PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A
MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

On a System of Functional Dynamics and Optics

W. H. Watson

Phil. Trans. R. Soc. Lond. A 1937 **236**, 155-190
doi: 10.1098/rsta.1937.0001

VI—On a System of Functional Dynamics and Optics

By W. H. WATSON, *McGill University*

(Communicated by J. S. FOSTER, F.R.S.—Received 30 March, 1936)

With a view to developing a transformation theory of electromagnetism, in a former paper,* the writer has outlined a theory of the motion of lines. The procedure adopted was to regard MAXWELL's equations (by which the space and time derivates of the field components are connected) as the analogue of the system of equations :

$$\frac{\partial p_x}{\partial t} = -\frac{\partial H}{\partial x}, \qquad \frac{\partial p_x}{\partial y} = \frac{\partial p_y}{\partial x}, \text{ etc.}, \qquad (1)$$

by which are connected the derivates of the momentum p_x, p_y, p_z, and energy H of a particle specified at each point (x, y, z, t) of space-time by means of a given solution of the corresponding HAMILTON-JACOBI equation. Thus a particular e.m. field for which **E** and **H** are specified at each point of space and time is to be regarded as belonging to a system of fields defined by the form of a function of the field components and of x, y, z, t, analogous to the Hamiltonian function in dynamics. This is equivalent to representing the field by means of a system of motions of a line, and is possible because of the form of Maxwell's equations for space in the absence of matter, just as it is possible to represent by means of a system of motions of a particle any vector field whose curl vanishes everywhere as in (1). However, it is not on account of the possible physical interest of the motion of lines *per se* that it is of value to study them, but because the mathematical scheme, which systematically deals with the dynamics of lines and with the corresponding extension of optics, appears to have the proper generality to embrace a transformation theory of electromagnetism, and thereby open the way for introducing the electronic charge as a modulus of discontinuity in the theory.

In the motion of a particle its coordinates depend on one independent variable, which is the time in Newtonian dynamics, whereas in the motions which we are to consider the coordinates x_1, x_2, x_3, x_4 $(x_4 = ict)$ of a point of the line in 4-dimensional space are now taken to be functions of two independent variables u and v. The purpose of this communication is to investigate in greater detail than in the paper referred to the kinematics and dynamics of moving lines (or circuits) and to develop the optics of functional waves in the hope that physical form may ultimately be substituted for mathematical formalism.

* 'Trans. Roy. Soc. Can.,' vol. 28, Sect. III, pp. 1–27 (1934), referred to as I.

This paper is divided into five sections, in the first of which kinematical aspects of the subject are discussed in some detail. In § 2 some additions are made to the matter on dynamics contained in I (§ 4), and in § 3 the corresponding optical system is considered. This latter study raises questions which lead to the reconsideration of dynamics ; in § 4 an attempt is made to construct a system in which the dynamical variables are taken to be functionals* not necessarily connected by linear relations as they are in § 2. The present treatment of these matters cannot be regarded as other than elementary, and even on this level is certainly far from complete. The mathematical structure which is common to the theory of optics and the transformation theory of dynamics appears in essence as the extension of PFAFF's problem, and is not simple. When one reflects on the extensive literature on PFAFF's problem and on the related theories of Pfaffian systems and of contact transformations, one is impressed by the possibility of analogous development relative to the extended problem and of the application of this development to physics.

In § 5 attention is once again directed to the possible connexion of functional dynamics and optics with electricity, and to the method of representing electromagnetic fields used in I to suggest a proper form of transformation theory of electromagnetic fields. It is shown how BORN's theory of a limiting electric field may be quite naturally associated with the invariance of the velocity of the functional waves analogous to light, when the functional variables undergo linear transformation. A physical explanation is offered of the independent variables u and v and it is shown that just as quaternions are appropriate to the discussion of special relativity so biquaternions are a suggestive means of expressing BORN's hypothesis. It is concluded that the idea of a transformation theory of electromagnetic fields may be combined with any theory of the same type as BORN's.

1—KINEMATICS

An electromagnetic field is explored, either by the motion of charged particles through it or by means of a secondary circuit, which for practical reasons is usually a closed one. On account of this and because of the mathematical arguments advanced in I,† we shall deal with the motion of a closed line, and unless qualification is made to the contrary, we shall, in dealing with moving lines, understand that they are closed ones. When u and v are varied, in accordance with the motion, the possible changes in x_1, x_2, x_3, x_4 must satisfy the Pfaffian system

$$\sum_s (x_r, x_s)\, dx_s = 0 \qquad (r, s = 1, 2, 3, 4) \qquad \text{where} \qquad (x_r, x_s) = \frac{\partial\,(x_r, x_s)}{\partial\,(u, v)}\,. \quad (2)$$

Of the equations (2) only two are independent, for

$$(x_1, x_2)\,(x_3, x_4) + (x_2, x_3)\,(x_1, x_4) + (x_3, x_1)\,(x_2, x_4) = 0, \quad . \quad . \quad . \quad (3)$$

* *Cf.* VOLTERRA, " Theory of Functionals," pp. 163–165, where a different system is considered.

