphi_{mn} + a_{mn}(\varphi_{m+1,n} - \varphi_{m-1,n}) + a_{nm}(\varphi_{m,n+1} - \varphi_{m,n-1}) + b_{mn}\int_0^n \varphi d\eta + b_{nm}\int_0^m \varphi d\zeta + .01\left[(\varphi_{m,n+1} - \varphi_{m,n-1})\int_0^m \varphi d\zeta + (\varphi_{m+1,n} - \varphi_{m-1,n})\int_0^n \varphi d\eta\right] + A_{mn} = 0. \quad (16)$$

In this equation the integrals may be evaluated by the simple trapezoidal rule since the function φ is very "smooth" although if more accuracy is desired a simple graphical method is conveniently employed.

TABLE I

ζ / η	$\phi(\eta, \zeta)$ 0	1	2	3	4	5	6	7	8	$f_0'(\eta)$
0	0	0	0	0	0	0	0	0	0	0
1	0	.58	1.00	1.00	.64	.25	.08	.02	.00	.330
2	0	1.00	1.60	1.46	.86	.28	.08	.02	.00	.630
3	0	1.00	1.46	1.23	.61	.16	.04	.01	.00	.846
4	0	.64	.86	.61	.24	.03	.01	.00		.955
5	0	.25	.29	.16	.03	.01	.00			.992
6	0	.08	.08	.04	.01	.00	.00			.999
7	0	.02	.02	.01	.00					1.000
8	0	00	.00	.00						1.000

The numerical procedure is this: guess values for φ at all points $m, n \leqq 8$. Replace the zero on the right side of Eq. (12) by Q_{mn} and compute each Q_{mn} (the residuals). Then revise the guesses for the φ_{mn} in such a way as to decrease the Q_{mn}, disregarding the changes in the values of the terms containing integrals. When considerable improvement has been made, recompute the Q_{mn} using the complete equation (12) and

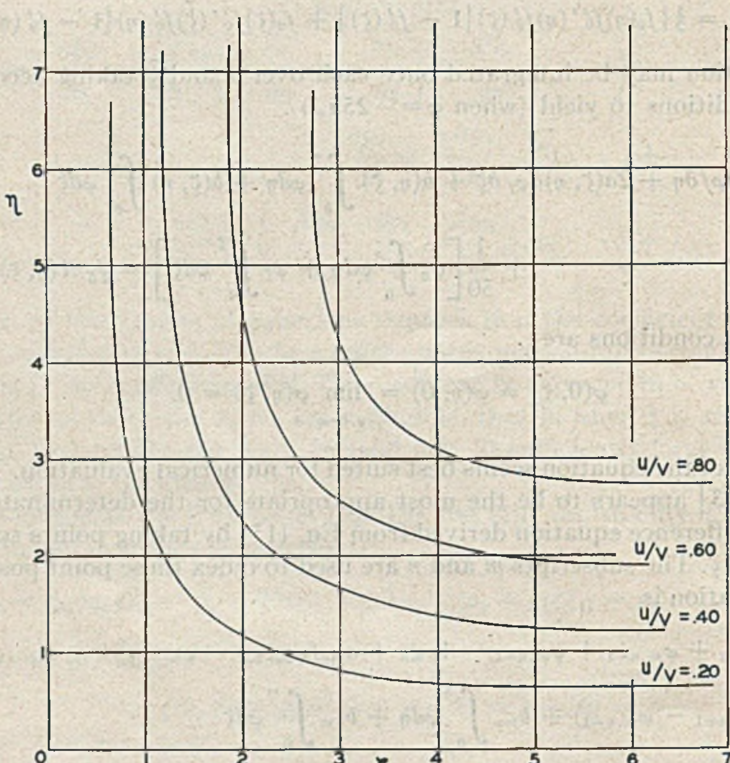

FIG. 1. Contours of constant U in corner boundary layer.

repeat the foregoing procedure. It is not necessary to get extremely accurate values of φ (especially since a, b, A are not known too finely) because the velocity $u_0 = f_0'(\eta) f_0'(\zeta) + \frac{1}{25}\varphi(\eta, \zeta)$ will be accurate to three places when φ is known to the one hundredths digit. The functions f_0' and φ are tabulated in Table I and contours of constant u_0 are shown in Fig. 1.

BIBLIOGRAPHY

1. L. Prandtl, *The mechanics of viscous fluids*, in Durand, "Aerodynamic Theory," vol. 3, J. Springer, Berlin, 1935, pp. 34–208.

2. Th. von Kármán, *The engineer grapples with non-linear problems*, Bull. Am. Math. Soc., 46, 615–683 (1940).

3. Howard Emmons, *The numerical solution of differential equations*, Q. Appl. Math., 2, 173–195 (1944).

THE TREATMENT OF SINGULARITIES OF PARTIAL DIFFERENTIAL EQUATIONS BY RELAXATION METHODS*

BY

H. MOTZ

University of Sheffield, England

1. Introduction. In the course of a study of boundary value problems arising in radiation theory and electrostatics, the treatment of singularities demanded special attention. In most problems of practical importance boundaries with sharp corners occur. Such sharp corners give rise to singularities of various types. When the computed function is bounded, but has a branch point at the corner, the difficulty is not serious. The use of a graded net with a finer mesh size near the corner is possible. Conformal transformation which automatically provides a finer net near corners is also successful. The mesh size near the corner should be of the order of magnitude of the radius of curvature of the corner, and when this is small a mathematical idealization involving infinitely sharp corners is preferable. The special treatment outlined in this note makes use of such an idealization and shortens the labour considerably. Special treatment is essential when the function approaches infinite values near the corner.

2. Plane harmonic functions. Solutions of $\nabla^2\phi=0$ are bounded when the boundary condition prescribes constant values near the corners. It can be shown that they are also bounded when $\partial\phi/\partial\nu$ is constant, where ν is the direction normal to the boundary. This type of boundary condition occurs e.g. when two plane harmonic functions ϕ and ψ are computed inside a boundary B for the purpose of a conformal transformation

$$x + iy = \phi(x, y) + i\psi(x, y)$$

and $\psi=$const. is specified at the boundary forming the corner. When expressed in polar coordinates r, ϑ centered at P (Fig. 1), the equation

$$\nabla^2\phi = 0 \tag{1}$$

becomes

$$r^2\frac{\partial^2\phi}{\partial r^2} + r\frac{\partial\phi}{\partial r} + \frac{\partial^2\phi}{\partial\vartheta^2} = 0. \tag{2}$$

With $\phi(r, \vartheta) = R(r)\cdot\Theta(\vartheta)$, the following equations for R and Θ are obtained

$$\frac{d^2\Theta}{d\vartheta^2} + n^2\Theta = 0, \tag{3}$$

$$r^2\frac{d^2R}{dr^2} + r\frac{dR}{dr} - n^2R = 0. \tag{4}$$

In these equations n^2 stands for a positive constant, and

$$\frac{d\Theta}{d\vartheta} = 0 \quad \text{when} \quad \vartheta = 0, \quad \vartheta = \alpha.$$

* Received Jan. 2, 1946.

Hence

$$\phi = \sum_{k=-\infty}^{\infty} A_k r^n \cos n\vartheta \tag{5}$$

where $n = \pi k/\alpha$ $(k = 0, \pm 1, \pm 2, \pm 3, \cdots)$.

In order to investigate the terms with negative exponent in this series, we exclude the corner by a small circle of radius ρ. On this circle $(\partial\phi/\partial r)_{\rho=0} = 0$. It is found that

$$A_{-s} = \rho^{2n} A_{+s} \qquad (s = 1, 2, 3, \cdots).$$

When $\rho \to 0$, the circle contracts towards the point P and the terms with negative exponents vanish. Thus ϕ will be represented by the series

$$\phi = A_0 + A_1 r^{\pi/\alpha} \cos \frac{\pi}{\alpha}\vartheta + A_2 r^{2\pi/\alpha} \cos \frac{2\pi}{\alpha}\vartheta + A_3 r^{3\pi/\alpha} \cos \frac{3\pi}{\alpha}\vartheta + \cdots \tag{6}$$

in the neighbourhood of P.

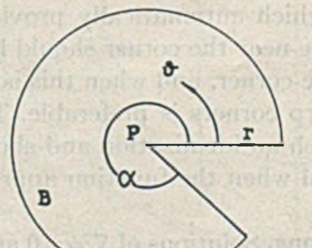

FIG. 1

3. Method of special treatment. The method of treatment will now be explained with reference to the example of a corner with $\alpha = 2\pi$, $\pi/\alpha = \frac{1}{2}$. In the treatment of two-dimensional problems by relaxation methods,[1,2] the function ϕ is computed at points of a net with small but finite mesh size. Let us denote by ϕ_0 the value of ϕ at such a point, by $\phi_1, \phi_2, \phi_3, \phi_4$ the values of ϕ at the nearest neighbouring points. The mesh length is a. At points where the function is regular, double Taylor expansion shows that

$$a^2\nabla^2\phi - \sum_{m=1}^{4} \phi_m + 4\phi_0 = -\frac{1}{12}\Delta_0^{\text{iv}}(x) - \frac{1}{12}\Delta_0^{\text{iv}}(y) \cdots, \tag{7}$$

where $\Delta_0^{\text{iv}}(x)$, $\Delta_0^{\text{iv}}(y)$, are the fourth central differences in the x and y direction, respectively, at the point where $\phi = \phi_0$. This expansion can only be used when the right hand side converges. At a singularity and its nearest neighbouring points this expansion is not valid. Figure 2 shows an example of a boundary where $\phi = 0$ on AE, $\phi = 1000$ on EB, and $\partial\phi/\partial\nu = 0$ on all other boundaries. The Taylor expansion fails at

[1] H. W. Emmons, *Numerical solution of partial differential equations*, Quarterly Appl. Math., **2**, 173–195 (1944).

[2] D. N. de G. Allen, D. G. Christopherson, L. Fox, J. R. Green, H. Motz, F. S. Shaw and R. V. Southwell, *Relaxation methods applied to engineering problems*, Phil. Trans. Royal Soc. London (A), **239**, 367–386, 419–537, 539–578 (1945).

P', Q', R', and S'. In order to obtain valid equations at these points, we consider series of the type (6) at the pivotal points P, Q, R, S

$$\phi = A_0 + A_1 r^{1/2} \cos \frac{\vartheta}{2} + A_2 r \cos \vartheta + A_3 r^{3/2} \cos 3/2\vartheta + \cdots . \tag{8}$$

Only the first four terms are retained. The units of r can be so chosen that $r=1$ at P, Q, R, S. In terms of ϕ_P, ϕ_Q, ϕ_R, ϕ_S one obtains

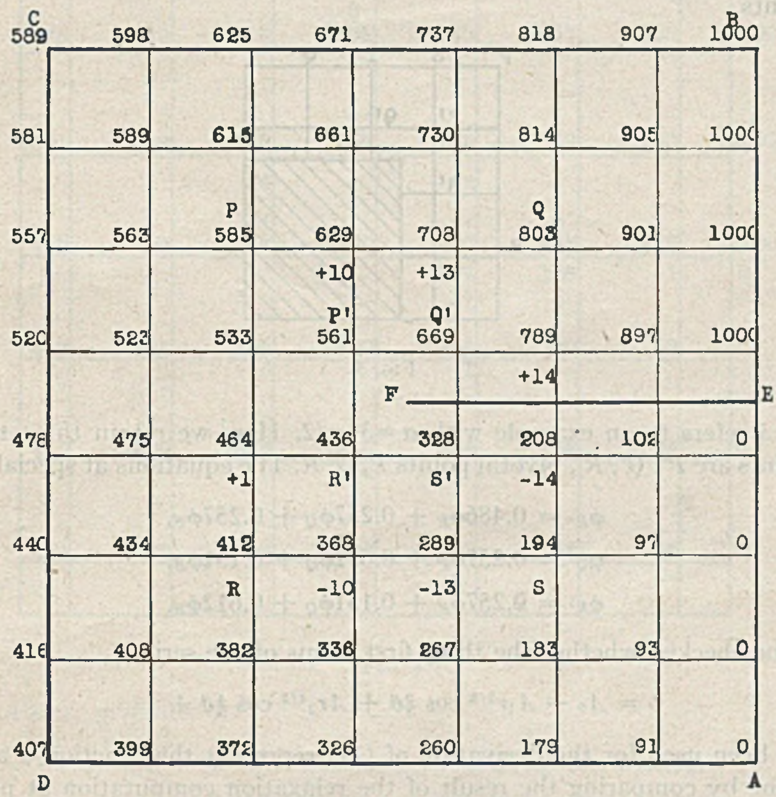

Fig. 2

$$\begin{aligned} A_0 &= 0.25\ (\phi_P + \phi_Q + \phi_R + \phi_S), \\ A_1 &= 0.191(\phi_P - \phi_R) - 0.462(\phi_S - \phi_Q), \\ A_2 &= 0.354(\phi_Q - \phi_P - \phi_R + \phi_S), \\ A_3 &= -\,0.191(\phi_Q - \phi_S) + 0.462(\phi_R - \phi_P). \end{aligned} \tag{9}$$

At the special points P', Q', R', and S', we find from (8)

$$\begin{aligned} \phi_{P'} &= 0.457\phi_P + 0.235\phi_Q + 0.209\phi_R + 0.099\phi_S, \\ \phi_{Q'} &= 0.235\phi_P + 0.593\phi_Q + 0.099\phi_R + 0.073\phi_S, \\ \phi_{R'} &= 0.209\phi_P + 0.099\phi_Q + 0.457\phi_R + 0.235\phi_S, \\ \phi_{S'} &= 0.099\phi_P + 0.073\phi_Q + 0.235\phi_R + 0.593\phi_S. \end{aligned} \tag{10}$$

These are the equations used at special points. The relaxation procedure is carried out normally everywhere, observing that equations (10) hold at special points. The residual at a special point due to an increment at a pivotal point P is therefore the product of this increment with the coefficient of ϕ_P in the equation which holds at the special point. This is in accordance with the usual relaxation procedure. The removal of a residual at a special point is particularly easy. It is simply subtracted from the value of ϕ at the special point in question. Due to this removal the usual residuals accrue at ordinary neighbouring points, but of course, no residuals are passed on to special points.

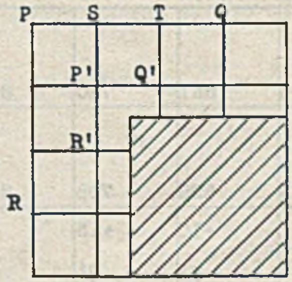

FIG. 3

Figure 3 refers to an example with $\alpha = 3 \cdot \pi/2$. Here we retain three terms only Special points are P', Q', R', pivotal points P, Q, R. The equations at special points are

$$
\begin{aligned}
\phi_{P'} &= 0.486\phi_P + 0.257\phi_Q + 0.257\phi_S, \\
\phi_{Q'} &= 0.257\phi_P + 0.612\phi_Q + 0.131\phi_S, \\
\phi_{R'} &= 0.257\phi_P + 0.131\phi_Q + 0.612\phi_S.
\end{aligned} \tag{11}
$$

It should be checked whether the three first terms of the series

$$\phi = A_0 + A_1 r^{2/3} \cos \tfrac{2}{3}\vartheta + A r_2^{4/3} \cos \tfrac{4}{3}\vartheta + \cdots \tag{12}$$

which has been used for the derivation of (11) represent the function ϕ adequately. This is done by comparing the result of the relaxation computation at points S, T, where ordinary difference equations have been used with the values of ϕ calculated by means of the first three terms of (12).

A similar check was carried out at analogous points in the example of Fig. 2. It was found that the agreement was not satisfactory. The errors have been recorded in Fig. 2 underneath the respective ϕ values. In this case it is possible (with the net shown in Fig. 4) to retain five terms of the series. Pivotal points are T, U, V, W, X; special points T', U', V', W', X', and the equations for ϕ at special points are

$$
\begin{aligned}
\phi_{T'} &= 0.546\,\phi_T + 0.313\,\phi_U + 0.062_5\phi_V + 0.062_5\phi_W + 0.016\phi_X, \\
\phi_{U'} &= 0.156\,\phi_T + 0.578\,\phi_U + 0.188\,\phi_V + 0.047\,\phi_W + 0.031\,\phi_X, \\
\phi_{V'} &= 0.031\,\phi_T + 0.188\,\phi_U + 0.562\,\phi_V + 0.188\,\phi_W + 0.031\,\phi_X, \\
\phi_{W'} &= 0.031\,\phi_T + 0.047\,\phi_U + 0.188\,\phi_V + 0.578\,\phi_W + 0.156\,\phi_X, \\
\phi_{X'} &= 0.016\,\phi_T + 0.062_5\phi_U + 0.062\,\phi_V + 0.313\,\phi_W + 0.546\,\phi_X.
\end{aligned} \tag{13}
$$

The result is again checked by comparing the result of the relaxation procedure with the ϕ values near the corner calculated from

$$\phi = A_0 + A_1 r^{1/2} \cos \tfrac{1}{2}\vartheta + A_2 r \cos \vartheta + A_3 r^{3/2} \cos \tfrac{1}{2}\vartheta + A_4 r^2 \cos 2\vartheta,$$

590	607	643	700	776	863	954	1000
573	589	623	U 682	765	857	952	1000
541	552	578	U' 639	745	849	950	1000
499	V 499	-5 V' 499	499	-2 T' 727	T 843	949	1000
				271 X'	156 X	51	0
457	446	420 +5	W' 359	254 +1	150	50	0
452	409	375	W 316	234	142	48	0
408	390	354	298	223	137	46	0

FIG. 4

where the A's are given by

$$\begin{aligned} A_0 &= 0.250(\phi_U + \phi_V + \phi_W) + 0.125(\phi_T + \phi_X), \\ A_1 &= 0.354(\phi_U - \phi_W) + 0.250(\phi_T - \phi_X), \\ A_2 &= -0.500\phi_V + 0.250(\phi_T + \phi_X), \\ A_3 &= 0.354(\phi_W - \phi_U) + 0.250(\phi_T - \phi_X), \\ A_4 &= 0.250(\phi_V - \phi_U - \phi_W) + 0.125(\phi_T + \phi_X). \end{aligned} \tag{14}$$

The agreement is now much better. In Fig. 4 the errors have been recorded. It will be noticed that the mesh points of Fig. 2 lie between those of Fig. 4. By interpolating values at midpoints of the meshes of Fig. 4, we find that the solutions given in the two figures are in fair agreement.

When the above test fails, a finer net should be used as a rule, because the calculation becomes rather cumbersome when more than five terms of the series are retained.

To obtain, without the special treatment, a result which differs from the one of Fig. 2 by less than 1% at any point of the net, the net near the corner would have to be 7 times as fine.

4. Other examples. As an example of a corner where the value of the function is specified, let us consider an electrostatic potential ψ. In this case the series is

$$\psi = A_0 + A_1 r^{1/2} \sin \tfrac{1}{2}\vartheta + A_2 r \sin \vartheta + A r^{3/2} \sin \tfrac{3}{2}\vartheta + \cdots . \tag{15}$$

The components of the electric field in Cartesian cordinates, E_x, E_y, are not bounded on a sharp corner when $\alpha > \pi$. Let us consider the term $r^n \cos n\vartheta$ of the series (5). E_x and E_y will contain terms $r^{n-1} \sin (n-1)\vartheta$, $r^{n-1} \cos (n-1)\vartheta$, respectively, and negative exponents of r will therefore occur when $\alpha > \pi$. The method outlined above can still be used to compute a function with such a singularity. In the case $\alpha = 2\pi$ the negative exponent $-\frac{1}{2}$ occurs. Terms with exponents $-\frac{1}{2}$ and $+\frac{1}{2}$ depend on ϑ in the same manner. It is therefore necessary to choose pivotal points which have not all the same distance from the corner.

The method is equally applicable to solutions of the wave equation

$$\nabla^2 \phi + k^2 \phi = 0. \tag{16}$$

In Cartesian coordinates and with $\xi = x/a$, $\eta = y/a$ (where a is the mesh length of the net), Eq. (16) becomes

$$\frac{\partial^2 \phi}{\partial \xi^2} + \frac{\partial^2 \phi}{\partial \eta^2} + k^2 a^2 \phi = 0. \tag{17}$$

Referring again to the case $\alpha = 2\pi$, $\partial\phi/\partial\nu = 0$ and retaining five terms, we see that the expression (5) holds at pivotal points. At the special points we have

$$\phi = A_0 + A_1 \frac{J_{1/2}(ka)}{J_{1/2}(2ka)} \cos \tfrac{1}{2}\vartheta + A_2 \frac{J_1(ka)}{J_1(2ka)} \cos \vartheta + A_3 \frac{J_{3/2}(ka)}{J_{3/2}(2ka)} \cos \tfrac{3}{2}\vartheta$$
$$+ A_4 \frac{J_2(ka)}{J_2(2ka)} \cos 2\vartheta,$$

where J_n are the Bessel functions of order n. When $ka < 0 \cdot 1$, the ratios $J_n(ka)/J_n(2ka)$ differ from $(\frac{1}{2})^n$ only in the third decimal. When the mesh length a is small compared with the wave length $l = 2\pi/k$, the special equations are therefore the same for solutions of the wave equation and those of Laplace's equation.

5. Conformal transformation. When a solution of more complicated differential equations, e.g. the equations of viscous flow, or $\nabla^4 F = 0$, is computed it is often an advantage to remove singularities at the boundary by a conformal transformation $\phi = \phi(x, y)$, $\psi = \psi(x, y)$. Let us suppose that it is desired to transform the interior of the boundary shTwn in Fig. 2 into the interior of a rectangle in the ϕ, ψ plane.

The lines $\phi =$const. at suitable intervals can be found from the ϕ-values recorded in Fig. 2. The lines $\psi' =$const. are orthogonal to the lines $\phi =$const. and are best computed separately.

The condition for orthogonality

$$\frac{\partial \phi}{\partial x} \frac{\partial \psi'}{\partial x} + \frac{\partial \phi}{\partial y} \frac{\partial \psi'}{\partial y}$$

is satisfied when

$$\frac{\partial \phi}{\partial x} = \lambda \frac{\partial \psi'}{\partial y}, \qquad \frac{\partial \phi}{\partial y} = -\lambda \frac{\partial \psi'}{\partial x} \tag{18}$$

where λ is a constant. From these equations it follows that

$$\frac{\partial^2 \phi}{\partial x^2} + \frac{\partial^2 \phi}{\partial y^2} = 0, \qquad \frac{\partial^2 \psi'}{\partial x^2} + \frac{\partial^2 \psi'}{\partial y^2} = 0. \tag{19}$$

The boundary conditions for ψ' are $\partial\psi'/\partial\nu = 0$ on EA, EB, $\psi' = 0$ on EF, $\psi' =$ const. on AD, DC, and CB. The last constant is arbitrary and may be given a convenient value, e.g. 1000 for three figure accuracy.

It is easily seen that it is necessary to determine the constant λ in (18), in order to carry on with the computation of the original equation (e.g. $\nabla^4 F = 0$). In the coordinates ϕ and $\psi = \lambda\psi'$, this equation becomes

$$\left(\frac{\partial^2}{\partial \phi^2} + \frac{\partial^2}{\partial \psi^2}\right)\left(\frac{\partial^2}{\partial \phi^2} + \frac{\partial^2}{\partial \psi^2}\right) F = 0.$$

The constant λ is determined by (18). These equations can be regarded as an estimate of λ at every point. Denoting the finite differences in the x and y directions at the mesh point i by

$$D_x\phi(i), \qquad D_y\phi(i), \qquad D_x\psi'(i), \qquad D_y\psi'(i),$$

the quantities $\delta_1(i)$ and $\delta_2(i)$ defined by

$$\delta_1(i) = D_x\phi(i) - \lambda D_y\psi'(i), \qquad \delta_2(i) = D_y\phi(i) + \lambda D_x\psi'(i) \tag{20}$$

are not all zero, but constitute a measure of the computational error. It is desired to find a mean value for λ for which the variance of the computational error is a minimum. The sum

$$\sum_i \left[\delta_1^2(i) + \delta_2^2(i)\right]$$

is thus minimized with respect to λ and the following expression for λ is obtained:

$$\lambda = \frac{\sum_i \left\{D_x\phi(i) D_y\psi'(i) - D_y\phi(i) D_x\psi'(i)\right\}}{\sum_i \left(\left\{D_x\psi'(i)\right\}^2 + \left\{D_y\psi'(i)\right\}^2\right)}. \tag{21}$$

It has been found that, with the help of this technique of separate computation of the two transformation functions and using the special treatment of corners at the boundary conformal transformations can be computed with great accuracy.

6. Acknowledgment. The numerical work for this paper was done by Miss L. Klanfer whose services were put at the author's disposal by the Directorate of Scientific Research of the British Admiralty.

THE GENERAL VARIATIONAL PRINCIPLE OF THE THEORY OF STRUCTURAL STABILITY*

BY

WILLIAM PRAGER

Brown University

1. Introduction. This paper is concerned with the general problem of structural stability in the elastic or plastic range. Two slightly different formulations of this problem are found in the literature. According to the first, one considers a deformable body which, initially, is free from stresses, and which is then subjected to a system of loads of gradually increasing intensity. As long as these loads are sufficiently small the equilibrium configuration which the body assumes under their influence will be stable; one asks for that intensity of the loads for which this equilibrium configuration first becomes unstable. According to the second formulation of the problem of structural stability, one considers a given configuration of a deformable body and an equilibrium system of body and surface stresses and asks whether, in the presence of these *initial* stresses, the given configuration is stable or not. This second point of view is adopted in this paper because:

(1) it clearly separates the stability problem from the problem of finding the stresses produced by the given loads, and

(2) the manner in which the initial stresses are produced is irrelevant for the solution of the stability problem. In particular, it is by no means necessary that the initial stresses are produced by loads which are applied to an otherwise stressfree body; they may be produced by temperature changes or may partly be due to previous overstraining of the body.

Once this second point of view is adopted, stress-strain relations enter into the discussion at one point only: we must be able to predict the infinitesimal changes in stress which correspond to the infinitesimal strains associated with a system of infinitesimal displacements from the considered equilibrium configuration. As the relations between these infinitesimal changes in the stresses and strains are essentially *linear*, the only difference between the elastic and plastic ranges consists in the fact that in the plastic range a different set of coefficients must be used in these linear relations according to whether the change of stress constitutes "loading" or "unloading," while no such distinction need be made in the elastic range.

In Section 2, the general problem of structural stability is reduced to an eigenvalue problem for the displacements from a configuration of indifferent equilibrium to a neighbouring configuration of this type. Except for the consideration of plastic deformations, we follow Biezeno and Hencky[1] in this derivation, but simplify the discussion by the systematic use of tensors. In Section 3, a variational principle is derived which is equivalent to the eigen-value problem formulated in Section 2. As an example for the application of this principle, the lateral buckling of an unevenly heated lamina is treated in Section 4.

---

* Received June 5, 1946.

[1] C. B. Biezeno and H. Hencky, Proc. Roy. Acad. Amsterdam, 31, 569–592 (1928).

2. The eigen-value problem associated with the general problem of structural stability. We consider a *given configuration* of a deformable body and an *equilibrium system* of body and surface stresses which is given to within an arbitrary factor λ. *If λ is sufficiently small, this equilibrium configuration will be stable; we ask for that value of λ for which it becomes indifferent, assuming that the additional stresses which are produced by infinitesimal displacements from the given equilibrium configuration are linearly related to the corresponding infinitesimal strains.* This critical value of λ will be called the *safety factor* of the considered equilibrium configuration. With respect to a system of rectangular Cartesian coordinates x_i, let us denote the components of the given stresses by $\lambda\sigma_{ij}$ and the components of an infinitesimal displacement from the given equilibrium configuration by u_i. If the unit vector along the outward normal to the surface is denoted by n_i, the surface stresses are

$$\lambda T_j = \lambda\sigma_{ij}n_i. \tag{1}$$

The quantities σ_{ij} must satisfy the equilibrium conditions

$$\sigma_{ij,i} = 0, \tag{2}$$

where the subscript i after the comma denotes differentiation with respect to x_i, and the usual summation convention regarding repeated subscripts is adopted.

The infinitesimal strain associated with the displacements u_i is given by

$$\epsilon_{ij} = \tfrac{1}{2}(u_{i,j} + u_{j,i}). \tag{3}$$

Since the relation between this strain and the corresponding additional stress τ_{ij} is assumed to be linear, we have

$$\tau_{ij} = C_{ijkl}\epsilon_{kl}, \tag{4}$$

where C_{ijkl} is a fourth order tensor which is symmetric with respect to i and j and with respect to k and l. If, in particular, τ_{ij} and ϵ_{ij} are assumed to be related to each other by the generalized law of Hooke, we have

$$C_{ijkl} = 2G_0\left(\delta_{ik}\delta_{kl} - \frac{\nu}{1-2\nu}\,\delta_{ij}\delta_{kl}\right), \tag{5a}$$

where G_0 denotes the modulus of rigidity, ν Poisson's ratio, and δ_{ij} is the Kronecker delta. If the body under consideration can be expected to behave like an isotropic elastic solid for an infinitesimal displacement from the given equilibrium configuration,[2] i.e. if the stresses $\lambda\sigma_{ij}$ do nowhere exceed the elastic limit of the material, the expression (5a) may be used in connection with the stress-strain relation (4). On the other hand, where the stresses $\lambda\sigma_{ij}$ exceed the elastic limit, different expressions must be used for C_{ijkl} according to whether the stresses τ_{ij} associated with the strains ϵ_{ij} constitute "loading" or "unloading." We reserve the complete discussion of suitable stress-strain relations beyond the elastic limit for another paper and give but one example here. Defining the *stress deviation* as $s_{ij}=\sigma_{ij}-\frac{1}{3}\sigma_{kk}\delta_{ij}$ and its *intensity* as $S=\frac{1}{2}s_{ij}s_{ij}$, we set

[2] M. A. Biot [J. Appl. Phys., 10, 860–864 (1939)] and, more recently, F. D. Murnaghan [Proc. Nat. Acad. Sci., 30, 244–247 (1944)] have pointed out that an elastic solid under initial stress can be *strictly* isotropic only if the initial stress is of the nature of a hydrostatic pressure. For the conventional structural materials, however, this small anisotropy caused by the initial stress can be disregarded as long as the initial stress does not exceed the elastic limit.

$$C_{ijkl} = 2G_0\left(\delta_{ik}\delta_{jl} - \frac{\nu}{1-2\nu}\delta_{ij}\delta_{kl}\right) - \frac{G_0 - G}{S} s_{ij}s_{kl} \quad \text{for} \quad s_{ij}\epsilon_{ij} > 0 \tag{5b}$$

and

$$C_{ijkl} = 2G_0\left(\delta_{ik}\delta_{jl} - \frac{\nu}{1-2\nu}\delta_{ij}\delta_{kl}\right) \quad \text{for} \quad s_{ij}\epsilon_{ij} < 0. \tag{5c}$$

Here G_0 denotes the value which the modulus of rigidity assumes in the elastic range, while $G=G(S)$ is the so-called *tangent modulus of rigidity*. In the elastic range $G=G_0$, and (5b) as well as (5c) reduce to (5a). The stress-strain relations which are obtained by substituting (5b) and (5c) into (4) were suggested by J. H. Laning in an unpublished paper (1942); they constitute a generalization of stress-strain relations which the present author had used in earlier papers.[3] We note that $C_{ijkl}=C_{klij}$, according to (5a), (5b), and (5c).

A generic particle with the coordinates x_i in the initial state has the coordinates $\bar{x}_i=x_i+u_i$ in the considered neighbouring state, and

$$d\bar{x}_i = (\delta_{ij} + u_{i,j})dx_j = (\delta_{ij} + \epsilon_{ij} + \omega_{ij})dx_j, \tag{6}$$

where the deformation ϵ_{ij} is defined by (3) and

$$\omega_{ij} = \tfrac{1}{2}(u_{i,j} - u_{j,i}) \tag{7}$$

is the rotation associated with the displacement u_i.

The infinitesimal force λdf_j which is transmitted across the surface element dS in the initial state equals

$$\lambda df_j = \lambda T_j dS = \lambda\sigma_{ij}n_i dS. \tag{8}$$

The force which is transmitted to the corresponding material element in the neighbouring state will be written in the form

$$\lambda d\bar{f}_j = \lambda\bar{\sigma}_{ij}n_i dS. \tag{9}$$

Note that the normal vector n_i and the area *dS in the initial state* are used in (9). This means that the stress tensor $\lambda\bar{\sigma}_{ij}$ is defined in the *Lagrangian* manner[4] with the initial state as the state of reference. Consequently, $\lambda\bar{\sigma}_{ij}$ is not a symmetric tensor; it will be written in the form

$$\lambda\bar{\sigma}_{ij} = \lambda\sigma_{ij} + \tau_{ij} + \tau'_{ij} + \tau''_{ij}, \tag{10}$$

where the terms τ_{ij}, τ'_{ij}, and τ''_{ij} are infinitesimal changes of stress defined in the following manner:

(1) the tensor τ_{ij} is symmetric; it represents the change of stress associated with the infinitesimal strain ϵ_{ij} and is given by Eq. (4);

(2) the tensor τ'_{ij}, too, depends on the strain ϵ_{ij}; it is antisymmetric and represents the change of stress necessary to restore the moment equilibrium which is expressed by the symmetry of σ_{ij} in the initial state and which is disturbed by the deformation;

[3] W. Prager, Proc. 5th Internat. Congr. Appl. Mech. Cambridge, Mass., 1938, pp. 234–237; Prikladnaia Matematika i Mekhanika 5, 419–430 (1941); Duke Math. J. 9, 228–233 (1942).

[4] H. Jeffreys has recently given a similar analysis using the Eulerian approach [Proc. Cambridge Phil. Soc. 38, 125–128 (1942)]. The Lagrangian approach seems more suitable, however, for the problem under consideration.

(3) the term τ''_{ij}, finally, depends on the rotation ω_{ij}; it represents the change of stress, with respect to the *fixed* coordinate axes, which is produced by this rotation.

Since only first order terms in ϵ_{ij} and ω_{ij} need be considered in the following analysis, the order in which the deformation ϵ_{ij} and the rotation ω_{ij} are applied is immaterial.