† Especially pp. 14–15.

the Jacobians being proportional to the direction-cosines of the surface element at (x_1, x_2, x_3, x_4) generated by the motion of the line. These quantities form a six-vector and are the analogues of the components of velocity in the motion of a point; when all of them vanish the x's depend on one variable only, the line does not generate a two-way manifold by its motion. This analogy between the Jacobians and the components of the velocity of a particle deserves further examination; it must be based on some sort of notion of the distance between two closed lines. (*a*) The distance between the centres of mass of the lines (supposed uniformly loaded) is clearly a measure of the separation of the lines in space, and this measure can be represented as a 4-vector. It is, however, really the distance between two points, and the same device will serve for measuring the distance between any two loci in hyperspace, be they linear or not. (*b*) For two figures in a plane, MAXWELL* introduced the conception of the geometrical mean distance R defined by

$$\log R = \iiiint \log r_{AB}\, dx\, dy\, dx'\, dy' \Big/ \iint dx\, dy \,.\, \iint dx'\, dy',$$

where r_{AB} is the distance between A, a point of the part of the plane enclosed by the boundary of the first figure, and B lying within the boundary of the second. This idea does not admit of generalization to lines in 3- or 4-space unless we are given the surfaces (2-dimensional manifolds) bounded by the lines in question over which we are to make the integrations, and this, of course, means that the distance so defined would be that between the portions of the two surfaces. On the other hand, one could apply MAXWELL's conception by replacing the integrals over the surfaces by line integrals, thus writing

$$\log R\,(L_1 L_2) = \iint \log r\, ds_1\, ds_2 \Big/ \int ds_1 \,.\, \int ds_2.$$

This definition is obviously just a formal variant of (*a*) which employed the arithmetical mean distance and is not suited for our purpose. It should be noted that both of these means give the distance as a functional. They have been mentioned in order to emphasize that perfectly definite conceptions have been considered for the measurement of distance between elements of space which are not mere points. To be physically useful, our ideas must escape from the bewildering generality of functional spaces. It is therefore desirable to go back to some form of expression for the distance between two points which will serve as a suitable pattern for the generalization to the case of lines. This is found, of course, in the form by which the components of the vector distance between two points P and Q are defined in terms of integrals along any simple curve C joining P and Q; thus

$$x_P - x_Q = \int_Q^P \frac{dx}{ds}\, ds = \int_Q^P dx, \quad \ldots \ldots \ldots \quad (4)$$

* Treatise, vol. II, p. 324.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

while the actual scalar distance can be put in the form

$$\int_Q^P \sum_r \frac{dx_r}{ds}\, dx_r \Big/ \sqrt{\sum_r \left(\frac{dx_r}{ds}\right)^2}, \qquad (5)$$

the number of values taken by r being equal to the number of dimensions of the space.

Quite analogous to these are the definitions of the components of the 6-vector **A** (in 3 dimensions 3-vector) as integrals along any tubular surface S which joins the lines L_1 and L_2, namely, the projections of S on the six coordinate planes, *e.g.*,

$$A_{yz} = \iint_S dy\, dz, \qquad (6)$$

integrated over the portion of S bounded by L_1 and L_2. Now if Σ_1 and Σ_2 are any two simple surfaces bounded by L_1 and L_2 respectively, on account of the fact that S, Σ_1, and Σ_2 make a closed surface, $\mathbf{A}_{yz} = (\Sigma_2)_{yz} - (\Sigma_1)_{yz}$ is the difference of a function of the line L_1 and a function of the line L_2. We shall call **A** the 6-distance between the lines L_1 and L_2. It is clearly not a measure of the separation of the lines in the ordinary sense for if L_1 can be transformed into L_2 by a simple translation, then each of the components of the 6-distance between L_1 and L_2 vanishes. On the other hand, the functional which corresponds to (5), namely

$$|\mathbf{A}| = \iint \frac{\Sigma\, (x_r, x_s)\, dx_r dx_s}{\sqrt{\Sigma\, (x_r, x_s)^2}}, \qquad (7)$$

which is the area of the tube surface connecting L_1 and L_2, does not vanish under the same circumstances. Conversely, it is possible for $|\mathbf{A}|$ to vanish when none of the components of the vector is equal to zero. The functional expressed by the double integral (7) will depend on the surface S which is chosen to connect L_1 and L_2, just as the integral (5) depends on the curve C chosen to connect the points P and Q. The scalar $|\mathbf{A}|$ can be made unique by choosing S to be the analogue of the straight line (or geodesic), that is, S must be a minimal surface. This connexion between minimal surfaces and linear functionals is a very important one for our purpose. The 3-dimensional case is somewhat more simple than the 4-dimensional one, so we shall deal with it first. Let the minimal surface S be defined by the equation

$$f(x, y, z) = 0. \qquad (8)$$

The scalar distance between L_1 and L_2 is the functional

$$\Phi = \iint \frac{f_x\, dydz + f_y\, dzdx + f_z\, dxdy}{\mathbf{Z}}, \qquad (9)$$

where $f_x = \partial f/\partial x$, etc., and $\mathbf{Z}^2 = f_x^2 + f_y^2 + f_z^2$.