The antisymmetric tensor τ'_{ij} depends only on ϵ_{ij}. To find its mathematical expression, it is therefore sufficient to consider a *pure homogeneous deformation*, i.e., a deformation for which $u_{i,j}$ is independent of the coordinates and $u_{i,j}=u_{j,i}=\epsilon_{ij}$. On account of (9), the equations of equilibrium for the deformed body are

$$\int \bar{\sigma}_{ij}n_i dS = 0, \qquad \int (\bar{\sigma}_{ij}\bar{x}_k - \bar{\sigma}_{ik}\bar{x}_j)n_i dS = 0$$

or

$$\int \bar{\sigma}_{ij,i}dv = 0, \qquad \int (\bar{\sigma}_{ij}\bar{x}_k - \bar{\sigma}_{ik}\bar{x}_j)_{,i}dv = 0.$$

Since these equations must hold not only for the entire body, but also for an arbitrary portion of it, we must have

$$\bar{\sigma}_{ij,i} = 0, \quad (11) \qquad (\bar{\sigma}_{ij}\bar{x}_k - \bar{\sigma}_{ik}\bar{x}_j)_{,i} = 0. \quad (12)$$

For the considered *pure* deformation, $\tau''_{ij}=0$ and

$$\bar{x}_{i,j} = \delta_{ij} + u_{i,j} = \delta_{ij} + \epsilon_{ij}.$$

Using the symmetry of the tensors σ_{ij} and τ_{ij} in addition to the Eqs. (10), (11), (2), and neglecting higher order terms, we may therefore write (12) in the form

$$\tau'_{ij} - \tau'_{ji} = 2\tau'_{ij} = \lambda(\sigma_{ik}\epsilon_{kj} - \sigma_{jk}\epsilon_{ki}). \quad (13)$$

The tensor τ''_{ij} depends only on ω_{ij}. To find its mathematical expression, it is sufficient to consider a rigid body rotation, i.e., a system of displacements u_i which depend linearly on the coordinates x_i and satisfy $u_{i,j}=-u_{j,i}=\omega_{ij}$. By this rotation the components of the infinitesimal force transmitted across a given surface element are transformed according to

$$d\bar{f}_i = (\delta_{ij} + u_{i,j})df_j = (\delta_{ij} + \omega_{ij})df_j = df_i + \omega_{ij}df_j. \quad (14)$$

For the considered rigid body rotation $\tau_{ij}=\tau'_{ij}=0$. Using (8), (9), and (10), we may therefore write (14) in the form

$$\tau''_{ij} = -\lambda\sigma_{ik}\omega_{kj}. \quad (15)$$

Returning now to the consideration of arbitrary infinitesimal displacements u_i, we write in accordance with (10), (13), and (15):

$$\lambda\bar{\sigma}_{ij} = \lambda\sigma_{ij} + \tau_{ij} + \tfrac{1}{2}\lambda(\sigma_{ik}\epsilon_{kj} - \sigma_{jk}\epsilon_{ki}) - \lambda\sigma_{ik}\omega_{kj}. \quad (16)$$

On account of (2), the equilibrium condition (11) furnishes therefore

$$[\tau_{ij} + \tfrac{1}{2}\lambda(\sigma_{ik}\epsilon_{kj} - \sigma_{jk}\epsilon_{ki}) - \lambda\sigma_{ik}\omega_{kj}]_{,i} = 0, \quad (17)$$

and the condition $d\bar{f}_i=df_i$ furnishes

$$[\tau_{ij} + \tfrac{1}{2}\lambda(\sigma_{ik}\epsilon_{kj} - \sigma_{jk}\epsilon_{ki}) - \lambda\sigma_{ik}\omega_{kj}]n_i = 0. \quad (18)$$

Except for our more general definition of the tensor τ_{ij}, Eqs. (17) and (18) agree with those derived by Biezeno and Hencky. Biot[5] obtained the same relations from his non-linear theory of elasticity, and Neuber[6] has recently discussed the formal relation of the differential equations (17) to the fundamental equations of elasticity. As was already pointed out by Biot, Eqs. (17) differ somewhat from the equations which Trefftz[7] derived using an unconventional definition of stress. If the given state of stress, $\lambda\sigma_{ij}$, is homogeneous and if the coordinate axes have the directions of the principal axes of this state of stress, Eqs. (17) reduce to the form given by Southwell.[8]

By means of (3), (4), and (7), the quantities ϵ_{ij}, τ_{ij}, and ω_{ij} can be expressed in terms of the first derivatives of the displacement u_i. In this manner an eigen-value problem for the displacement u_i is obtained. The smallest eigen-value λ is the desired safety factor for the given distribution of initial stresses. We refrain from formulating this eigen-value problem explicitly, because in all but the most simple cases its exact solution would hardly seem possible.

3. The variational principle associated with the general problem of structural stability. The form of Eqs. (17) and (18) suggests the existence of an equivalent variational principle from which approximate solutions of stability problems can be obtained. Indeed, let us establish the Euler equations and natural boundary conditions of the variational problem

$$\delta \int [C_{pqrs}\epsilon_{pq}\epsilon_{rs} + \lambda\sigma_{pq}(u_{r,p}u_{r,q} - \epsilon_{rp}\epsilon_{rq})]dv = 0, \tag{19}$$

where only the displacements u_p and hence strains ϵ_{pq} are to be varied, but not the stresses σ_{pq} and the coefficients C_{pqrs} which depend on the stresses. If the integrand of the left-hand side of (19) is denoted by F, the Euler equations and natural boundary conditions are

$$\frac{\partial}{\partial x_i}\left(\frac{\partial F}{\partial u_{j,i}}\right) = 0, \tag{20}$$

$$\frac{\partial F}{\partial u_{j,i}} n_i = 0. \tag{21}$$

Since

$$\frac{\partial \epsilon_{pq}}{\partial u_{j,i}} = \tfrac{1}{2}(\delta_{jp}\delta_{iq} + \delta_{ip}\delta_{jq}),$$

we have

$$\frac{\partial F}{\partial u_{j,i}} = 2C_{ijkl}\epsilon_{kl} + \lambda[2\sigma_{ik}u_{j,k} - \sigma_{ik}\epsilon_{jk} - \sigma_{jk}\epsilon_{ik}]$$

$$= 2\tau_{ij} + \lambda[\sigma_{ik}\epsilon_{kj} - \sigma_{jk}\epsilon_{ki} - 2\sigma_{ik}\omega_{kj}].$$

Equations (20) and (21) thus are indeed identical with (17) and (18).

The variational principle (19) can be used in very much the same manner in which the principles of minimum potential energy and minimum complementary energy are used in elasticity:[9] by reasonable assumptions concerning the displacements u_i

[5] M. A. Biot, Phil. Mag. (7), **27**, 468–489 (1939).

[6] H. Neuber, Z. angew. Math. Mech. **23**, 321–330 (1943). The author is indebted to Professor E. Reissner for the reference to this paper.

[7] E. Trefftz, Z. angew. Math. Mech. **13**, 160–165 (1933).

[8] R. V. Southwell, Phil. Trans. Roy. Soc. London (A), **213**, 187–244 (1913).

[9] See, for instance, E. Volterra, Atti Accad. Lincei, Rend. (6), **20**, 424–428, 463–467 (1934); **21**, 14–19 (1935); **23**, 329–332 (1936).

the class of admitted functions is restricted and the variational problem simplified. In using this technique, we must see to it that the restrictions imposed on the displacements u_i do not rule out the possibility of fulfilling the boundary conditions (17).

4. An example. To illustrate the manner of application of the variational principle formulated in Section 3, let us discuss the lateral buckling of an elastic, prismatic beam of the length l which is built in at both ends. We assume that the cross section of this beam is doubly symmetric. Taking the origin of the coordinates at one end of the beam, we let the axis of x_1 coincide with the axis of the beam and the axes of x_2 and x_3 with the axes of symmetry of the cross section $x_1=0$. To simplify the expression (5a) for the coefficients C_{ijkl}, we shall assume that $\nu=0$. This assumption is in conformity with the spirit of the engineering theory of the bending of beams; in using it we must keep in mind that Young's modulus E_0 equals twice the modulus of rigidity G_0 if $\nu=0$.

As to the initial state of stress, let us consider the case where

$$\sigma_{11} = cx_2, \tag{22}$$

while all other components of σ_{ij} vanish. The constant c in (22) obviously has the dimension of a stress divided by a length. In an originally unstressed beam with built-in ends a stress distribution of the type (22) can be produced by changes of temperature which vary linearly with x_2. If the width of the beam (measured in the direction of x_3) is small in comparison to its height, (measured in the direction of x_2) the stresses (22) may produce lateral buckling. The infinitesimal displacements associated with this type of instability may be described in the following manner: a generic cross section x_1 of the beam undergoes a translation $u(x_1)$ in the direction of the x_3-axis, a rotation $-u'(x_1)$ about the x_2-axis which makes the cross section remain normal to the bent centerline of the beam, and, simultaneously, a rotation $-\theta(x_1)$ about the x_1-axis; in addition to this rigid body displacement the cross section undergoes a warping $-w(x_2, x_3)\theta'(x_1)$ which is associated with the twist $-\theta'(x_1)$. The corresponding displacement components are

$$u_1 = -x_3u'(x_1) - w(x_2, x_3)\theta'(x_1), \qquad u_2 = x_3\theta(x_1), \qquad u_3 = u(x_1) - x_2\theta(x_1). \tag{23}$$

Note that on account of the assumption $\nu=0$ the longitudinal extension $\partial u_1/\partial x_1$ is not accompanied by any lateral contraction. Particularly simple expressions for u_2 and u_3 are thus obtained. The matrices of the derivatives $u_{i,j}$ and of the strains ϵ_{ij} therefore are

$$u_{i,j} = \begin{bmatrix} -x_3u'' - w\theta'' & -\theta'\partial w/\partial x_2 & -u' - \theta'\partial w/\partial x_3 \\ x_3\theta' & 0 & -\theta \\ u' - x_2\theta' & \theta & 0 \end{bmatrix}, \tag{24}$$

$$\epsilon_{ij} = \begin{bmatrix} -x_3u'' - w\theta'' & \frac{1}{2}\theta'(x_3 - \partial w/\partial N_2) & -\frac{1}{2}\theta'(x_2 + \partial w/\partial x_3) \\ \frac{1}{2}\theta'(x_3 - \partial w/\partial x_2) & 0 & 0 \\ -\frac{1}{2}\theta'(x_2 + \partial w/\partial x_3) & 0 & 0 \end{bmatrix}. \tag{25}$$

Since $\sigma_{ij}=0$ unless $i=j=1$, we need only $u_{k1}u_{k1}-\epsilon_{k1}\epsilon_{k1}$ for the evaluation of the term with the factor λ in (19). Now, for a doubly symmetric cross section the warping function w is odd in x_2 as well as in x_3. Taking account of this fact, and keeping in mind that σ_{11} is odd in x_2 and even in x_3, we find that

$$\int \sigma_{11}(u_{k1}u_{k1} - \epsilon_{k1}\epsilon_{k1})dv = -2c\int u'\theta' x_2^2 dv = -2cI_3\int_0^l u'\theta' dx_1, \tag{26}$$

where I_3 denotes the moment of inertia of the cross section with respect to the x_3-axis.

We now proceed to the evaluation of term $C_{pqrs}\epsilon_{pq}\epsilon_{rs}$ in (19). With $\nu=0$, Eq. (5a) takes the form $C_{ijkl}=2G_0\delta_{ik}\delta_{jl}$ and the stress-strain relation (4) reduces to

$$\tau_{ij} = 2G_0\epsilon_{ij}. \tag{27}$$

In applying this, we shall replace $2G_0$ by E_0 whenever $i=j$. In view of (25), we have

$$C_{pqrs}\epsilon_{pq}\epsilon_{rs} = \tau_{pq}\epsilon_{pq} = E_0(x_3u'' + w\theta'')^2 + 4G_0(\epsilon_{12}^2 + \epsilon_{13}^2), \tag{28}$$

where ϵ_{12} and ϵ_{13} depend on the twist θ' and on the warping w per unit twist in precisely the same manner as in the case of pure torsion. In this case, however, the integral of $4G_0(\epsilon_{12}^2+\epsilon_{13}^2)$ over the cross section equals $G_0C\theta'^2$, where G_0C denotes the torsional stiffness of the beam. Adopting the warping w per unit twist found in the case of pure torsion, and setting*

$$\Gamma = \int w^2 dA, \tag{29}$$

where dA denotes the area element of the cross section, we obtain

$$\int C_{pqrs}\epsilon_{pq}\epsilon_{rs}dv = E_0I_2\int_0^l u''^2 dx_1 + E_0\Gamma\int_0^l \theta''^2 dx_1 + G_0C\int_0^l \theta'^2 dx_1, \tag{30}$$

where I_2 is the moment of inertia of the cross section with respect to the x_3-axis.

Substituting the expressions (26) and (30) into (19), we obtain

$$E_0I_2u^{\mathrm{IV}} + \lambda cI_3\theta'' = 0, \qquad E_0\Gamma\theta^{\mathrm{IV}} - G_0C\theta'' + \lambda cI_3u'' = 0 \tag{31}$$

as the Euler equations for our problem, and

$$\theta'' = 0 \quad \text{for} \quad x_1 = 0 \quad \text{and} \quad x_1 = l \tag{32}$$

as the natural boundary conditions. In addition to these natural boundary conditions, we have the imposed boundary conditions

$$\theta = u = u' = 0 \quad \text{at} \quad x_1 = 0 \quad \text{and} \quad x_1 = l. \tag{33}$$

The safety factor λ is found as the lowest eigen-value of the problem formulated by Eqs. (31), (32) and (33).

---

* Note that for the doubly symmetric section considered here the point x_1, 0, 0 is the shear center of the cross section x_1. Since w is odd with respect to x_2 and x_3, we have $w=0$ at this point. These remarks identify the definition (29) with that given by J. N. Goodier, Eng. Exp. Station, Cornell University, Bulletin No. 27 (1941), p. 9.

UNSTABLE SOLUTIONS OF A CLASS OF HILL DIFFERENTIAL EQUATIONS*

BY
GABRIEL HORVAY
McDonnell Aircraft Corporation

1. Introduction. Linear differential equations with periodic coefficients play an important role in problems of engineering and physics. The best-known of these equations is Mathieu's equation. A somewhat more complicated equation is

$$\frac{d^2v}{d\psi^2} + [\theta_{-2}e^{-2i\psi} + \theta_{-1}e^{-i\psi} + \theta_0 + \theta_1 e^{i\psi} + \theta_2 e^{2i\psi}]v = 0 \tag{1a}$$

which reduces to Mathieu's equation for

$$\theta_0^* = \theta_0, \qquad \theta_{-1}^* = \theta_{-1} = \theta_1, \qquad \theta_{-2} = \theta_2 = 0,$$

where the asterisk is used to denote the conjugate complex quantity.

This paper is concerned with the determination of the solutions

$$v(\psi) = e^{\sigma\psi} \sum_{-\infty}^{+\infty} c_k e^{ik\psi} \tag{2}$$

of Eq. (1a) subject to the restrictions

$$\theta_0^* = \theta_0, \qquad \theta_{-1}^* = \theta_1, \qquad \theta_{-2}^* = \theta_2 \tag{1b}$$

and

$$\theta_1 = O(\mu), \qquad \theta_2 = O(\mu^2), \tag{3}$$

where μ is a small positive quantity. It will be seen that solution of the problem involves the determination of the "characteristic exponent" σ from the equation

$$\sin i\pi\sigma = \sqrt{\mathcal{D}} \sin \pi\sqrt{\theta_0}, \tag{4}$$

where $\mathcal{D}$ denotes the expansion

$$\mathcal{D} = 1 + C_\delta \delta + C_\epsilon \epsilon + C_\eta \eta + C_{\delta^2}\delta^2 + C_{\delta\epsilon}\delta\epsilon + \cdots \tag{5}$$

in the three real combinations

$$\delta = \theta_{-1}\theta_1, \qquad \epsilon = \theta_{-2}\theta_2, \qquad \eta = \tfrac{1}{2}(\theta_1^2\theta_{-2} + \theta_{-1}^2\theta_2) \tag{6a}$$

of the four quantities, real and imaginary parts of θ_1 and θ_2. $\mathcal{D}$ is a power series in μ^2 since

$$\delta = O(\mu^2), \qquad \epsilon = O(\mu^4), \qquad \eta = O(\mu^4). \tag{6b}$$

The coefficients C of the series depend on θ_0 alone.

The numerical evaluation of the coefficients of the expansion is the principal aim of this paper. This is best accomplished by first re-expressing the "doubly infinite" *Hill determinant* $\mathcal{D}$ in terms of its "simply infinite" principal subdeterminants D_n,

---

* Received June 13, 1946.

$$\mathcal{D} = f(D_0, D_1, D_2, \cdots), \tag{7}$$

and then expanding D_n into the series

$$D_n = 1 + \overset{n}{A}_\delta \delta + \overset{n}{A}_\epsilon \epsilon + \overset{n}{A}_\eta \eta + \overset{n}{A}_{\delta^2} \delta^2 + \cdots. \tag{8}$$

The coefficients of the expansions (7) and (8) are tabulated in Tables II and I respectively for a convenient range of θ_0. For the sake of simpler printing the notation

$$\overset{n}{A}_{\delta^i \epsilon^j \eta^k} \equiv \{n, \delta^i \epsilon^j \eta^k\} \tag{8'}$$

will be used whenever the subscript of A becomes excessively long.

The practical solution of Eq. (1) is carried out in four steps. First, the determinants D_n, Eqs. (8), are evaluated by means of Table I. Next $\mathcal{D}$, Eq. (7), is determined from Table II. The third step consists in solving Eq. (4) or one of its variants (13a, b, c) for σ, and the last step is the determination of the coefficients c_k of solution (2). A convenient method for carrying out this last step is discussed in Section 2. The derivations of the formulas for $\{n, \delta^i \epsilon^j \eta^k\}$ and for the coefficients of (7) are presented in Section 3. A numerical example is given in Section 4.

The present paper is based on a study which was recently undertaken at the McDonnell Aircraft Corp. under the sponsorship of the Bureau of Aeronautics, U. S. Navy Department. The study was prompted by recent instances of control difficulties of some helicopters and rotor blade failures of others. As will be shown in a separate paper,[1] the natural modes in which hinged rotor blades flap can be represented by solutions of Eq. (1) multiplied by suitable damping factors. It will be found that the stability of the blade motion decreases as the speed of advance of the helicopter increases (as μ increases). Nevertheless, instability does not set in, because an aerodynamic damping effect outweighs, at all feasable speeds, the tendency towards instability which results from the flapping motion.

The writer's thanks are due to his colleague, Elizabeth J. Spitzer, for checking the derivations and the numerical work. The writer also wishes to express his indebtedness to Messrs. W. R. Foote, H. Poritsky and J. J. Slade, who in their paper on rotational instability of shafts[2] applied a Laplace expansion to a doubly infinite determinant, and thus suggested the present approach.

2. Method of solution. The solution of Eq. (1a) is assumed in the standard form

$$v(\psi) = e^{\sigma\psi} \sum_{-\infty}^{+\infty} c_k e^{ik\psi}. \tag{2}$$

Substitution of expression (2) into Eq. (1a) leads to the infinite set of homogeneous equations for the coefficients $c_k(\sigma)$:

$$\begin{array}{lll}
\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot & & \cdot\,, \\
k=-2: & \theta_2 c_{-4} + \theta_1 c_{-3} + [(\sigma - 2i)^2 + \theta_0] c_{-2} + \theta_{-1} c_{-1} + \theta_{-2} c_0 & = 0, \\
k=-1: & \theta_2 c_{-3} + \theta_1 c_{-2} + [(\sigma - i)^2 + \theta_0] c_{-1} + \theta_{-1} c_0 + \theta_{-2} c_1 & = 0, \quad (9) \\
k=0: & \theta_2 c_{-2} + \theta_1 c_{-1} + [\sigma^2 + \theta_0] c_0 + \theta_{-1} c_1 + \theta_{-2} c_2 & = 0,
\end{array}$$

---

[1] G. Horvay, *Rotor blade flapping motion*, to be published soon.

[2] W. R. Foote, H. Poritsky and J. J. Slade, *Critical speeds of a rotor with unequal shaft flexibilities, mounted in bearings of unequal flexibility*, Journal of Applied Mechanics, 10, A77, 1943.

$$k=-1: \qquad \theta_2 c_{-1}+\theta_1 c_0+[(\sigma+i)^2+\theta_0]c_1+\theta_{-1}c_2+\theta_{-2}c_3 \qquad =0,$$

$$k=-2: \qquad \theta_2 c_0+\theta_1 c_1+[(\sigma+2i)^2+\theta_0]c_2+\theta_{-1}c_3+\theta_{-2}c_4=0,$$

$$\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot$$

The equations are consistent if their determinant, $\Delta(\sigma)$, vanishes. The consistency criterion

$$\Delta(\sigma) = 0 \tag{10}$$

can be expressed in the much simpler form[3,4]

$$\sin i\pi\sigma = \pm \sqrt{\mathcal{D}} \sin \pi\sqrt{\theta_0}, \tag{4}$$

where

$$\mathcal{D} \equiv \Delta(0) = \begin{vmatrix} \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & 1 & \theta_{-1}y_2 & \theta_{-2}y_2 & 0 & 0 & \cdot \\ \cdot & \theta_1 y_1 & 1 & \theta_{-1}y_1 & \theta_{-2}y_1 & 0 & \cdot \\ \cdot & \theta_2 y_0 & \theta_1 y_0 & 1 & \theta_{-1}y_0 & \theta_{-2}y_0 & \cdot \\ \cdot & 0 & \theta_2 y_1 & \theta_1 y_1 & 1 & \theta_{-1}y_1 & \cdot \\ \cdot & 0 & 0 & \theta_2 y_2 & \theta_1 y_2 & 1 & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{vmatrix} \tag{11a}$$

is the determinant of system (9) for $\sigma=0$ when each equation is divided by the coefficient of the diagonal term, and

$$y_k = \frac{1}{\theta_0 - k^2}\cdot \tag{11b}$$

$\mathcal{D}$ is either positive or negative, and so is θ_0. Thus the quantity $\sqrt{\mathcal{D}} \sin \pi\sqrt{\theta_0}$ is either real or pure imaginary. In the first case set

$$q = \sqrt{\mathcal{D}} \sin \pi\sqrt{\theta_0} = -\sqrt{-\mathcal{D}} \sinh \pi\sqrt{-\theta_0}. \tag{12a}$$

In the second case set

$$q' = q/i = -\sqrt{-\mathcal{D}} \sin \pi\sqrt{\theta_0} = -\sqrt{\mathcal{D}} \sinh \pi\sqrt{-\theta_0}. \tag{12b}$$

Then the solution of the transcendental equation (4) is given by

$$\pm\sigma = \frac{1}{\pi} \log (q/i + \sqrt{1-q^2}) + mi \qquad (m = 0, \pm 1, \pm 2, \cdots),$$

$$= \frac{i}{\pi} \arctan \frac{-q}{\sqrt{1-q^2}} + mi \qquad \text{for } -1 \leqq q \leqq 1, \tag{13a}$$

$$= \frac{1}{\pi} \log (q + \sqrt{q^2-1}) + (m - \tfrac{1}{2})i \quad \text{for } q \leqq -1, q \geqq 1, \tag{13b}$$

$$= \frac{1}{\pi} \log (q' + \sqrt{q'^2+1}) + mi \qquad \text{for } q \text{ imaginary}. \tag{13c}$$

---

[3] Whittaker and Watson, *A course of modern analysis*, Cambridge, 1927, p. 416.

[4] M. J. O. Strutt, *Lamésche, Mathieusche und verwandte Funktionen in Physik und Technik*, Springer 1932 (Edward Bros., 1944), p. 22.

Once $\mathcal{D}$ is known, the calculation of σ from (13) is simple matter. For any m there are two solutions σ_1 and σ_2 which differ only in sign

$$\sigma_2 = -\sigma_1. \tag{14a}$$

Let σ_1 be the solution with the positive real part

$$\sigma_1 = \sigma_r + i\sigma_i, \qquad \sigma_r \geqq 0. \tag{14b}$$

The function $v_1(\psi)$ (or $v_2(\psi)$) associated with σ_1 (or with σ_2) is readily obtained by placing σ_1 (or σ_2) into the system (9) which is now limited to the equations $k=-N$, $-N+1, \cdots, -1, +1, \cdots, +N$. Assuming $c_k=0$ for $k<-N$ and $k>+N$, one can solve the $2N$ equations for $c_{-N}, c_{-N+1}, \cdots, c_{-1}, c_1, \cdots, c_N$ in terms of the arbitrary constant c_0, and then use equation $k=0$ as a check. The greater is N, the more harmonics are taken into account, and the more accurate is the solution. In practice the calculations are most conveniently carried out by solving equations $k=N$ and $k=-N$ for c_N and c_{-N} in terms of the variables c_{N-1}, c_{N-2} and c_{-N+1}, c_{-N+2}, respectively. The results are then substituted into equations $k=N-1$ and $k=-N+1$; similarly c_{N-1} and c_{-N+1} are determined. Continuing the process one finally arrives at equations $k=+1$ and $k=-1$ involving the two variables c_1 and c_{-1} only, and the parameter c_0 which can be assumed as 1. One eliminates one of the unknowns, say c_{-1} determines from the real and imaginary parts of the remaining equation the real and imaginary parts of c_1, and then, retracing the steps, obtains in succession the numerical values of $c_{-1}, c_2, c_{-2}, \cdots, c_N, c_{-N}$.

Evidently, in principle, it is immaterial what mi ($m=0, \pm 1, \pm 2, \cdots$) is used in the σ of Eqs. (9). For instance, the set of equations $k=-N$ to $+N$ with $m=2$, the set of equations $k=-N-2$ to $+N-2$ with $m=4$, and the set of equations $k=-N+3$ to $+N+3$ with $m=-1$ are identical. Thus, as one passes to the limit $N\to\infty$, any m and any $2N+1$ adjoining equations will lead to the same function $v_1(\psi)$ [or $v_2(\psi)$]. In practice, where one is limited to a finite number of equations, $2N+1$, it is best to use the centrally located equations $k=-N$ to $+N$ with an m which makes c_0 the dominating term in the series (2).

In general v_1, associated with σ_1, and v_2, associated with σ_2, are linearly independent functions. An exceptional case arises when $\sigma_r=0$, and σ_i is an integral multiple of $\frac{1}{2}$. Then substitution of σ_1 and σ_2 into the system (9) leads to the same function

$$v_1 = e^{i\nu\psi}\sum_{-\infty}^{+\infty} c_k e^{ik\psi}, \qquad (\nu = 0 \text{ or } \tfrac{1}{2}). \tag{15a}$$

The second, linearly independent, solution is now a "quasiperiodic function":[4a]

$$v_2 = e^{i\nu\psi}\left[\psi\sum_{-\infty}^{+\infty} c_k e^{ik\psi} + \sum_{-\infty}^{+\infty} d_k e^{ik\psi}\right], \qquad (\nu = 0, \tfrac{1}{2}). \tag{15b}$$

For convenience the functions (15a) will be called "purely periodic" functions.

Determination of the purely periodic solutions (15a) forms the subject matter of most investigations on Mathieu and Hill differential equations. The purely periodic

---

[4a] M. J. O. Strutt, *loc. cit.*, p. 23. As an exception there may be two purely periodic solutions. For instance, for $\theta_0=4$, $\theta_1=\theta_2=0$, one obtains $v_1=\cos 2\psi$, $v_2=\sin 2\psi$.

solutions are usually of greatest interest, because they separate the μ-regions of stability ($\sigma_r=0$; v_1 and v_2 are oscillatory) from the μ-regions of instability ($\sigma_r>0$, $v_1\rightarrow\infty$ as $\psi\rightarrow\infty$). A purely periodic solution can be obtained, cf. Eqs. (12), (13), only when $q'=0$ ($\mathcal{D}=0$, or perhaps $\theta_0=k^2$), or when $q=\pm 1$ ($\mathcal{D}$ and θ_0 are such that $\sqrt{\mathcal{D}}\sin\pi\sqrt{\theta_0}=1$ or $\sqrt{-\mathcal{D}}\sinh\pi\sqrt{-\theta_0}=1$). In general $\theta_0=k^2$ does not provide a purely periodic function.

In the present analysis the principal interest is attached to the unstable solutions

$$v_1(\psi) = e^{\sigma_r\psi}\sum_{-\infty}^{+\infty} c_k e^{i(\sigma_i+k)\psi}, \qquad (\sigma_r>0) \tag{16}$$

of (1) which, after multiplication by a damping factor $e^{-n\psi/2}$, are still stable. These solutions are in the "transition region" which extends from the μ-value for which $v(\psi)$ is purely periodic to the μ-value for which $e^{-n\psi/2}v(\psi)$ is purely periodic. It will be seen in Reference 1 that a rapidly advancing helicopter usually operates in the transition region.

3. Expansion of Hill's infinite determinant. It will be convenient to call the determinant $\mathcal{D}$, Eq. (11a), a doubly infinite determinant to indicate that it extends to infinity both upward and downward. Simply infinite determinant are the principal subdeterminants of $\mathcal{D}$,

$$D_n = \begin{vmatrix} 1 & \theta_{-1}y_n & \theta_{-2}y_n & 0 & \cdot \\ \theta_1 y_{n+1} & 1 & \theta_{-1}y_{n+1} & \theta_{-2}y_{n+1} & \cdot \\ \theta_2 y_{n+2} & \theta_1 y_{n+2} & 1 & \theta_{-1}y_{n+2} & \cdot \\ 0 & \theta_2 y_{n+3} & \theta_1 y_{n+3} & 1 & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{vmatrix}; \quad E_n = \begin{vmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & 1 & \theta_{-1}y_{n+3} & \theta_{-2}y_{n+3} & \\ \cdot & \theta_1 y_{n+2} & 1 & \theta_{-1}y_{n+2} & \theta_{-1}y_{n+2} \\ \cdot & \theta_2 y_{n+1} & \theta_1 y_{n+1} & 1 & \theta_{-1}y_{n+1} \\ \cdot & 0 & \theta_2 y_n & \theta_1 y_n & 1 \end{vmatrix}. \tag{17a, b}$$

The first extends to infinity downward, the second upward. Simply infinite determinants are also the auxiliary subdeterminants

$$S_n = \begin{vmatrix} \theta_1 y_n & \theta_{-1}y_n & \theta_{-2}y_n & 0 & \cdot \\ \theta_2 y_{n+1} & 1 & \theta_{-1}y_{n+1} & \theta_{-2}y_{n+1} & \cdot \\ 0 & \theta_1 y_{n+2} & 1 & \theta_{-1}y_{n+2} & \cdot \\ 0 & \theta_2 y_{n+3} & \theta_1 y_{n+3} & 1 & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{vmatrix}; \quad T_n = \begin{vmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & 1 & \theta_{-1}y_{n+3} & \theta_{-2}y_{n+3} & 0 \\ \cdot & \theta_1 y_{n+2} & 1 & \theta_{-1}y_{n+2} & 0 \\ \cdot & \theta_2 y_{n+1} & \theta_1 y_{n+1} & 1 & \theta_{-2}y_{n+1} \\ \cdot & 0 & \theta_2 y_n & \theta_1 y_n & \theta_{-1}y_n \end{vmatrix}. \tag{18a, b}$$

S_n differs from D_n only in the first column; T_n differs from E_n only in the rightmost column.

One readily establishes the recurrence relations

$$D_n = D_{n+1} - \theta_{-1}y_n S_{n+1} + \theta_1\theta_{-2}y_n y_{n+1}S_{n+2} - \epsilon y_n y_{n+2}D_{n+3} + \epsilon^2 y_n y_{n+1}y_{n+2}y_{n+3}D_{n+4}, \tag{19a}$$

$$S_n = \theta_1 y_n D_{n+1} - \theta_{-1}\theta_2 y_n y_{n+1}D_{n+2} + \epsilon y_n y_{n+1}S_{n+2}. \tag{19b}$$

A Laplace expansion of the doubly infinite determinant $\mathcal{D}$ along a dividing line between row $k=0$ and $k=-1$ leads to the following expression involving only simply infinite determinants of type D_n, E_n, S_r, T_n:

$$\begin{aligned} \mathcal{D} = {} & D_0E_1 - S_0T_1 + y_0y_1(\theta_{-1}\theta_2 D_1T_2 + \theta_1\theta_{-2}S_1E_2) \\ & - \epsilon(y_0y_2D_1E_3 + y_1^2D_2E_2 + y_0y_1S_1T_2) \\ & + \epsilon^2(y_0y_1y_2y_3D_1E_4 + y_0y_1^2y_2E_2D_3 + y_0y_1^2y_2D_2E_3). \end{aligned} \tag{20}$$

Since, by virtue of (1b)

$$E_n = D_n^* = D_n, \tag{21a}$$

$$T_n = S_n^* \qquad (\neq S_n, \text{ when } \theta_1^* \neq \theta_1,\ \theta_2^* \neq \theta_2) \tag{21b}$$

and by virtue of (2) and (19b)

$$S_n T_m = O(\mu^2), \tag{21c}$$

a replacement of E_n by D_n and a repeated insertion of (19b) into (20) gradually eliminates all but the D_n type of determinants from the expression for $\mathcal{D}$. It is also found that $\theta_1, \theta_{-1}, \theta_2, \theta_{-2}$ appear only in the combinations δ, ϵ, η given in (6a). Thus, by virtue of (6b), the expansion of $\mathcal{D}$ progresses in powers of μ^2. Using the notation

$$y_{0112} = y_0 \overset{2}{y_1} y_2, \tag{22a}$$

one finds that to μ^{10} terms

$$\begin{aligned} \mathcal{D} = D_0 D_1 &- \delta y_{01} D_1 D_2 - \epsilon(y_{02} D_1 D_3 + y_{11} \overset{2}{D_2}) + \eta(2 y_{011} \overset{2}{D_2} + 2 y_{012} D_1 D_3) \\ &- \delta\epsilon(4 y_{0112} D_2 D_3 + 2 y_{0123} D_1 D_4) + \epsilon^2 (y_{0123} D_1 D_4 + 2 y_{0112} D_2 D_3) \\ &+ \epsilon\eta(4 y_{01123} D_2 D_4 + 2 y_{01234} D_1 D_5 + 2 y_{01122} \overset{2}{D_3}) \\ &- \delta\epsilon^2 (4 y_{011234} D_2 D_5 + 4 y_{011223} D_3 D_4 + 2 y_{012345} D_1 D_6) + \cdots . \end{aligned} \tag{23}$$

The same process can also be carried out for the simply infinite determinant D_n. Disregarding the exceptional case $\theta_0 = k^2$ ($y_k = \infty$), one finds that to μ^{10} terms

$$\begin{aligned} D_0 = D_1 &- \delta y_{01} D_2 + (-\epsilon y_{02} + 2\eta y_{012}) D_3 + (-2\delta\epsilon + \epsilon^2) y_{0123} D_4 \\ &+ 2\epsilon\eta y_{01234} D_5 - 2\delta\epsilon^2 y_{012345} D_6 + \cdots , \end{aligned} \tag{24a}$$

and D_n is obtained from D_0 by increasing the subscripts in the latter's expression by n. (24b)

It will be convenient to introduce at this point the notation

$$\sum y_{356} = y_{356} + y_{467} + y_{578} + \cdots , \tag{22b}$$

$$\sum y_{12} \sum y_{356} = y_{12} \sum y_{356} + y_{23} \sum y_{467} + y_{34} \sum y_{578} + \cdots . \tag{22c}$$

Noting that

$$\lim_{n\to\infty} D_n = 1, \tag{25}$$

one obtains, by repeated application of (24),

$$\begin{aligned} D_0 &= (1 - \delta y_{01}) D_2 + (-\delta y_{12} - \epsilon y_{02} + 2\eta y_{012}) D_3 + \cdots \\ &= [1 - (y_{01} + y_{12})\delta - y_{02}\epsilon + 2 y_{012}\eta] D_3 + \cdots \qquad (26a) \\ &= 1 + \overset{0}{A}_\delta \delta + \overset{0}{A}_\epsilon \epsilon + \overset{0}{A}_\eta \eta + \cdots , \qquad (26b) \end{aligned}$$

where

$$\overset{0}{A}_\delta = -\sum y_{01}, \quad \overset{0}{A}_\epsilon = -\sum y_{02}, \quad \overset{0}{A}_\eta = 2\sum y_{012}, \quad \overset{0}{A}_{\delta^2} = \sum y_{01} \sum y_{23}, \cdots . \tag{27}$$

The coefficient of a general term, like $\delta\epsilon^2$, is obtained as follows: Equation (24a) gives rise to the following symbolic products containing $\delta\epsilon^2$:

$$[-\delta y_{01}][-\epsilon y_{02}][-\epsilon y_{02}], \tag{a}$$

$$[-\delta y_{01}][\epsilon^2 y_{0123}], \tag{b}$$

$$[-\epsilon y_{02}][-2\delta\epsilon y_{0123}], \tag{c}$$

$$[1][-2\delta\epsilon^2 y_{012345}]. \tag{d}$$

It is found that (a) contributes

$$-\sum y_{01}\sum y_{24}\sum y_{57}-\sum y_{02}\sum y_{34}\sum y_{57}-\sum y_{02}\sum y_{35}\sum y_{67} \tag{28a}$$

to $A^0_{\delta\epsilon^2}$. (Note that no subscripts can be repeated, nor can any be skipped as one passes from one $\sum$ to the next $\sum$; furthermore (a) gives rise to three distinct summation expressions, because $y_{k,k+1}$ can appear in the first place, in the second place, and in the third place.) The relation (b) yields

$$-\sum y_{01}\sum y_{2345}-\sum y_{0123}\sum y_{45}, \tag{28b}$$

which (c) yields

$$2\sum y_{02}\sum y_{3456}+2\sum y_{0123}\sum y_{46} \tag{28c}$$

and (d) yields

$$-2\sum y_{012345}. \tag{28d}$$

The contributions (28a, b, c, d) sum up to $\{0, \delta\epsilon^2\}$. By increasing the subscripts of the expressions (28) by 2, one obtains

$$\begin{aligned} A^0_{\delta\epsilon^2} = &-\sum y_{23}\sum y_{46}\sum y_{79}-\sum y_{24}\sum y_{56}\sum y_{79}-\sum y_{24}\sum y_{57}\sum y_{89}\\ &-\sum y_{23}\sum y_{4567}-\sum y_{2345}\sum y_{67}+2\sum y_{24}\sum y_{5678}\\ &+2\sum y_{2345}\sum y_{68}-2\sum y_{234567}. \end{aligned} \tag{29}$$

The determination of the other $\{n, \delta^i\epsilon^j\eta^k\}$ is similar.