In virtue of the differential equation for minimal surfaces, namely, in the usual notation,

$$(1+q^2)\,r-2pqs+(1+p^2)\,t=0, \quad \ldots \ldots \ldots \quad (10)$$

the corresponding homogeneous equation to which (10) is equivalent, and which must be satisfied by f, is

$$\frac{\partial}{\partial x}\left(\frac{f_x}{Z}\right)+\frac{\partial}{\partial y}\left(\frac{f_y}{Z}\right)+\frac{\partial}{\partial z}\left(\frac{f_z}{Z}\right)=0. \quad \ldots \ldots \ldots \quad (11)$$

Consequently the necessary relation*

$$\frac{\partial}{\partial x}\left(\frac{d\Phi}{d\,(yz)}\right)+\frac{\partial}{\partial y}\left(\frac{d\Phi}{d\,(zx)}\right)+\frac{\partial}{\partial z}\left(\frac{d\Phi}{d\,(xy)}\right)=0 \quad \ldots \ldots \quad (12)$$

is fulfilled, the derivates of Φ, by their definition, being given by

$$\frac{d\Phi}{d\,(yz)}=\frac{f_x}{Z}, \qquad \frac{d\Phi}{d\,(zx)}=\frac{f_y}{Z}, \qquad \frac{d\Phi}{d\,(xy)}=\frac{f_z}{Z}. \quad \ldots \ldots \quad (13)$$

Now if $Z=1$, f satisfies LAPLACE's equation ; accordingly the function f and the functional Φ are then conjugate and Φ is a harmonic functional.† On account of (13) the derivates of the harmonic functional Φ will satisfy in addition to (12) the system of equations

$$\frac{\partial}{\partial y}\left(\frac{d\Phi}{d\,(yz)}\right)=\frac{\partial}{\partial x}\left(\frac{d\Phi}{d\,(zx)}\right), \text{ etc.} \quad \ldots \ldots \ldots \quad (14)$$

The corresponding relations in 4-dimensional space are derived as follows. The area of S connecting L_1 and L_2 is

$$\Phi=\iint \sum_{rs} N_{rs}\,dx_r\,dx_s, \quad \ldots \ldots \ldots \ldots \quad (15)$$

where

$$N_{rs}=(x_r, x_s)/[\sum_{rs}(x_r, x_s)^2]^{\frac{1}{2}},$$

any two surface coordinates u and v being used as independent variables in the calculation of the Jacobians. The condition that S be a minimal surface is that the symbolic differential under the double sign of integration in (15) must be a total symbolic differential. When the equations of the minimal surface have the form

$$f\,(x_1, x_2, x_3, x_4)=0 \qquad g\,(x_1, x_2, x_3, x_4)=0 \quad (16)$$

and hence

$$N_{12}=\frac{\partial\,(f,g)}{\partial\,(x_3, x_4)}\Big/\left\{\sum_{rs}\left[\frac{\partial\,(f,g)}{\partial\,(x_r, x_s)}\right]^2\right\}^{\frac{1}{2}}, \text{ etc.,} \quad \ldots \ldots \ldots \quad (17)$$

* VOLTERRA, "Theory of Functionals," p. 82.

† VOLTERRA, "Theory of Functionals," p. 88.

the conditions are the four equations

$$\frac{\partial N_{rs}}{\partial x_t} + \frac{\partial N_{st}}{\partial x_r} + \frac{\partial N_{tr}}{\partial x_s} = 0. \qquad (t \neq r, s). \quad \ldots \ldots \quad (18)$$

The equations (18) can also be deduced from the Eulerian equations of the variation problem by laborious transformation.

Now

$$N_{rs} = \frac{d\Phi}{d\,(x_r x_s)}, \quad \ldots \ldots \quad (19)$$

so that (18) gives the relations analogous to (12) which must be fulfilled by the functional derivates of Φ.

If we write

$$N^{rs} = \frac{\partial\,(f, g)}{\partial\,(x_r, x_s)} \Big/ \left\{ \sum_{rs} \left[\frac{\partial\,(f, g)}{\partial\,(x_r, x_s)} \right]^2 \right\}^{\frac{1}{2}}, \quad \ldots \ldots \quad (20)$$

then

$$\Psi = \iint \sum_{rs} N^{rs}\, dx_r\, dx_s \quad \ldots \ldots \quad (21)$$

is the functional conjugate to Φ, provided that

$$\sum_{rs} \left[\frac{\partial\,(f, g)}{\partial\,(x_r, x_s)} \right]^2 = 1, \quad \ldots \ldots \quad (22)$$

for the necessary relations

$$\frac{\partial}{\partial x_r}(N^{st}) + \frac{\partial}{\partial x_s}(N^{tr}) + \frac{\partial}{\partial x_t}(N^{rs}) = 0, \qquad (t \neq r, s), \quad \ldots \ldots \quad (23)$$

are satisfied identically. It will be noticed that Ψ is isogenic to each of the two functions f and g. This means that we can define a whole series of related functionals, of which

$$F = \int f\, d\Psi = \iint \sum_{rs} f N^{rs}\, dx_r\, dx_s \quad \ldots \ldots \quad (24)$$

is typical.

When the condition (22) is satisfied, equations (23) can be regarded as imposing on N_r the restrictions

$$\sum_s \frac{\partial N_{rs}}{\partial x_s} = 0, \quad \ldots \ldots \quad (25)$$

and it is obvious that equations (18) and (25) are just MAXWELL's equations in the absence of matter, whenever the N_{rs} are suitably correlated with the field components, under the condition that the electric and magnetic vectors are mutually perpendicular.* Equations (25) are analogous to (14). Ψ, of course, vanishes when the integration is over any portion whatever of the surface S defined by (16) and (18).

* I, p. 13.