The numerical values of the coefficients $\{n, \delta^i\epsilon^j\eta^k\}$ are given in Table I to 5 decimal places, for θ_0 ranging from $+0.9$ to -1.0 (the interesting range in helicopter theory). In the evaluation of $\{n, \delta^i\epsilon^j\eta^k\}$ the first 51 y_m were taken into account. (The accuracy obtainable is thus equivalent to the use of a 101-row approximant to $\mathfrak{D}$.) y_0 to y_{20} were computed in some instances to 6, in some instances to 7 decimal places; y_{21} to y_{50} were computed to 7 decimal places. It is expected that the entries of Table I are in error by not more than 2 units in the fifth decimal place.[5]

It is readily seen that the present method is not limited to Eq. (1), but can be extended to the general Hill differential equation where $\theta_m e^{im\psi}$ form a convergent series.

For the special case of Mathieu's equation ($\epsilon=\eta=0$), one finds by (23) and (27) that

$$\begin{aligned} \mathfrak{D} = D_1(D_0-\delta y_{01}D_2) &= 1-2\delta\sum_{k=0}^{\infty}\frac{1}{\theta_0-k^2}\cdot\frac{1}{\theta_0-(k+1)^2}+O(\delta^2)\\ &= 1-2\delta\,\frac{\pi\cot\pi\sqrt{\theta_0}}{(4\theta_0-1)\sqrt{\theta_0}}+O(\delta^2). \end{aligned} \tag{30}$$

[5] An experienced computer can calculate a column of Table I in somewhat less than a day.

This formula was used by H. Bremekamp in 1926 in a study of the flow of electrons in metals.[6]

4. Example (a). Given $\theta=0.2$, $\theta_1=0.19685+0.33465i$, $\theta_2=0.03875-0.10258i$. Determine v_1, v_2. One finds $\delta=0.15074$, $\epsilon=0.01202$, $\eta=-0.01635$, and, by Table I, $D_0=1.8422$, $D_1=0.9434$, $D_2=0.9934$, $D_3=0.9980$, $D_4=0.999$, $D_5=D_6=1.000$. Likewise, by Table II, $\mathfrak{D}=2.3291$. Therefore, $q=\sqrt{\mathfrak{D}}\sin\pi\sqrt{\theta_0}=1.5052$ and by (13b) $\sigma_1=0.30782+i/2$, $\sigma_2=-0.30782-i/2$. The associated functions $v_1(\psi)$ and $v_2(\psi)$ are determined from the equation system (9). Normalizing to $c_0=1$, and using the equations $k=-4$ to $k=+4$, one obtains[7]

$$v_1 = (-0.1898 + 1.9817i)e^{+0.3078\psi}\left\{-0.0958\cos\frac{\psi}{2} + \sin\frac{\psi}{2} + 0.2076\cos\frac{3\psi}{2} - 0.0083\sin\frac{3\psi}{2} - 0.0125\cos\frac{5\psi}{2} - 0.0038\sin\frac{5\psi}{2} + 0.0008\cos\frac{7\psi}{2} + 0.0019\sin\frac{7\psi}{2} + \cdots\right\}$$

$$v_2 = (1.6652 + 0.7467i)e^{-0.3078\psi}\left\{\cos\frac{\psi}{2} + 0.4484\sin\frac{\psi}{2} + 0.2107\cos\frac{3\psi}{2} + 0.0188\sin\frac{3\psi}{2} + 0.0041\cos\frac{5\psi}{2} + 0.0101\sin\frac{5\psi}{2} + 0.0005\cos\frac{7\psi}{2} + 0.0020\sin\frac{7\psi}{2} + \cdots\right\}.$$

5. Example (b). Given $\theta_0=0$, $\theta_1=0.37249+0.63323i$, $\theta_2=0.13875-0.36728i$. Determine σ. One finds

$$q = \pi\sqrt{\mathfrak{D}\theta_0}\cdot\sin\pi\sqrt{\theta_0}/\pi\sqrt{\theta_0} = \pi\sqrt{0.4392} = 2.082,$$
$$\sigma_1 = -\sigma_2 = 0.4339 + i/2.$$

---

[6] M. J. O. Strutt, *loc. cit.*, p. 26.

[7] Note that the use of $\sigma_1=0.30782-i/2$ leads to the above expression of v_1 when c_1 is normalized to 1, and to the conjugate complex of the above when c_0 is normalized to 1.

TABLE I.* Numerical values of $\{n, \delta^i\epsilon^j\eta^k\}$.

θ_0 =	.9	.8	.7	.6	.5	.45	.4	.35
$\{0, \delta\}$	7.83180	4.63590	3.70199	3.38325	3.38203	3.48245	3.65865	3.92977
$\{0, \epsilon\}$	−.90665	−.24871	.00257	.16467	.30899	.38655	.47423	.57886
$\{0, \eta\}$	6.36577	3.51924	2.63695	2.27050	2.14607	2.15140	2.20218	2.30622
$\{0, \delta^2\}$	−.55015	−.30117	−.22352	−.19037	−.17858	−.17822	−.18162	−.18937
$\{0, \delta\epsilon\}$	.52473	.27858	.20047	.16578	.15052	.14789	.14837	.15229
$\{0, \delta\eta\}$	−.15952	−.08611	−.06307	−.05310	−.04908	−.04867	−.04928	−.05105
$\{0, \delta^3\}$	.00265	.00143	.00104	.00087	.00081	.00080	.00081	.00083
$\{0, \epsilon^2\}$	−.41427	−.22658	−.16801	−.14320	−.13403	−.13372	−.13622	−.14199
$\{0, \epsilon\eta\}$	−.00590	−.00363	−.00301	−.00278	−.00280	−.00292	−.00306	−.00330
$\{0, \eta^2\}$	−.00638	−.00342	−.00249	−.00208	−.00191	−.00189	−.00190	−.00196
$\{0, \delta^2\epsilon\}$	−.00463	−.00248	−.00178	−.00150	−.00136	−.00133	−.00133	−.00136
$\{0, \delta^2\eta\}$	.00049	.00025	.00019	.00015	.00014	.00014	.00013	.00014
$\{0, \delta\epsilon^2\}$	.00317	.00170	.00123	.00103	.00095	.00093	.00094	.00096
$\{0, \delta\epsilon\eta\}$	−.00021	−.00012	−.00009	−.00008	−.00007	−.00007	−.00007	−.00007
$\{1, \delta\}$	−3.27931	−1.61410	−1.05991	−.78342	−.61797	−.55795	−.50802	−.46584
$\{1, \epsilon\}$	−1.26507	−.63934	−.43033	−.32553	−.26241	−.23943	−.22021	−.20392
$\{1, \eta\}$	−.80268	−.38700	−.24905	−.18048	−.13964	−.12488	−.11264	−.10233
$\{1, \delta^2\}$	.04439	.02131	.01366	.00986	.00759	.00678	.00610	.00553
$\{1, \delta\epsilon\}$	−.01646	−.00763	−.00471	−.00328	−.00243	−.00213	−.00187	−.00166
$\{1, \delta\eta\}$	.00739	.00351	.00223	.00160	.00122	.00109	.00098	.00088
$\{1, \delta^3\}$	−.00009	−.00004	−.00003	−.00002	−.00002	−.00001	−.00001	−.00001
$\{1, \epsilon^2\}$	.03150	.01513	.00970	.00701	.00540	.00482	.00434	.00394
$\{1, \epsilon\eta\}$	.00132	.00063	.00041	.00030	.00024	.00020	.00019	.00017
$\{1, \eta^2\}$	.00019	.00009	.00006	.00004	.00002	.00002	.00002	.00002
$\{2, \delta\}$	−.05351	−.05160	−.04981	−.04813	−.04654	−.04579	−.04505	−.04434
$\{2, \epsilon\}$	−.03050	−.02958	−.02872	−.02791	−.02714	−.02678	−.02642	−.02607
$\{2, \eta\}$	−.00619	−.00591	−.00565	−.00541	−.00519	−.00508	−.00498	−.00487
$\{2, \delta^2\}$	.00025	.00024	.00023	.00021	.00021	.00020	.00020	.00019
$\{2, \delta\epsilon\}$	.00001	.00001	.00001	.00001	.00001	.00001	.00002	.00002
$\{2, \delta\eta\}$	.00002	.00002	.00002	.00002	.00002	.00002	.00002	.00002
$\{2, \epsilon^2\}$	.00017	.00017	.00016	.00015	.00015	.00014	.00014	.00013
$\{3, \delta\}$	−.01368	−.01349	−.01330	−.01311	−.01293	−.01284	−.01275	−.01267
$\{3, \epsilon\}$	−.00914	−.00903	−.00892	−.00881	−.00871	−.00866	−.00861	−.00856
$\{3, \eta\}$	−.00092	−.00090	−.00088	−.00086	−.00086	−.00084	−.00084	−.00083
$\{3, \delta^2\}$	.00003	.00003	.00003	.00003	.00003	.00003	.00003	.00003
$\{3, \delta\epsilon\}$	.00001	.00001	.00001	.00001	.00001	.00001	.00001	.00001
$\{3, \epsilon^2\}$	.00002	.00002	.00002	.00002	.00002	.00002	.00002	.00002
$\{4, \delta\}$	−.00551	−.00546	−.00543	−.00538	−.00534	−.00532	−.00530	−.00528
$\{4, \epsilon\}$	−.00401	−.00399	−.00396	−.00393	−.00391	−.00390	−.00388	−.00387
$\{4, \eta\}$	−.00024	−.00024	−.00023	−.00023	−.00023	−.00023	−.00023	−.00023
$\{4, \delta\epsilon\}$	.00001	.00001	.00001	.00001	.00001	.00001	.00001	.00001
$\{5, \delta\}$	−.00276	−.00275	−.00273	−.00272	−.00271	−.00270	−.00269	−.00269
$\{5, \epsilon\}$	−.00213	−.00212	−.00211	−.00210	−.00209	−.00209	−.00208	−.00208
$\{5, \eta\}$	−.00008	−.00008	−.00008	−.00008	−.00008	−.00008	−.00008	−.00008
$\{6, \delta\}$	−.00158	−.00157	−.00157	−.00156	−.00155	−.00155	−.00155	−.00155
$\{6, \epsilon\}$	−.00126	−.00126	−.00126	−.00125	−.00125	−.00125	−.00125	−.00125
$\{6, \eta\}$	−.00003	−.00003	−.00003	−.00003	−.00003	−.00003	−.00003	−.00003

θ_0 =	.3	.25	.2	.15	.1	.05	0**
$\{0, \delta\}$	4.33216	4.93480	5.87874	7.49588	10.78515	20.74569	1.000000
$\{0, \epsilon\}$	.71097	.88889	1.14867	1.57391	2.41481	4.92154	.250000
$\{0, \eta\}$	2.48046	2.75850	3.21013	4.00081	5.62957	10.59568	.500000
$\{0, \delta^2\}$	−.20279	−.22456	−.26021	−.32296	−.45255	−.84825	−.039865
$\{0, \delta\epsilon\}$	.16054	.17498	.19957	.24378	.33621	.62017	.028683
$\{0, \delta\eta\}$	−.05432	−.05977	−.06883	−.08490	−.11824	−.22028	−.010290
$\{0, \delta^3\}$	.00089	.00098	.00112	.00138	.00192	.00357	.000166
$\{0, \epsilon^2\}$	−.15201	−.16827	−.19494	−.24186	−.33884	−.63496	−.029835
$\{0, \epsilon\eta\}$	−.00364	−.00416	−.00496	−.00630	−.00905	−.01736	−.000836
$\{0, \eta^2\}$	−.00208	−.00228	−.00262	−.00322	−.00448	− 00832	−.000388
$\{0, \delta^2\epsilon\}$	−.00144	−.00158	−.00180	−.00221	−.00305	−.00566	−.000262
$\{0, \delta^2\eta\}$	.00014	.00015	.00018	.00022	.00032	.00059	.000028
$\{0, \delta\epsilon^2\}$	.00103	.00112	.00129	.00160	.00222	.00412	.000194
$\{0, \delta\epsilon\eta\}$	−.00007	−.00007	−.00008	−.00009	−.00012	−.00023	−.000012
$\{1, \delta\}$	−.42975	−.39853	−.37126	−.34726	−.32596	−.30694	−.28987
$\{1, \epsilon\}$	−.18993	−.17778	−.16712	−.15769	−.14929	−.14176	−.13496
$\{1, \eta\}$	−.09354	−.08595	−.07934	−.07355	−.06844	−.06388	−.05980
$\{1, \delta^2\}$	.00505	.00463	.00426	.00395	.00366	.00341	.00319
$\{1, \delta\epsilon\}$	−.00149	−.00134	−.00120	−.00109	−.00099	−.00091	−.00083
$\{1, \delta\eta\}$	.00080	.00073	.00067	.00062	.00057	.00053	.00050
$\{1, \delta^3\}$	−.00001	−.00001	−.00001	−.00001	−.00001	−.00001	−.00001
$\{1, \epsilon^2\}$	.00359	.00330	.00303	.00281	.00261	.00243	.00227
$\{1, \epsilon\eta\}$	.00014	.00014	.00012	.00012	.00011	.00011	.00010
$\{1, \eta^2\}$	.00002	.00002	.00001	.00001	.00001	.00001	.00001

* Note 1. $\{n, \delta^i\epsilon^j\eta^k\}$ which are less than .00001 in magnitude, or for which $n > 6$, are not shown.

** Note 2. For $\theta_0=0$ ($y_0=\infty$) the coefficients $\frac{A_\delta^0}{y_0}$, $\frac{A_\epsilon^0}{y_0}$, $\cdots$ of $\frac{D_s}{y_0}$ are given.

TABLE I. (Continued)

θ =	.3	.25	.2	.15	.1	.05	0**
$[2, \delta]$	−.04365	−.04297	−.04232	−.04168	−.04106	−.04045	−.03987
$[2, \epsilon]$	−.02573	−.02540	−.02507	−.02475	−.02445	−.02415	−.02385
$[2, \eta]$	−.00477	−.00468	−.00458	−.00450	−.00442	−.00433	−.00425
$[2, \delta^2]$	.00019	.00018	.00018	.00018	.00017	.00017	.00017
$[2, \delta\epsilon]$	.00002	.00002	.00002	.00002	.00002	.00002	.00002
$[2, \delta\eta]$	.00002	.00002	.00002	.00002	.00002	.00002	.00002
$[2, \epsilon^2]$	.00013	.00013	.00013	.00013	.00012	.00012	.00012
$[3, \delta]$	−.01258	−.01250	−.01241	−.01233	−.01225	−.01217	−.01209
$[3, \epsilon]$	−.00851	−.00846	−.00842	−.00837	−.00832	−.00827	−.00823
$[3, \eta]$	−.00082	−.00081	−.00080	−.00079	−.00078	−.00078	−.00077
$[3, \delta^2]$	.00003	.00002	.00002	.00002	.00002	.00002	.00002
$[3, \delta\epsilon]$	.00001	.00001	.00001	.00001	.00001	.00001	.00001
$[3, \epsilon^2]$	.00002	.00002	.00002	.00002	.00002	.00002	.00002
$[4, \delta]$	−.00526	−.00524	−.00522	−.00520	−.00518	−.00516	−.00514
$[4, \epsilon]$	−.00386	−.00385	−.00383	−.00382	−.00381	−.00380	−.00378
$[4, \eta]$	−.00022	−.00022	−.00022	−.00022	−.00022	−.00022	−.00022
$[5, \delta]$	−.00268	−.00268	−.00267	−.00266	−.00266	−.00265	−.00264
$[5, \epsilon]$	−.00207	−.00207	−.00207	−.00206	−.00206	−.00205	−.00205
$[5, \eta]$	−.00008	−.00008	−.00008	−.00008	−.00008	−.00008	−.00008
$[6, \delta]$	−.00155	−.00155	−.00154	−.00154	−.00154	−.00153	−.00153
$[6, \epsilon]$	−.00124	−.00124	−.00124	−.00024	−.00124	−.00123	−.00123
$[6, \eta]$	−.00003	−.00003	−.00003	−.00003	−.00003	−.00003	−.00003

θ_0 =	−.05	−.1	−.15	−.2	−.25	−.3	−.35	−.4
$[0, \delta]$	−19.32207	−9.35137	−6.04482	−4.40273	−3.42538	−2.77963	−2.32283	−1.98371
$[0, \epsilon]$	−5.06707	−2.56220	−1.72447	−1.30380	−1.05015	−.88013	−.75802	−.66590
$[0, \eta]$	−9.46238	−4.48740	−2.84360	−2.03119	−1.55045	−1.23488	−1.01322	−.84989
$[0, \delta^2]$	.75141	.35491	.22402	.15940	.12120	.09616	.07860	.06569
$[0, \delta\epsilon]$	−.53194	−.24724	−.15352	−.10745	−.08037	−.06274	−.05042	−.04144
$[0, \delta\eta]$	.19281	.09051	.05679	.04017	.03036	.02396	.01947	.01617
$[0, \delta^3]$	−.00311	−.00146	−.00092	−.00065	−.00048	−.00038	−.00031	−.00026
$[0, \epsilon^2]$	.56222	.26550	.16755	.11919	.09061	.07188	.05874	.04908
$[0, \epsilon\eta]$	.01609	.00775	.00500	.00363	.00281	.00226	.00188	.00160
$[0, \eta^2]$	.00724	.00338	.00211	.00149	.00112	.00088	.00072	.00059
$[0, \delta^2\epsilon]$	.00488	.00227	.00141	.00098	.00074	.00059	.00047	.00039
$[0, \delta^2\eta]$	−.00048	−.00023	−.00014	−.00010	−.00008	−.00006	−.00006	−.00004
$[0, \delta\epsilon^2]$	−.00357	−.00167	−.00103	−.00073	−.00055	−.00044	−.00035	−.00029
$[0, \delta\epsilon\eta]$	.00023	.00010	.00006	.00004	.00003	.00003	.00002	.00001
$[1, \delta]$	−.27445	−.26046	−.24772	−.23607	−.22538	−.21553	−.20643	−.19800
$[1, \epsilon]$	−.12880	−.12318	−.11805	−.11332	−.10897	−.10494	−.10120	−.09772
$[1, \eta]$	−.05614	−.05282	−.04981	−.04707	−.04457	−.04227	−.04016	−.03821
$[1, \delta^2]$	.00299	.00281	.00264	.00249	.00235	.00223	.00211	.00201
$[1, \delta\epsilon]$	−.00076	−.00069	−.00064	−.00059	−.00054	−.00050	−.00047	−.00043
$[1, \delta\eta]$	.00046	.00043	.00040	.00038	.00036	.00034	.00032	.00030
$[1, \delta^3]$	−.00001	−.00001	−.00001	−.00001				
$[1, \epsilon^2]$	.00213	.00200	.00188	.00178	.00168	.00159	.00151	.00143
$[1, \epsilon\eta]$	.00009	.00009	.00008	.00008	.00007	.00007	.00007	.00006
$[1, \eta^2]$	.00001	.00001	.00001	.00001	.00001	.00001	.00001	
$[2, \delta]$	−.03929	−.03873	−.03819	−.03766	−.03714	−.03663	−.03614	−.03566
$[2, \epsilon]$	−.02357	−.02328	−.02301	−.02275	−.02249	−.02223	−.02198	−.02174
$[2, \eta]$	−.00417	−.00409	−.00401	−.00394	−.00387	−.00380	−.00373	−.00367
$[2, \delta^2]$	.00016	.00016	.00016	.00015	.00015	.00015	.00015	.00014
$[2, \delta\epsilon]$	.00002	.00002	.00002	.00002	.00002	.00002	.00002	.00002
$[2, \delta\eta]$	.00002	.00002	.00001	.00001	.00001	.00001	.00001	.00001
$[2, \epsilon^2]$	.00011	.00011	.00011	.00011	.00011	.00010	.00010	.00010
$[3, \delta]$	−.01201	−.01193	−.01185	−.01178	−.01170	−.01163	−.01155	−.01148
$[3, \epsilon]$	−.00818	−.00814	−.00809	−.00805	−.00801	−.00796	−.00792	−.00788
$[3, \eta]$	−.00077	−.00076	−.00075	−.00075	−.00074	−.00073	−.00073	−.00072
$[3, \delta^2]$	.00002	.00002	.00002	.00002	.00002	.00002	.00002	.00002
$[3, \delta\epsilon]$	.00001	.00001	.00001	.00001	.00001	.00001	.00001	.00001
$[3, \epsilon^2]$	.00002	.00002	.00001	.00001	.00001	.00001	.00001	.00001
$[4, \delta]$	−.00512	−.00510	−.00509	−.00507	−.00505	−.00503	−.00501	−.00500
$[4, \epsilon]$	−.00377	−.00376	−.00375	−.00374	−.00372	−.00371	−.00370	−.00369
$[4, \eta]$	−.00022	−.00022	−.00021	−.00021	−.00021	−.00021	−.00021	−.00021
$[5, \delta]$	−.00264	−.00263	−.00262	−.00262	−.00261	−.00261	−.00260	−.00259
$[5, \epsilon]$	−.00204	−.00204	−.00204	−.00203	−.00203	−.00202	−.00202	−.00202
$[5, \eta]$	−.00008	−.00008	−.00008	−.00008	−.00008	−.00008	−.00008	−.00008
$[6, \delta]$	−.00153	−.00153	−.00152	−.00152	−.00152	−.00152	−.00151	−.00151
$[6, \epsilon]$	−.00123	−.00123	−.00123	−.00122	−.00122	−.00122	−.00122	−.00122
$[6, \eta]$	−.00003	−.00003	−.00003	−.00003	−.00003	−.00003	−.00003	−.00003

TABLE I. (Continued)

θ_0 =	−.45	−.5	−.55	−.6	−.7	−.8	−.9	−1.0
$[0, \delta]$	−1.72274	−1.51622	−1.34910	−1.21138	−.99849	−.84237	−.72362	−.63067
$[0, \epsilon]$	−.59385	−.53588	−.48819	−.44823	−.38497	−.33706	−.29947	−.26917
$[0, \eta]$	−.72520	−.62733	−.54879	−.44863	−.38670	−.31617	−.26349	−.22300
$[0, \delta^2]$	.05584	.04813	.04195	.03691	.02924	.02374	.01965	.01652
$[0, \delta\epsilon]$	−.03464	−.02936	−.02516	−.02176	−.01666	−.01306	−.01043	−.00846
$[0, \delta\eta]$	.01368	.01172	.01016	.00888	.00695	.00559	.00457	.00380
$[0, \delta^3]$	−.00022	−.00019	−.00016	−.00014	−.00011	−.00009	−.00007	−.00006
$[0, \epsilon^2]$	.04172	.03595	.03133	.02757	.02184	.01772	.01467	.01233
$[0, \epsilon\eta]$	.00139	.00121	.00107	.00096	.00078	.00064	.00055	.00047
$[0, \eta^2]$	.00050	.00043	.00037	.00032	.00025	.00020	.00016	.00013
$[0, \delta^2\epsilon]$	.00033	.00028	.00024	.00021	.00016	.00013	.00010	.00008
$[0, \delta^2\eta]$	−.00003	−.00003	−.00002	−.00002	−.00001	−.00001	−.00001	−.00001
$[0, \delta\epsilon^2]$	−.00025	−.00021	−.00018	−.00016	−.00012	−.00010	−.00008	−.00007
$[0, \delta\epsilon\eta]$	.00001	.00001	.00001	.00001				
$[1, \delta]$	−.19017	−.18288	−.17608	−.16971	−.15815	−.14793	−.13882	−.13067
$[1, \epsilon]$	−.09448	−.09144	−.08859	−.08591	−.08102	−.07664	−.07272	−.06917
$[1, \eta]$	−.03641	−.03473	−.03318	−.03173	−.02911	−.02682	−.02480	−.02300
$[1, \delta^2]$	.00191	.00182	.00173	.00166	.00151	.00139	.00128	.00118
$[1, \delta\epsilon]$	−.00040	−.00037	−.00034	−.00032	−.00028	−.00024	−.00020	−.00018
$[1, \delta\eta]$	.00029	.00027	.00026	.00025	.00022	.00020	.00018	.00017
$[1, \epsilon^2]$	.00136	.00130	.00124	.00118	.00108	.00099	.00092	.00085
$[1, \epsilon\eta]$	.00006	.00006	.00006	.00006	.00005	.00005	.00004	.00003
$[2, \delta]$	−.03519	−.03473	−.03428	−.03384	−.03300	−.03219	−.03141	−.03067
$[2, \epsilon]$	−.02150	−.02126	−.02103	−.02081	−.02037	−.01995	−.01955	−.01917
$[2, \eta]$	−.00361	−.00354	−.00349	−.00342	−.00331	−.00320	−.00310	−.00150
$[2, \delta^2]$	.00014	.00014	.00014	.00013	.00013	.00012	.00012	.00012
$[2, \delta\epsilon]$	.00002	.00002	.00002	.00002	.00002	.00002	.00002	.00002
$[2, \delta\eta]$	.00001	.00001	.00001	.00001	.00001	.00001	.00001	.00001
$[2, \epsilon^2]$	.00010	.00010	.00010	.00009	.00009	.00009	.00008	.00008
$[3, \delta]$	−.01141	−.01134	−.01127	−.01120	−.01106	−.01093	−.01080	−.01067
$[3, \epsilon]$	−.00784	−.00779	−.00775	−.00771	−.00763	−.00755	−.00748	−.00740
$[3, \eta]$	−.00072	−.00071	−.00071	−.00070	−.00068	−.00067	−.00066	−.00065
$[3, \delta^2]$	.00002	.00002	.00002	.00002	.00002	.00002	.00002	.00002
$[3, \delta\epsilon]$	.00001	.00001	.00001	.00001	.00001	.00001	.00001	.00001
$[3, \epsilon^2]$	.00001	.00001	.00001	.00001	.00001	.00001	.00001	.00001
$[4, \delta]$	−.00498	−.00496	−.00494	−.00492	−.00489	−.00485	−.00482	−.00479
$[4, \epsilon]$	−.00368	−.00366	−.00365	−.00364	−.00362	−.00360	−.00358	−.00356
$[4, \eta]$	−.00021	−.00020	−.00020	−.00020	−.00020	−.00020	−.00020	−.00020
$[5, \delta]$	−.00259	−.00259	−.00258	−.00257	−.00256	−.00255	−.00253	−.00252
$[5, \epsilon]$	−.00201	−.00201	−.00200	−.00200	−.00199	−.00198	−.00197	−.00197
$[5, \eta]$	−.00008	−.00008	−.00008	−.00008	−.00008	−.00007	−.00007	−.00007
$[6, \delta]$	−.00151	−.00151	−.00150	−.00150	−.00150	−.00149	−.00149	−.00149
$[6, \epsilon]$	−.00122	−.00121	−.00121	−.00121	−.00121	−.00120	−.00120	−.00120
$[6, \eta]$	−.00003	−.00003	−.00003	−.00003	−.00003	−.00003	−.00003	−.00003

TABLE II. Expansion of $\mathfrak{D}$ (row 1 × row 2 × row 3, see Eq. 23). Numerical tabulation of the coefficients in row 3.

θ_0	row 1 row 2 row 3	1 D_0D_1 1	δ D_1D_2 $-y_{01}$	ϵ D_1D_3 $-y_{02}$	 D_2^2 $-y_{11}$	η D_2^2 $2y_{011}$	 D_1D_3 $2y_{012}$	$\delta\epsilon$ D_2D_3 $-4y_{0111}$	 D_1D_4 $-2y_{0112}$
.5		1	4.00000	.57143	−4.00000	16.00000	2.28571	9.14285	.26891
.45		1	4.04040	.62598	−3.30579	14.69238	2.27628	8.27740	.26623
.4		1	4.16667	.69445	−2.77778	13.88889	2.31482	7.71606	.26916
.35		1	4.39561	.78278	−2.36687	13.52495	2.40856	7.41094	.27845
.3		1	4.76190	.90090	−2.04082	13.60543	2.57400	7.35428	.29586
.25		1	5.33333	1.06667	−1.77778	14.22222	2.84445	7.58519	.32508
.2		1	6.25000	1.31579	−1.56250	15.62500	3.28948	8.22369	.37380
.15		1	7.84314	1.73160	−1.38408	18.45445	4.07436	9.58673	.46037
.1		1	11.11111	2.56410	−1.23457	24.69135	5.69800	12.66222	.64023
.05		1	21.05263	5.06329	−1.10803	44.32135	10.65956	22.44119	1.19100
0*		0	1.00000	.25000	0	2.00000	.50000	1.00000	.05556
−.05		1	−19.04762	−4.93827	− .90703	−36.28118	−9.40624	−17.91666	−1.03936
−.10		1	−9.09091	−2.43902	− .82645	−16.52891	−4.43458	−8.06287	− .48732
−.15		1	−5.79710	−1.60643	− .75644	−10.08191	−2.79379	−4.85876	− .30533
−.2		1	−4.16667	−1.19048	− .69444	−6.94444	−1.98412	−3.30686	− .21566
−.25		1	−3.20000	− .94118	− .64000	−5.12000	−1.50588	−2.40941	− .16280
−.3		1	−2.56410	− .77519	− .59171	−3.94477	−1.19260	−1.83478	− .12823
−.35		1	−2.11640	− .65681	− .54870	−3.13541	− .97306	−1.44157	− .10407
−.4		1	−1.78571	− .56818	− .51020	−2.55102	− .81169	−1.15955	− .08635
−.45		1	−1.53256	− .49938	− .47562	−2.11388	− .68880	− .95006	− .07289
−.5		1	−1.33333	− .44444	− .44444	−1.77778	− .59259	− .79012	− .06238
−.55		1	−1.17302	− .39960	− .41623	−1.51357	− .51561	− .66530	− .05399
−.6		1	−1.04167	− .36232	− .39063	−1.30208	− .45290	− .56612	− .04718
−.7		1	− .84034	− .30395	− .34602	− .98863	− .35759	− .42069	− .03687
−.8		1	− .69444	− .26042	− .30864	− .77160	− .28935	− .32150	− .02953
−.9		1	− .58480	− .22676	− .27701	− .61557	− .23869	− .25126	− .02411
−1.0		1	− .50000	− .20000	− .25000	− .50000	− .20000	− .20000	− .02000

θ_0	row 1 row 2 row 3	ϵ^2 D_1D_4 y_{0111}	 D_2D_3 $2y_{0112}$	 D_2D_4 $4y_{01112}$	$\epsilon\eta$ D_1D_5 $2y_{01113}$	 D_3^2 $2y_{01122}$	 D_2D_5 $-4y_{011113}$	$\delta\epsilon^2$ D_3D_4 $-4y_{011122}$	 D_1D_6 $-2y_{011114}$
.5		−.13445	−4.57143	1.07563	.01735	1.30612	.06941	.30732	.00071
.45		− 13312	−4.13870	96812	.01712	1.16583	.06226	.27271	.00070
.4		−.13458	−3.85803	.89721	.01726	1.07168	.05752	.24922	.00070
.35		−.13922	−3.70547	.85676	.01779	1.01520	.05475	.23498	.00072
.3		−.14793	−3.67714	.84532	.01885	.99382	.05384	.22846	.00076
.25		−.16254	−3.79260	.86688	.02064	1.01136	.05504	.23117	.00083
.2		−.18690	−4.11185	.93450	.02366	1.08206	.05915	.24592	.00095
.15		−.23019	−4.79337	1.08323	.02904	1.24503	.06834	.28136	.00117
.1		−.32011	−6.33111	1.42273	.04027	1.62336	.08948	.36480	.00162
.05		−.59550	−11.22059	2.50737	.07468	2.84066	.15722	.63478	.00300
0*		−.02778	−.50000	.11111	.00347	.12500	.00694	.02778	.00014
−.05		.51968	8.95833	−1.97974	− 06476	−2.21193	−.12335	−.48882	−.00260
−.1		.24366	4.03144	−.88604	−.03026	−.98328	−.05503	−.21610	−.00121
−.15		.15267	2.42938	−.53101	−.01891	−.58539	−.03288	−.12796	−.00075
−.2		.10783	1.65343	−.35943	−.01331	−.39367	−.02218	−.08558	−.00053
−.25		.08140	1.20470	−.26048	−.01002	−.28346	−.01603	−.06129	−.00040
−.3		.06412	.91739	−.19728	−.00787	−.21335	−.01210	−.04588	−.00031
−.35		.05203	.72078	−.15418	−.00637	−.16570	−.00943	−.03544	−.00025
−.4		.04318	.57978	−.12336	−.00527	−.13177	−.00752	−.02804	−.00021
−.45		.03644	.47503	−.10054	−.00443	−.10675	−.00611	−.02259	−.00017
−.5		.03119	.39506	−.08317	−.00378	−.08779	−.00504	−.01848	−.00015
−.55		.02700	.33265	−.06967	−.00326	−.07311	−.00421	−.01531	−.00013
−.6		.02359	.28306	−.05897	−.00284	−.06154	−.00355	−.01282	−.00011
−.7		.01843	.21035	−.04337	−.00221	−.04475	−.00260	−.00923	−.00009
−.8		.01476	.16075	−.03281	−.00176	−.03349	−.00195	−.00684	−.00007
−.9		.01206	.12563	−.02538	−.00143	−.02564	−.00150	−.00518	−.00006
−1.0		.01000	.10000	−.02000	−.00118	−.02000	−.00118	−.00400	−.00005

* Note 1. For $\theta_0=0$ ($y_0=\infty$) the coefficients of $\mathfrak{D}/y_0$ are given.

ON THE MECHANICAL BEHAVIOUR OF METALS IN THE STRAIN-HARDENING RANGE*

BY

G. H. HANDELMAN, C. C. LIN AND W. PRAGER

Brown University

1. Introduction. The present paper is concerned with certain stress-strain relations purporting to describe the mechanical behaviour of quasi-isotropic metals in the strain-hardening range. As a preparation for a more precise characterization of these relations, let us consider the tension test of a metal like copper or aluminum which *does not flow under a constant stress, but exhibits strain hardening.* If the test involves loading only, i.e., if the reduced tensile stress[1] σ or the tensile strain ϵ increase throughout the test, the resulting diagram of reduced stress versus strain will have the general appearance of the curve OPQ in Fig. 1. On the other hand, if the test specimen is unloaded after a certain point, such as P, has been reached along this curve, the stress-strain diagram for unloading is found to be very nearly a straight line PA which is parallel to the tangent of the curve OPQ at O. After complete unloading, the specimen shows a permanent extension which corresponds to the permanent strain represented by OA.