Let us now return to the kinematical discussion which prompted the above consideration of geometrical matters, and consider first the non-relativistic case of a line in 3-space whose form and position depend on the time, the second independent variable being the parameter u by means of which the points of the line are distinguished. The locus of the moving line is a tube whose equations given in parametric form are:

$$x = x(u, t), \qquad y = y(u, t), \qquad z = z(u, t). \quad \ldots \ldots \quad (26)$$

A particular line on this tube is given by setting up a relation between u and t, which, of course, is represented by a line in the 2-dimensional space with coordinates u and t. Suppose that the lines L_1 and L_2 of xyz-space correspond respectively to the lines λ_1 and λ_2 of ut-space. The average velocity of the line in its motion from L_1 to L_2, on analogy with the velocity of a particle, is the ratio

Distance of L_1 from L_2 /Distance of λ_1 from λ_2.

The limiting value of this ratio as $\lambda_2 \rightarrow \lambda_1$ in the neighbourhood of (u, t) yields the analogue of instantaneous velocity, namely, the 3-vector whose components are

$$w_x = \frac{\partial (y, z)}{\partial (u, t)}, \qquad w_y = \frac{\partial (z, x)}{\partial (u, t)}, \qquad w_z = \frac{\partial (x, y)}{\partial (u, t)},$$

and whose magnitude $|w|$ is given by

$$|w| = |\mathbf{A}|\ |\mathbf{B}| \sin (\mathbf{AB}), \quad \ldots \ldots \ldots \ldots \quad (27)$$

where $\mathbf{A}$ is the vector $\left(\frac{\partial x}{\partial t}, \frac{\partial y}{\partial t}, \frac{\partial z}{\partial t}\right)$ and $\mathbf{B}$ the vector $\left(\frac{\partial x}{\partial u}, \frac{\partial y}{\partial u}, \frac{\partial z}{\partial u}\right)$. Since, under our assumptions, $\mathbf{A} \sin (\mathbf{A\,B})$ is the resolved part of $\mathbf{A}$ normal to the line at (x, y, z) and $|\mathbf{B}| = \frac{\partial s}{\partial u}$ where s denotes the length of arc of the line from some reference point to (x, y, z), it is not inappropriate to refer to the Jacobians as the components of the velocity of the line at the point (x, y, z).

In the 4-dimensional case, $\mathbf{V}$ is the 6-vector which is the vector product of two 4-vectors. We can make a reduction analogous to that given above for 3 dimensions by introducing two variables s and σ such that when x_4 is varied s does not change, there being no such restriction on σ, which, however, is chosen to have the dimensions of interval.

$$\mathbf{V}^2 = \Sigma\ (x_r, x_s)^2 = \Sigma \left[\frac{\partial (x_r, x_s)}{\partial (s, \sigma)} (s, \sigma)\right]^2 \qquad (r, s = 1, 2, 3, 4),$$

summed over all combinations of r and s. Thus

$$\mathbf{V}^2 = (s, \sigma)^2 \left\{\sum_{rs} \left[\frac{\partial (x_r, x_s)}{\partial (s, \sigma)}\right]^2 - c^2 \sum_r \left(\frac{\partial x_r}{\partial s}\right)^2 \left(\frac{dt}{d\sigma}\right)^2\right\} \qquad (r, s = 1, 2, 3)$$

$$= (s, \sigma)^2 \left\{\left[\sum_r \left(\frac{\partial x_r}{\partial s}\right)^2\right]\left[\sum_r \left(\frac{\partial x_r}{\partial \sigma}\right)^2 - c^2 \left(\frac{dt}{d\sigma}\right)^2\right] - \left[\sum_r \frac{\partial x_r}{\partial \sigma} \cdot \frac{\partial x_r}{\partial s}\right]^2\right\}. \quad \ldots \quad (28)$$

Now let

$$\left(\frac{\partial s_1}{\partial \sigma}\right)^2 = c^2\left(\frac{dt}{d\sigma}\right)^2 - \sum_r \left(\frac{\partial x_r}{\partial \sigma}\right)^2, \quad \ldots \ldots \quad (29)$$

then

$$(\mathbf{V})^2 = -(s, \sigma)^2 \left(\frac{\partial s_1}{\partial \sigma}\right)^2 \sin^2 (ss_1). \quad \ldots \ldots \quad (30)$$

On choosing v to be x_4 we obtain for the purpose of illustration

$$(\mathbf{V}) = \frac{\partial s}{\partial u} \cdot \frac{\partial \sigma}{\partial t} \cdot \frac{\partial s_1}{\partial \sigma} \sin (ss_1) = \frac{\partial s}{\partial u}\frac{\partial s_1}{\partial t} \sin (ss_1) \quad \ldots \ldots \quad (31)$$

which is of the same form as in the 3-dimensional case (27).

In a more direct fashion we can bring out the analogy in question by means of functional derivation. For example, $\mathbf{A}_{yz}$ the yz component of the 6-distance of the moving line from its initial configuration is given by (6) and is a function of the line λ in uv space. Its functional derivate with respect to the uv plane is

$$\frac{d\mathbf{A}_{yz}}{d\,(uv)} = \frac{\partial\,(y, z)}{\partial\,(u, v)}. \quad \ldots \ldots \quad (32)$$

Thus functional differentiation with respect to the pair of independent variables u and v in the motion of a line is analogous to ordinary differentiation with respect to the time in the motion of a particle. But the analogy is a restricted one, for so long as $\mathbf{A}_{yz}$ is a linear functional, the derivate (32) is a function of the point (x, y, z) only, and there is no place for the second operation of differentiation, whereas there is in passing from the velocity to the acceleration of a particle.