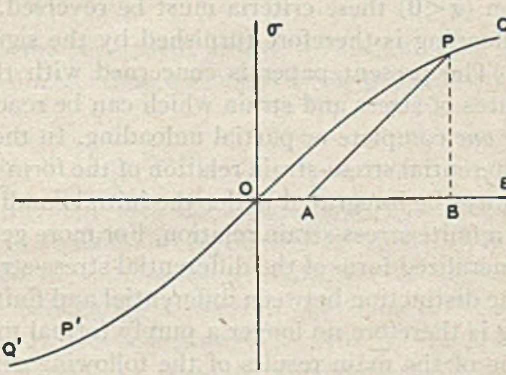

Fig. 1. Typical curve of reduced stress vs. strain.

To simplify the discussion, let us assume at present that the material is incompressible. A longitudinal extension ϵ of the isotropic specimen is then accompanied by a uniform lateral contraction of the magnitude $\epsilon/2$. If the discussion is restricted to states of stress and strain which can be reached by a *single* loading followed by *one* complete or partial unloading at the most, the mechanical behaviour of the material in simple tension is therefore completely defined by the curve OPQ. It will be assumed in the following that for the materials under consideration the stress-strain diagram in simple compression ($OP'Q'$ in Fig. 1) is obtained by reflecting the curve OPQ with respect to the origin O, and that the practically important portion of the curve $Q'OQ$, i.e., the portion corresponding to small and moderate strains, is represented with sufficient accuracy by a development of the form

$$\epsilon = \sigma + a_3\sigma^3 + a_5\sigma^5 + \cdots, \tag{1}$$

where $a_3, a_5, \cdots$ are constants. (The coefficient of the linear term on the right-hand side of (1) must be unity since σ is the reduced stress. No even powers of σ can occur

* Received September 17, 1946.

[1] The reduced stress is defined as the quotient of the stress by Young's modulus.

on the right-hand side of (1), because the stress-strain diagrams for tension and compression are assumed to be congruent.)

In the case of simple tension or compression, the mechanical behaviour of the material during the first loading is readily represented by a *finite* relation of the form (1); the behaviour during the first unloading, however, is most naturally represented by the *differential* stress-strain relation

$$d\epsilon = d\sigma, \tag{2}$$

for this form avoids explicit reference to the state of stress at which the unloading began. Accordingly, it is often convenient to write Eq. (1), too, in differential form:

$$d\epsilon = \alpha(\sigma)d\sigma. \tag{3}$$

Here, $\alpha(\sigma) = d\epsilon/d\sigma = 1 + 3a_3\sigma^2 + 5a_5\sigma^4 + \cdots$ equals the quotient of Young's modulus by the so-called *tangent modulus*. To arrive at a complete analytical description of the mechanical behaviour of the material in simple tension and compression, we must supplement the preceding equations by analytical criteria for loading and unloading. For tension ($\sigma > 0$) loading corresponds to $d\sigma > 0$ and unloading to $d\sigma < 0$; for compression ($\sigma < 0$) these criteria must be reversed. A satisfactory criterion for loading and unloading is therefore furnished by the sign of $\sigma d\sigma = d(\frac{1}{2}\sigma^2)$.

The present paper is concerned with the extension of this analysis to general states of stress and strain which can be reached by a *single* loading followed at most by *one* complete or partial unloading. In the case of simple tension or compression, a differential stress-strain relation of the form (3) which is valid for the first loading can always be integrated under the initial condition $\epsilon = 0$ for $\sigma = 0$ and is thus equivalent to a finite stress-strain relation. For more general states of stress, however, a suitably generalized form of the differential stress-strain relation (3) may be integrable or not. The distinction between differential and finite stress-strain relations for the first loading is therefore no longer a purely formal matter, but acquires physical significance. One of the main results of the following discussion consists in the remark that the assumption of a finite stress-strain relation for the first loading is incompatible with certain postulates concerning the mechanical behaviour under those changes of stress which constitute neither loading nor unloading. This is shown in Section 3. Sections 2 and 4 are devoted to the discussion of finite and differential stress-strain relations, respectively. Section 5 gives a method of correlating experimental results with the present theory. Finally, Section 6 contains a discussion of the limitations of the theory.

2. Finite stress-strain relations. Using rectangular Cartesian coordinates x_i, ($i = 1, 2, 3$), we denote the displacement from the standard state by u_i, the strain by ϵ_{ij} and the reduced stress by σ_{ij}. For the small deformations to which the following discussion is restricted, the strain ϵ_{ij} is given by

$$\epsilon_{ij} = \tfrac{1}{2}(u_{i,j} + u_{j,i}), \tag{4}$$

where $u_{i,j}$ stands for $\partial u_i/\partial x_j$, etc. Adopting the usual summation convention regarding repeated subscripts, we define the *mean normal strain* as

$$e = \tfrac{1}{3}\epsilon_{ii}, \tag{5}$$

and the *strain deviation* as

$$e_{ij} = \epsilon_{ij} - e\delta_{ij}, \tag{6}$$

where δ_{ij} is the Kronecker delta. Similarly, the *reduced mean normal stress* s and the *deviation* s_{ij} *of the reduced stress* are defined as

$$s = \tfrac{1}{3}\sigma_{ii} \tag{7}$$

and

$$s_{ij} = \sigma_{ij} - s\delta_{ij}. \tag{8}$$

According to the definitions of the deviations e_{ij} and s_{ij}, we have

$$e_{ii} = 0, \qquad s_{ii} = 0. \tag{9}$$

The task of generalizing the finite stress-strain relation (1) is simplified by the remark that the first term on the right-hand side represents that part of the total strain ϵ which is recovered upon *complete* unloading. The remaining terms on the right-hand side of (1) accordingly represent the permanent strain. In Fig. 1 the *total* strain is represented by the segment OB, the *recoverable* strain by AB, and the *permanent* strain by OA.

Setting

$$\epsilon_{ij} = \epsilon'_{ij} + \epsilon''_{ij}, \tag{10}$$

where ϵ'_{ij} denotes the recoverable and ϵ''_{ij} the permanent strain, we may assume that the recoverable strain is related to the reduced stress by means of the generalized law of Hooke:

$$\epsilon'_{ij} = (1 + \nu)s_{ij} + (1 - 2\nu)s\delta_{ij}. \tag{11}$$

Here ν denotes Poisson's ratio. We are then left with the task of supplementing (11) by a relation which expresses the permanent strain occurring during the first loading in terms of the reduced stress. For an *isotropic* material, this relation can only contain scalar constants in addition to the tensors ϵ''_{ij}, σ_{ij} and δ_{ij}, and their invariants. Furthermore, the principal axes of ϵ''_{ij} and σ_{ij} must coincide. Under the pressures commonly encountered in the testing of materials, no permanent change of volume is observed, i.e., $\epsilon''_{ii}=0$ and $\epsilon''_{ij}=e''_{ij}$. A state of hydrostatic pressure therefore does not produce any permanent strain, and two states of stress which differ only by a state of hydrostatic pressure may be expected to produce identical permanent strains. The permanent strain ϵ''_{ij} is thus independent of s and depends only on the deviation s_{ij}. Furthermore, if the stress-strain diagrams for simple tension and simple compression are congruent, a reversal of the signs of all stresses may be expected to produce a mere reversal of the signs of all principal strains. Finally, if the ratios of the principal stresses are kept constant during the loading process, the ratios of the principal permanent strains, too, can be expected to remain constant.

In a recent paper,[2] W. Prager established the most general stress-strain relation which is compatible with the preceding postulates. With the notations

$$J_2 = \tfrac{1}{2}s_{ij}s_{ji}, \qquad J_3 = \tfrac{1}{3}s_{ij}s_{jk}s_{ki}, \tag{12}$$

and

$$t_{ij} = s_{ik}s_{kj} - \tfrac{2}{3}J_2\delta_{ij}, \tag{13}$$

Prager's stress-strain relation can be written in the form

[2] W. Prager, *Strain-hardening under combined stresses*, J. Appl. Phys. **16**, 837–840 (1945).

$$\epsilon_{ij}'' = F(J_2, J_3^2)[P(J_2, J_3^2)s_{ij} + Q(J_2, J_3^2)J_3 t_{ij}], \tag{14}$$

where P and Q must be homogeneous in the components of the stress deviation, the degree of P exceeding that of Q by 4. The expressions (12) are second and third order invariants of the stress deviation s_{ij} (the first order invariant s_{ii} vanishes). The tensor (13) is the deviation of the square $s_{ik}s_{kj}$ of the stress deviation s_{ij}.

Combining (11) and (14), we obtain the desired generalization of the finite stress-strain relation (1):

$$\epsilon_{ij} = (1 + \nu)s_{ij} + (1 - 2\nu)s\delta_{ij} + F(J_2, J_3^2)[P(J_2, J_3^2)s_{ij} + Q(J_2, J_3^2)J_3 t_{ij}]. \tag{15}$$

3. Neutral changes of stress. Inadmissibility of finite stress-strain relations. In the case of simple tension or compression the sign of $\sigma d\sigma = d(\frac{1}{2}\sigma^2)$ proved to be a satisfactory criterion for loading and unloading. Accordingly, one might consider the possibility of using the sign of $\sigma_{ji}d\sigma_{ji}$ as a criterion in the general case. If, however, the term "loading" is reserved for such changes of stress which are accompanied by a change of the permanent strain, this criterion is not satisfactory. Indeed, on account of (8) and the second Eq. (9), we have

$$\sigma_{ij}d\sigma_{ij} = (s_{ij} + s\delta_{ij})(ds_{ij} + ds\delta_{ij}) = s_{ij}ds_{ij} + 3sds. \tag{16}$$

If loading were to correspond to $\sigma_{ij}d\sigma_{ij} > 0$, a change of stress for which $ds_{ij} = 0$ might therefore constitute loading in spite of the fact that such a change of stress is not accompanied by a change of the permanent strain. To avoid this difficulty, we shall use the sign of $s_{ij}ds_{ij} = dJ_2$ as the desired criterion, an increase of J_2 corresponding to loading, a decrease to unloading.

Whereas for uniaxial stress any change of stress constitutes either loading or unloading, we have three kinds of change of stress in the general case, according to whether J_2 increases, remains constant, or decreases. An infinitesimal change of stress for which $dJ_2 = 0$, will be called a *neutral* change of stress. For instance, any change of stress which affects only the mean normal stress, but leaves the stress deviation untouched, is a neutral change of stress. A more interesting example of a neutral change of stress is given by

$$\sigma_{ij} = \begin{Vmatrix} \sigma & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{Vmatrix}, \qquad d\sigma_{ij} = \begin{Vmatrix} 0 & d\tau & 0 \\ d\tau & 0 & 0 \\ 0 & 0 & 0 \end{Vmatrix}. \tag{17}$$

Equation (17) represents the stress system which arises from a combined tension and torsion test of a thin walled circular cylinder. Specifically, consider such a test piece which is pulled to an arbitrary tensile stress σ. If the traction is then kept constant and a small torque applied, the resulting systems of stress and increments of stress are represented by Eq. (17).

Let us now suppose that for the first loading ($dJ_2 > 0$) we have the finite stress-strain relation (15) and for unloading ($dJ_2 < 0$) the generalized-law of Hooke in the differential form

$$d\epsilon_{ij} = (1 + \nu)ds_{ij} + (1 - 2\nu)ds\delta_{ij}. \tag{18}$$

The simultaneous use of the stress-strain relations (15) and (18) will lead to obvious difficulties, unless these relations give identical strain increments for neutral changes of stress. We shall show that, in general, this *continuity condition* is not fulfilled. Indeed, if (15) is written in differential form, the first two terms on the right-hand side equal the right-hand side of (18); the continuity condition therefore requires the vanishing of the remaining terms on the right-hand side of the differential form of Eq. (15). Consider, for example, the stress and increment of stress given by Eq. (17). A simple computation will show that

$$s_{ij} = \begin{Vmatrix} \frac{2}{3}\sigma & 0 & 0 \\ 0 & -\frac{1}{3}\sigma & 0 \\ 0 & 0 & -\frac{1}{3}\sigma \end{Vmatrix}, \qquad t_{ij} = \begin{Vmatrix} \frac{2}{9}\sigma^2 & 0 & 0 \\ 0 & -\frac{1}{9}\sigma^2 & 0 \\ 0 & 0 & -\frac{1}{9}\sigma^2 \end{Vmatrix},$$

$$ds_{ij} = \begin{Vmatrix} 0 & d\tau & 0 \\ d\tau & 0 & 0 \\ 0 & 0 & 0 \end{Vmatrix}, \qquad dt_{ij} = \begin{Vmatrix} 0 & \frac{1}{3}\sigma d\tau & 0 \\ \frac{1}{3}\sigma d\tau & 0 & 0 \\ 0 & 0 & 0 \end{Vmatrix},$$

as well as $dJ_2 = dJ_3 = 0$. For this special case the differential form of Eq. (14) reduces to

$$d\epsilon_{ij}'' = F(J_2, J_3^2)\left[P(J_2, J_3^2)ds_{ij} + Q(J_2, J_3^2)J_3\, dt_{ij}\right], \tag{19}$$

where J_2 and J_3 are evaluated for an arbitrary state of pure tension. Since this state of stress satisfies the condition for a neutral change of stress ($dJ_2 = 0$), $d\epsilon_{ij}''$ must vanish. We find then, upon substituting the values of ds_{ij} and dt_{ij} previously computed, that

$$F(J_2, J_3^2)\left[P(J_2, J_3^2) + \tfrac{1}{3}Q(J_2, J_3^2)J_3\, \sigma\right] = 0. \tag{20}$$

Now let us return to the finite stress-strain relation, Eq. (15), for the case of pure tension. The first component of the strain tensor (the other non-vanishing terms differ from this only by a constant factor) becomes

$$\epsilon_{11} = \sigma + \tfrac{2}{3}F(J_2, J_3^2)\left[P(J_2, J_3^2)\sigma + \tfrac{1}{3}Q(J_2, J_3^2)J_3\sigma^2\right]. \tag{21}$$

The invariants appearing in Eq. (20) have been evaluated for an arbitrary state of pure tension. Consequently, Eq. (20) is valid for pure tension and the second term in Eq. (21) equals zero. (A similar remark holds true for each of the other non-vanishing strain components.) Therefore, the stress-strain relation will reduce to Hooke's law for pure tension if the continuity condition is to be fulfilled. On the other hand, we have seen in Section 1 that the stress-strain law for tension need not be linear. Thus the most general *finite* stress-strain law coupled with Hooke's law for unloading will not be sufficiently flexible to represent a tensile test if the continuity condition is to be fulfilled. It is necessary, therefore, to turn to differential stress-strain relations if both loading and unloading are to be adequately represented.

4. Differential stress-strain relations. A system of differential stress-strain relations can be obtained from the properties discussed in Section 2 provided certain of these are rewritten in such a way as to be directly applicable in differential form. We shall assume that given the components of the stress tensor σ_{ij} and the increments $d\sigma_{ij}$ there correspond unique strain increments $d\epsilon_{ij}$. This implies that the increment in strain, $d\epsilon_{ij}$, depends only on the state of stress at the given instant, σ_{ij}, and the increment in stress, $d\sigma_{ij}$, and is independent of the way in which this state of stress has been achieved provided only loading has taken place. In particular, we shall assume that this dependence is such that the increments in strain are linear functions of the increments in stress. Thus $d\epsilon_{ij}$ can be written in the form

$$d\epsilon_{ij} = (1 + \nu)ds_{ij} + (1 - 2\nu)ds\delta_{ij} + c_{ijkl}\, d\sigma_{kl}, \tag{22}$$

where the fourth order tensor c_{ijkl} is a function of σ_{ij} only. For unloading, the material is assumed to satisfy the differential form of Hooke's law given in Eq. (18). Loading is supposed to take place when $dJ_2 > 0$ and unloading occurs for $dJ_2 < 0$. For a neutral change of stress, $dJ_2 = 0$, the continuity condition requires that Eqs. (18) and (22) coincide. Consequently,

$$c_{ijkl}\, d\sigma_{kl} = 0 \quad \text{whenever} \quad dJ_2 = s_{kl}\, ds_{kl} = 0.$$

Since s_{kl} is a deviator, dJ_2 may also be written in the form $dJ_2 = s_{kl}\, d\sigma_{kl}$. Thus the linear form in $d\sigma_{kl}$, $c_{ijkl}\, d\sigma_{kl}$, must vanish whenever $s_{kl}\, d\sigma_{kl}$ vanishes. The coefficients of $d\sigma_{kl}$ in the two forms must be proportional or

$$c_{ijkl} = C_{ij}s_{kl},$$

where the second order tensor C_{ij} is a function of σ_{kl} alone. The stress-strain relations then become

$$\left.\begin{aligned} d\epsilon_{ij} &= (1 + \nu)ds_{ij} + (1 - 2\nu)ds\delta_{ij} + C_{ij}dJ_2, && \text{when} \quad dJ_2 \geqq 0; \\ d\epsilon_{ij} &= (1 + \nu)ds_{ij} + (1 - 2\nu)ds\delta_{ij}, && \text{when} \quad dJ_2 \leqq 0. \end{aligned}\right\} \tag{23}$$

In a certain sense, the term C_{ij} measures the permanent deformation. Indeed, let us consider the infinitesimal cycle of stress which results when first $d\sigma_{ij}$ is applied and then $-d\sigma_{ij}$. We assume, in addition, that the material is being loaded when $d\sigma_{ij}$ is applied. The permanent increment in strain $d\epsilon_{ij}''$ will then be

$$d\epsilon_{ij}'' = C_{ij}\, dJ_2. \tag{24}$$

Since the permanent strain is independent of a state of hydrostatic stress for pressures within the range normally encountered in testing of materials, the tensor C_{ij} can only be a function of the components of the stress deviator rather than the stress tensor itself. Furthermore, there can be no permanent change in volume; that is, $d\epsilon_{ii}'' = 0$ or $C_{ii} = 0$. Since the tensors $d\epsilon_{ij}$, ds_{ij}, and δ_{ij} are symmetric, C_{ij} will also be symmetric. In addition, a reversal of the signs of all the stresses is assumed to produce a reversal of sign of all the strain increments. This implies that C_{ij} must be an odd function of the stress components and thus will vanish when all the s_{ij} vanish.