In order to obtain acceleration by means of the operator $d/d\,(uv)$, we have to create a functional from the components $\mathbf{V}_{rs} = (x_r, x_s)$ of the velocity at a point of the line. This can be done in the following way to which the method used in I* is equivalent. Consider a system of particle orbits of given energy in a conservative field of force. If ϕ is the potential energy per unit mass of particle, the rth component of acceleration

$$\dot{v}_r = -\frac{\partial \phi}{\partial x_r} = \sum_s \frac{\partial v_s}{\partial x_r} v_s,$$

since $\frac{1}{2} \sum_s v_s^2 + \phi =$ a constant. This suggests that we can set up a Hamiltonian system of motions of a line, and define the acceleration as the vector whose 4 components are given by

$$a_r = \sum_{is} \frac{\partial \mathrm{V}_{is}}{\partial x_r} \mathrm{V}_{is}.$$

If the form $\sum_{,} \mathrm{V}_{rs}\, dx_r\, dx_s$ is a total symbolic differential,

$$\frac{\partial \mathrm{V}_{ir}}{\partial x_s} + \frac{\partial \mathrm{V}_{rs}}{\partial x_i} - \frac{\partial \mathrm{V}_{is}}{\partial x_r} = 0,$$

* Pp. 17–18.

therefore,

$$a_r = \sum_{is} \left(\frac{\partial V_{rs}}{\partial x_i} + \frac{\partial V_{ir}}{\partial x_s}\right)(x_i, x_s)$$

$$= \sum_s \frac{\partial (V_{rs}, x_s)}{\partial (u, v)}. \quad (33)$$

That is

$$a_r = \frac{d}{d\,(uv)} \iint \sum_s dV_{rs}\, dx_s. \quad (34)$$

The functional appearing in (34) might be called the functional velocity of the line, to distinguish it from velocity at a point, which is properly speaking an intensity and not a property of the whole line. In his discussion of the exploration of the electromagnetic field by means of a secondary circuit, MAXWELL* introduced the corresponding distinction between the electromotive force in the whole circuit and the electromotive intensity at a point of the circuit.

We can also show the formal relation between the acceleration at a point of a moving line and the acceleration of a particle by another method. After carrying out the differentiations, we obtain,

$$a_r = \sum_s ((x_r, x_s), x_s)$$

$$= G\frac{\partial^2 x_r}{\partial u^2} - 2F\frac{\partial^2 x_r}{\partial u \partial v} + E\frac{\partial^2 x_r}{\partial v^2} + \left(\frac{\partial G}{\partial u} - \frac{\partial F}{\partial v}\right)\frac{\partial x_r}{\partial u} + \left(\frac{\partial E}{\partial v} - \frac{\partial F}{\partial u}\right)\frac{\partial x_r}{\partial v}, \quad (35)$$

where, in the usual notation of differential geometry,

$$E = \sum_r \left(\frac{\partial x_r}{\partial u}\right)^2, \qquad F = \sum_r \frac{\partial x_r}{\partial u} \cdot \frac{\partial x_r}{\partial v}, \qquad G = \sum_r \left(\frac{\partial x_r}{\partial v}\right)^2. \quad (36)$$

From (35), in the 3-dimensional case, if

$$H^2 = (y, z)^2 + (z, x)^2 + (x, y)^2 \quad (37)$$

and

$$(y, z) = HX, \text{ etc.}$$

$$a_x = \frac{\partial}{\partial u}\left(G\frac{\partial x}{\partial u} - F\frac{\partial x}{\partial v}\right) + \frac{\partial}{\partial v}\left(E\frac{\partial x}{\partial v} - F\frac{\partial x}{\partial u}\right)$$

$$= -\frac{\partial}{\partial u}\left[H\left(Y\frac{\partial z}{\partial v} - Z\frac{\partial y}{\partial v}\right)\right] + \frac{\partial}{\partial v}\left[H\left(Y\frac{\partial z}{\partial u} - Z\frac{\partial y}{\partial u}\right)\right]$$

$$= H[-(Y, z) + (Z, y)] + Y(z, H) - Z(y, H)$$

$$= K_m H^2 X + Y(z, H) - Z(y, H), \quad (38)$$

on making use of the well-known relations of differential geometry,† where K_m is the mean curvature of the surface. When H is a constant during the motion and

* "Treatise," vol. II, p. 242.

† *See*, for example, EISENHART, "Differential Geometry," p. 116, equation (13), and p. 123, equation (33).

when the acceleration also vanishes, then K_m must be equal to zero, that is, the line describes a minimal surface in its motion. Under the same condition of constant H, but with non-vanishing acceleration, (38) is quite analogous to the expression for the acceleration of a particle moving with constant speed in a circular path. Clearly one could find other points of analogy which are associated with particular choices of the parameters u and v. For instance, if we had $E = G = 1$ and $F = 0$

$$a_r = \frac{\partial^2 x_r}{\partial u^2} + \frac{\partial^2 x_r}{\partial v^2},$$

and the Laplacian operator replaces $(d/dt)^2$ when we pass from the particle to the line.

The most important result for us is the emergence of the motion of a line on a minimal surface as the analogue of the unaccelerated rectilinear motion of a particle.

2—Dynamics

When we pass now from kinematics to dynamics, it is necessary to introduce the quantity corresponding to momentum in the dynamics of a particle, and we define as the momentum of the line at one of its points

$$P_{rs} = \kappa\ (x_r, x_s), \quad \ldots \ldots \quad (39)$$

where κ has certain physical dimensions not yet specified.

It was indicated in I* that if " force " is defined as the vector whose rth component is

$$\Sigma_s\ (P_{rs}, x_s), \quad \ldots \ldots \quad (40)$$

and if this is derivable from the potential function $V(x_1, x_2, x_3, x_4)$, then the equations of motion (due to Volterra) are

$$(x_r, x_s) = \frac{\partial H}{\partial p_{rs}} \qquad \Sigma_s\ (p_{rs}, x_s) = -\frac{\partial H}{\partial x_r}, \quad \ldots \ldots \quad (41)$$

where

$$H\ (p_{23}, p_{31}, p_{12}, p_{14}, p_{24}, p_{34}, x_1, x_2, x_3, x_4) = \frac{1}{2\kappa} \Sigma_{rs}\, p_{rs}^2 + V$$

is the Hamiltonian function and p_{rs} is the canonical momentum conjugate to or associated with $x_r x_s$.

These equations of motion are connected with the variational problem

$$\delta \iint [H - V]^{\frac{1}{2}}\, dS = 0, \quad \ldots \ldots \quad (42)$$

when H is constant, dS being the element of area of the tube on which the line moves. In the case of free motion, that is, when the potential function V is constant, then,

* P. 18.

on account of (38) (re-written for 4 dimensions), the line describes a minimal surface, provided that κ, like the mass of the corresponding particle, is constant. Given the initial position of the line, and the initial velocity at each point of this line being specified, one can find the tube trajectory of the moving line by the use of SCHWARZ's method, for, when the velocity is specified, one can at once deduce the direction cosines of the required minimal surface at each point of the given initial line, and these data are sufficient to determine the minimal surface uniquely.* These considerations are of importance as they bear also on the analogue of FERMAT's Principle which must be the basis of the ray optics associated with the system of dynamics with which we are dealing.

It is desirable to investigate the physical significance of the idea of force introduced in equation (40). For this purpose it is convenient to consider the line analogue of the motion of an electrified particle of charge e in a given electromagnetic field. It will be recalled that it is necessary to transform the Hamiltonian function for the non-electrical case by substituting for each canonical momentum p_r, the sum of p_r, and the corresponding component of the potential-momentum 4-vector. This vector is the generalization of potential energy in accordance with special relativity and is equal to e times the electromagnetic 4-vector potential. Let us consider, therefore, in the case of the line motion, the Hamiltonian function

$$\mathrm{H} = \frac{1}{2\kappa} \sum_{rs} (p_{rs} - \lambda \xi_{rs})^2 \qquad \begin{pmatrix} r, s = 1, 2, 3, 4 \\ r \neq s \end{pmatrix}, \quad \ldots \ldots \quad (43)$$

where $\xi_{rs} = -\xi_{sr}$, ξ_{rs} are given functions of x_1, x_2, x_3, x_4, and λ is a given constant. The potential V has been put equal to zero, for in what follows it plays no essential part.

From the equations of motion (41), it is evident that

$$p_{rs} - \lambda \xi_{rs} = \mathrm{P}_{rs} = \kappa\,(x_r, x_s) \quad \ldots \ldots \ldots \quad (44)$$

$$\sum_s (\mathrm{P}_{rs}, x_s) = \sum_s (p_{rs}, x_s) - \lambda \sum_s (\xi_{rs}, x_s) = -\frac{\partial \mathrm{H}}{\partial x_r} - \lambda \sum_s (\xi_{rs}, x_s). \quad \ldots \quad (45)$$

Now

$$\sum_s (\xi_{rs}, x_s) = \sum_s \sum_{ik} \frac{\partial\,(\xi_{rs}, x_s)}{\partial\,(x_i, x_k)}\,(x_i, x_k) \qquad \text{summed over all combinations of } i, k$$

$$= \sum_{ik} \left(\frac{\partial \xi_{rk}}{\partial x_i} - \frac{\partial \xi_{ri}}{\partial x_k}\right)(x_i, x_k),$$

and

$$\frac{\partial \mathrm{H}}{\partial x_r} = -\lambda \sum_{is} \frac{\partial \mathrm{H}}{\partial p_{is}} \cdot \frac{\partial \xi_{is}}{\partial x_r} = -\lambda \sum \frac{\partial \xi_{is}}{\partial x_r}\,(x_i, x_s).$$

Hence

$$\sum_s (\mathrm{P}_{rs}, x_s) = \lambda \sum_{is} (x_i, x_s) \left(\frac{\partial \xi_{is}}{\partial x_r} + \frac{\partial \xi_{ri}}{\partial x_s} + \frac{\partial \xi_{sr}}{\partial x_i}\right). \quad \ldots \ldots \quad (46)$$

* DARBOUX, "Théorie Générale des Surfaces," Bk. III, chap. viii.