The material is supposed to become orthotropic under the stress σ_{ij} in the sense that the C_{ij} can be represented as a power series in the stress deviator s_{ij} with *scalar* coefficients. These coefficients are either constants or else functions of the invariants of s_{ij}, i.e., functions of J_2 and J_3. It is convenient at this point to change from the subscript notation for tensors to Gibbs' notation; the tensor C_{ij} will be denoted by **C** and s_{ij} by **S**. The multiplications indicated below are the usual matrix multiplications. Under the assumptions stated above, the tensor **C** can be written as

$$\mathbf{C} = \sum_{n=0}^{\infty} a_{2n+1}(J_2, J_3)\mathbf{S}^{2n+1}. \tag{25}$$

We note that only odd powers appear in Eq. (25) since **C** is assumed to be an odd function of the stresses. Equation (25) can be simplified further by the Hamilton-Cayley theorem which states that the tensor S must satisfy its own characteristic equation.[3] For the stress deviator S, this implies that

$$S^3 = J_2 S + J_3 I, \tag{26}$$

where I is the unit tensor. Through Eq. (26), we can reduce[4] any power of S greater than the second to a linear combination of I, S, and S^2 with coefficients which are functions of J_2 and J_3. For example, consider the reduction of the power S^5. According to Eq. (26),

$$S^4 = J_2 S^2 + J_3 S;$$

thus

$$S^5 = J_2 S^3 + J_3 S^2 = J_3 S^2 + J_2^2 S + J_2 J_3 I.$$

In general, we can rewrite Eq. (25) as

$$C = a(J_2, J_3)S^2 + b(J_2, J_3)S + c(J_2, J_3)J_3 I.$$

We recall that $C_{ii}=0$; since S is a deviator, this implies that

$$2a(J_2, J_3)J_2 + 3c(J_2, J_3)J_3 = 0,$$

or

$$c(J_2, J_3) = -\frac{2J_2}{3J_3} a(J_2, J_3).$$

Consequently,

$$C = a(J_2, J_3)[S^2 - \tfrac{2}{3}J_2 I] + b(J_2, J_3)S.$$

The expression appearing in square brackets is just the tensor t_{ij} which was defined in Eq. (13). Returning now to the subscript notation we can write the tensor C_{ij} as

$$C_{ij} = a(J_2, J_3)t_{ij} + b(J_2, J_3)s_{ij}.$$

A further simplification can be made by noting that C_{ij} must be an odd function of the stress components. Since J_2 is even, J_3 odd, t_{ij} even, and s_{ij} odd, we must have

$$C_{ij} = p(J_2, J_3^2)s_{ij} + q(J_2, J_3^2)J_3 t_{ij}.$$

Thus the complete differential stress-strain relations can be written in the form[5]

$$\left.\begin{aligned} d\epsilon_{ij} &= (1+\nu)ds_{ij} + (1-2\nu)ds\delta_{ij} + [p(J_2, J_3^2)s_{ij} + q(J_2, J_3^2)J_3 t_{ij}]dJ_2 && \text{when } dJ_2 \geqq 0 \\ d\epsilon_{ij} &= (1+\nu)ds_{ij} + (1-2\nu)ds\delta_{ij}, && \text{when } dJ_2 \leqq 0. \end{aligned}\right\} \tag{27}$$

5. Further study of the stress strain relations. Experimental determination. In this section we shall discuss the relation between the differential form of the stress-strain relation (cf. (27))

---

[3] M. Bôcher, *Introduction to higher algebra*, The Macmillan Co., New York, 1907, p. 296.

[4] This technique has been used recently by Marcus Reiner, Am. J. Math. 67, 350–362 (1945) and W. Prager, *loc. cit.*

[5] These relations contain, as special cases, the stress-strain laws developed by W. Prager, Proc. Fifth International Congress of Applied Mechanics, Cambridge, Mass., 1938, pp. 234–237, and by J. H. Laning in an unpublished paper (1942).

$$d\epsilon_{ij}'' = \{p(J_2, J_3^2)s_{ij} + q(J_2, J_3^2)J_3t_{ij}\}dJ_2, \quad dJ_2 \geqq 0 \tag{28}$$

and the integral form

$$\epsilon_{ij}'' = F(J_2, J_3^2)[P(J_2, J_3^2)s_{ij} + Q(J_2, J_3^2)J_3t_{ij}], \tag{14}$$

which holds only when the ratios of the principal stresses are kept constant during the loading process, i.e., if

$$s_{ij} = ks_{ij}^{(0)}, \tag{29}$$

where $s_{ij}^{(0)}$ is fixed while k is the scalar variable. We shall then show how a series of tests necessary to establish the Lode diagram will be sufficient to determine the stress-strain relations completely. First of all, it is convenient to bring out the homogeneity properties in the relations (28) and (14) by introducing the symbols

$$\alpha = J_3^2/J_2^3, \qquad \gamma_{ij} = J_3\, t_{ij}/J_2^2, \tag{30}$$

where α is dimensionless, while γ_{ij} has the same dimensions as J_2. The relation (14) can be written in the form

$$\epsilon_{ij}'' = \lambda(J_2, \alpha)\{s_{ij} + \beta(\alpha)\gamma_{ij}\}, \tag{31}$$

where

$$\lambda(J_2, \alpha) \equiv F(J_2, J_3^2)P(J_2, J_3^2), \qquad \beta(\alpha) \equiv J_2^2Q(J_2, J_3^2)/P(J_2, J_3^2). \tag{32}$$

Note that β is independent of J_2, because of the homogeneity relation between P and Q established in Section 2.

With a similar change of notation, the relation (28) can be written as

$$d\epsilon_{ij}'' = G(J_2, \alpha)\{s_{ij} + \beta'(\alpha)\gamma_{ij}\}dJ_2, \qquad dJ_2 \geqq 0 \tag{33}$$

where

$$G(J_2, \alpha) \equiv p(J_2, J_3^2), \qquad \beta'(\alpha) \equiv J_2^2q(J_2, J_3^2)/p(J_2, J_3^2). \tag{34}$$

Since we did not establish a homogeneity relation between p and q, we cannot immediately conclude that β' is independent of J_2. However, we shall see immediately that this is true and that indeed

$$\beta' \equiv \beta. \tag{35}$$

We shall also show that $G(J_2, \alpha)$ may be obtained from $\lambda(J_2, \alpha)$ by the relation

$$G(J_2, \alpha) = \frac{\lambda}{2J_2} + \frac{\partial\lambda}{\partial J_2}. \tag{36}$$

The relations (35) and (36) will then determine the differential relation (33) *completely* once the integral relation (31) is known by a series of experiments of the *special* type (29). It is to be noted that the functions $G(J_2, \alpha)$ and $\beta(\alpha)$ in (33) determined through the use of (31) will by no means restrict the application of (33) to processes connected in any manner with (29).

To establish the relations (35) and (36), consider the application of (31) and (33) to a process of the type (29). Let $d\epsilon_{ij}''$ be the change in ϵ_{ij}'' corresponding to a change dk. Then

$$dJ_2 = 2J_2 \frac{dk}{k}, \qquad d\alpha = 0; \tag{37}$$

and (31) gives

$$d\epsilon_{ij}'' = \frac{\partial \lambda}{\partial J_2} dJ_2 \{s_{ij} + \beta \gamma_{ij}\} + \lambda \{s_{ij} + \beta \gamma_{ij}\} \frac{dk}{k},$$

while (33) yields

$$d\epsilon_{ij}'' = G(J_2, \alpha) \{s_{ij} + \beta' \gamma_{ij}\} dJ_2.$$

Equating coefficients of s_{ij} and γ_{ij}, we obtain the relations (35) and (36).

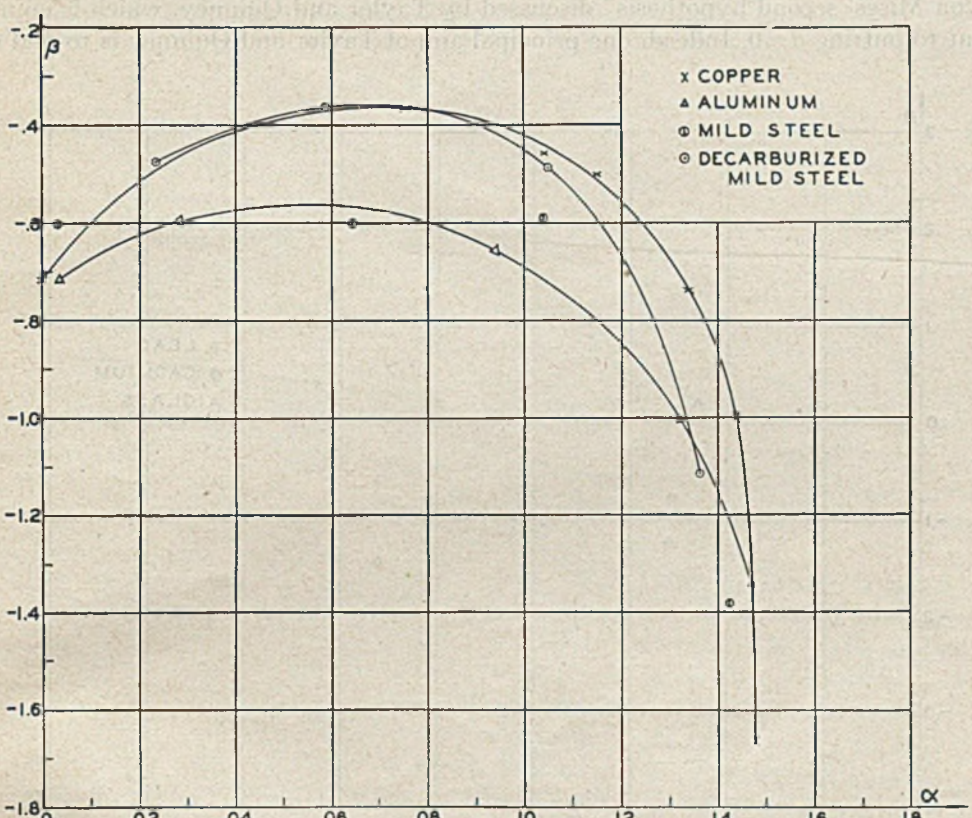

FIG. 2a. The $\alpha - \beta$ diagram (Eq. (38)) of the experimental results of Taylor and Quinney for copper, aluminum, mild steel and decarburized mild steel. The data for mild steel are too scattered for a definite curve to be drawn.

The experimental determination of the stress strain relations can then be reduced to that of (31) alone. This can be done by a series of tests of the type (29), which is of the class described by Lode,[6] Taylor-Quinney[7] and Hohenemser-Prager.[8] Indeed,

[6] W. Lodge, Forschungsarbeiten a.d. Gebiete d. Ingenieurwesens, No. 303, VDI-Verlag, Berlin, 1928.

[7] G. I. Taylor and H. Quinney, Phil. Trans. Roy. Soc. London (A) **230**, 323–362 (1931).

[8] K. Hohenemser and W. Prager, Z. angew. Math. Mech. **12**, 1–14 (1932). An English translation of this paper is available as R.T.P. Translation No. 2468 (Durand Reprinting Committee, in care of California Institute of Technology, Pasadena 4, Calif.).

the relation $\beta(\alpha)$ is merely another presentation of Lode's diagram. It can be easily verified that α, β are related to Lode's parameters[9] μ and ν by the relations

$$\left.\begin{aligned} \alpha &= \frac{4}{27}\,\frac{\mu^2(9-\mu^2)^2}{(3+\mu^2)^3}, \\ \beta &= \frac{9(3+\mu^2)^2}{2(9-\mu^2)}\,\frac{1-\nu/\mu}{\mu^2(1+2\nu/\mu)-3}. \end{aligned}\right\} \tag{38}$$

This new system has the advantage that β gives directly the extent of deviation from "von Mises' second hypothesis" discussed by Taylor and Quinney, which is equivalent to putting $\beta\equiv 0$. Indeed, one principal aim of Taylor and Quinney is to find out

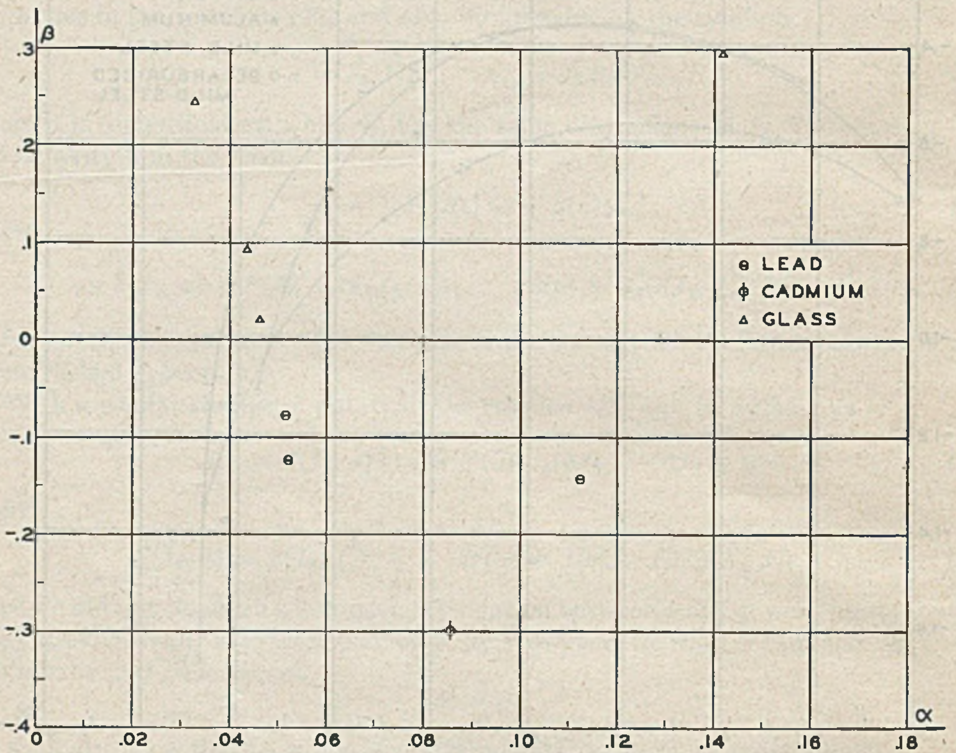

FIG. 2b. The $\alpha-\beta$ diagram (Eq. (38)) of the experimental results of Taylor and Quinney for lead, cadmium and glass. The data are too few to allow any curve to be drawn.

this extent and is therefore to determine the value of β. Figs. 2(a) and 2(b) show the results of Taylor and Quinney converted into the (α, β) diagram. This diagram reveals any experimental error more strongly, since β is essentially related to the slope of the (μ, ν) curve; e.g.,

---

[9] W. Lode, *loc. cit.*, pp. 1 and 12.

$$\frac{d\nu}{d\mu} = \frac{9 - 4\beta}{9 + 2\beta} \quad \text{for} \quad \mu \to 1,$$

$$\frac{d\nu}{d\mu} = 1 + \frac{2\beta}{3} \quad \text{for} \quad \mu = 0.$$

It should be noted in passing that α must satisfy the inequality

$$0 < \alpha < 4/27 \tag{39}$$

to insure real values of μ.

Having determined $\beta(\alpha)$ from (38), we may determine $\lambda(J_2, \alpha)$ by noting that (cf. (31))

$$I_2 \equiv \tfrac{1}{2}\epsilon_{ij}''\epsilon_{ij}'' = \lambda^2 J_2\left\{1 + 3\alpha\beta + \tfrac{1}{3}\alpha^2\beta^2\right\}. \tag{40}$$

For each loading process given by (29) the value of α is fixed, and (40) gives the dependence of λ^2 on J_2 if I_2 is determined for given values of J_2. A series of tests with different principal axes will then give the further dependence of λ on α.

6. **Concluding remarks.** In closing, we note some of the limitations of the stress-strain relations developed in this paper. It has been pointed out previously that these equations have been developed to cover the case of one loading followed by at most *one* unloading. This restriction is quite essential, for relations (27) are not applicable for a second loading. For example, if we consider a simple tensile test, the stress-strain diagram obtained from (27) for the second loading would be a mere translation of the diagram for the first loading. This does not agree with the experimental results. Secondly, we note that these equations apply only to metals which exhibit strain-hardening. They are not applicable, for example, to materials which yield under constant shearing stress or satisfy von Mises' yield condition, $J_2 = \text{const.}$

It is hoped that the results presented here will provoke experimental work to test their validity. Among the various features which should be tested are two assumptions made in developing the differential stress-strain laws. The first hypothesis (cf. Section 4) states that the increments in strain are uniquely determined by the components of the stress tensor σ_{ij} and the increments $d\sigma_{ij}$ without reference to the previous history of loading provided only one loading has taken place followed by at most one unloading. The range in which such a hypothesis is valid should be explored empirically. Secondly, the assumption involved in the transition from Eqs. (24) to (25) should be examined carefully. According to these two relations, the principal axes of the increment in permanent strain $d\epsilon_{ij}''$ will coincide with the principal axes of the existing state of stress s_{ij} independent of the increments in stress $d\sigma_{ij}$, provided only loading takes place. This conclusion should be tested by experiment.