Now, provided that the ξ_{rs} are not the components of the curl of a vector, the coefficients of (x_i, x_s) in equation (46) form the vector whose 4 components are

$$\zeta^k = \frac{\partial \xi_{is}}{\partial x_r} + \frac{\partial \xi_{ri}}{\partial x_s} + \frac{\partial \xi_{sr}}{\partial x_i} = \sum_s \frac{\partial \xi^{ks}}{\partial x_s},$$

$k \neq i, r, s$, on raising the indices in the usual way, and writing V^{kr} for (x_i, x_s). Thus

$$\sum_s (P_{rs}, x_s) = \lambda \sum_k \zeta^k V^{kr}, \qquad (47)$$

which are quite analogous to the equations of motion of a particle in a given electromagnetic field ; in (47) the ζ^k fill the role played by the field components for the particle. Now suppose that the ξ^{ks} are the components of an electromagnetic field, the indices being assigned in the usual way, then ζ^k are the components of the charge and current density vector. On account of (47) and (44),

$$\sum_s (P_{rs}, x_s) = \frac{\lambda}{\kappa} \sum_k (p^{kr}\zeta^k - \lambda \xi^{kr}\zeta^k), \qquad (48)$$

and since

$$\sum_k \xi^{rk}\zeta^k = F_r$$

is the rth component of the mechanical force per unit volume of the field ξ^{ks}, we see that a mechanical interpretation can be given to the concept " force " introduced in (40). Let p_{rs} be assigned values throughout the 4-space of the x's by selecting one solution of the extended HAMILTON-JACOBI equation* and equating p_{rs} to the corresponding derivative $\dfrac{d\Phi}{d(x_r, x_s)}$, of the functional Φ which is the solution in question. On account of the necessary relations between the functional derivatives (*cf.* equations (18) and (19)), the p^{rs} are the components of a field having zero charge and current, so that

$$\sum_s (P_{rs}, x_s) = \frac{\lambda}{\kappa} \sum_k P^{kr}\zeta^k = \frac{\lambda^2}{\kappa} f_r, \qquad (49)$$

where f_r is the rth component of the force per unit volume of the electromagnetic field whose components are P^{kr}/λ associated with the particular solution Φ of the extended HAMILTON-JACOBI equation. Further, the Hamiltonian function similarly derived from Φ is λ^2/κ times the action density in this field.

(It is perhaps not without interest to note some of the possible sets of physical dimensions that may be assigned to λ and κ. Suppose that the product uv is of the dimensions [L T], then making use of equations (39) and (49) we have,

$$[\lambda^2] = [\kappa^2]\,[L]^3\,[M]^{-1} \qquad \text{or} \qquad [\lambda] = [\kappa]\,[L]^2\,[V]\,[Q]^{-1},$$

* *See* I, p. 22.

where [V] and [Q] denote velocity and electric charge respectively. Hence, if

(i) $[\kappa] = [M]\,[L]^{-1}$, $[\lambda] = [Q]\,[V]^{-1}$

(ii) $[\kappa] = [Q]\,[L]^{-1}$, $[\lambda] = [L]\,[V]$

(iii) $[\kappa] = [Q]\,[L]^{-2}$, $[\lambda] = [V]$

(iv) $[\kappa] = [Q]\,[L]^{-1}\,[V]^{-1}$ (electrokinetic momentum), $[\lambda] = [L]$.)

So far we have considered the motion of one line only. In a system of two lines it is necessary to set up some physical principle specifying how they act on each other. Probably the simplest assumption that we can make is that the sum of the forces defined by (40) must be zero, for all possible pairs of values of u and v. In an obvious notation this is

$$\sum_s (P_{rs}, x_s) + \sum_s (Q_{rs}, y_s) = 0. \qquad (50)$$

On performing double integration with respect to u and v between two lines λ_1, λ_2 of uv-space we obtain the analogue of the conservation of momentum, namely,

$$\iint \sum_s dP_{rs}\, dx_s + \iint \sum_s dQ_{rs}\, dy_s = 0. \qquad (51)$$

The application of this result is straightforward if the lines do not act on each other except at the discontinuity where a collision takes place. In the general case, however, the interaction between the lines would be given by means of a potential functional (instead of function) (*see* § 4), symmetrical in the two lines. A simple case of such a functional is the coefficient of mutual induction between two circuits. We shall not pursue the matter any further here.

In what follows reference is made to the transformation structure of this dynamical scheme, discussed in I. The required development of ideas is made in connexion with the analogue of optics, now to be considered.

3—Optics

Associated with the system of dynamics discussed in § 2 there is a geometrical or ray optics in which the rays are 2-dimensional manifolds. The connexion between the optical and dynamical systems is set up by Hamilton's Principle for lines (42), which becomes the optical law analogous to Fermat's Principle, namely*

$$\delta \iint \mu\, dS = 0. \qquad (52)$$

For a homogeneous medium, the rays are therefore minimal surfaces, this principle is analogous to that of the rectilinear propagation of light. From (52) the laws of reflexion and refraction can be deduced. In 3-dimensional space, for instance, let

* In this connexion the status of Fermat's Principle in optics should be kept in mind. *See* Cartan, " Invariants Intégraux," chap. xix.

a ray (surface) pass through the line L_1, be reflected at the plane P in the line L_0, and pass through the line L_2. If S_1 is the area of the minimal surface joining L_1 to L_0 and S_2 that joining L_2 and L_0 and if $d/d\sigma$ denotes functional differentiation with respect to the plane P, then (52) requires that

$$\frac{dS_1}{d\sigma}+\frac{dS_2}{d\sigma}=0, \quad \ldots \ldots \ldots \quad (53)$$

for each point of L_0. Evidently by taking the ordinary image L_1' of L_1 with respect to P, L_0 can be found as the intersection of the plane P with the minimal surface which joins L_1' and L_2.

The corresponding law of refraction is

$$\mu_2\frac{dS_2}{d\sigma}-\mu_1\frac{dS_1}{d\sigma}=0, \quad \ldots \ldots \ldots \quad (54)$$

where μ_1, S_1, μ_2, S_2 have their obvious meanings.

One can imagine a system of geometrical optics based on these principles and developed to an extent limited only by mathematical difficulties ; but although one can form a clear picture of the paths of the rays (in 3-space at least), one is not so happily placed with regard to the possible nature of the disturbance which is to be propagated over them, or of the physical processes by which the propagation might take place. The satisfactory answering of these two questions involves first the construction of a system of optics analogous to ordinary physical or wave optics, and then, once this is understood, the examination of experience to find instances to which the theory may be applied. At present we are concerned with the former task only, and, in the manner converse to that which aided HAMILTON to form the transformation theory of dynamics on the basis of his optical knowledge, we shall endeavour to arrive at the required optical ideas starting from the corresponding transformation theory of the dynamics of lines. Let us recall that in the natural extension of the HAMILTON-JACOBI theory * it is necessary to introduce a functional in place of the characteristic function of classical theory. Accordingly, the quantity specifying the corresponding optical disturbance is necessarily a functional : whether it is to be scalar or vector we are not yet in a position to decide. We shall confine ourselves to the question of phase, which is closely related to the dynamical transformation theory, and, in order to retain the possibility of geometrical aids to thought, we shall restrict the study to 3 dimensions for the present.

In ordinary optics, FERMAT'S Principle is equivalent physically to the agreement in phase of disturbances arriving at the same time at an image focus from all possible directions lying within a certain cone. With this for a guide, we see from the minimum principle (52) that the phase of the disturbance at L_2 differs from that at L_1 on the same ray by an amount equal to the area of the ray surface bounded by L_1 and L_2, provided that the " epoch " at which the disturbance at L_1 is considered, is

* Due to VOLTERRA and discussed in I, p. 16 and pp. 19–22.

the same as that at L_2. Since there are now two independent variables, " epoch " refers to a line in the uv-plane (a particular equation connecting u and v) and the disturbance is a function of a line L in xyz-space and a function of a line λ in uv-space, which may be written

$$\mathrm{F}\,[x\,(s), y\,(s), z\,(s)\;;\;u\,(w), v\,(w)].$$

What corresponds to time interval in ordinary optics is the " distance " between the two closed lines λ_1 and λ_2 in uv-space. Hence the phase difference for the lines L_1, L_2 on the same ray at the epochs λ_1 and λ_2 respectively is the functional

$$\Phi = a\iint d\mathrm{S} - b\iint du\,dv. \qquad (55)$$

The first integral of (55) is over the minimal surface joining L_1 and L_2, provided that the surface in question is a ray of the optical disturbance, and the second integral is the distance between λ_1 and λ_2. The constants a and b roughly correspond to wave-number and frequency respectively. At each point of the ray and each point of uv-space the " wave " will have a certain velocity specified in the same way that the velocity of a line at one of its points was specified in § 1. This arises as follows. Having in mind the propagation of ordinary visible waves, let us imagine some feature of the disturbance in the functional wave. Corresponding to a particular epoch λ_1, this feature occupies a certain set of lines (L_1) of xyz-space and for another epoch λ_2, it lies at the set (L_2). On a particular optical ray of the system there is one representative of (L_1) and one of (L_2). The propagation of the feature of the disturbance along a ray is therefore equivalent to the motion of a line along the ray from L_1 to L_2, and the velocity of propagation is to be calculated in the following way. For example, the (yz) component of velocity at $(x, y, z\;;\;u, v)$ is

$$\frac{b}{a\,\dfrac{d\mathrm{S}}{d\,(yz)}} = \frac{d\Phi}{d\,(uv)}\Big/\frac{d\Phi}{d\,(yz)}, \qquad (56)$$

for when $dydz$ and $dudv$ are in the above ratio, then Φ is unaffected by the simultaneous infinitesimal deformation of L and λ.

Continuing to deal with lines L on the same ray, suppose that the phase difference is measured from the line L_0. For a given epoch, the lines L which have the same phase satisfy the equation

$$\iint d\mathrm{S} = \text{constant}, \qquad (57)$$

integration being between L_0 and L on the minimal surface. Now since dS can be put in the form dy_1dy_2, it is readily seen that the solution of (57) can be reduced to the solution of the problem

$$\int_0^{2\pi} r^2\,d\theta = \mathrm{A}, \qquad (58)$$

where r is a function of θ to be determined and A is a constant. $r(\theta)$ must be periodic, and if $r_1(\theta)$ and $r_2(\theta)$ are solutions of (58), then $\frac{1}{\sqrt{2}}\{r_1(\theta)+r_2(\theta)\}$ is also a solution provided that $r_1(\theta)$ and $r_2(\theta)$ fulfil the condition

$$\int_0^{2\pi} r_1(\theta)\, r_2(\theta)\, d\theta = 0, \quad \ldots \ldots \quad (59)$$

i.e., $r_1(\theta)$, $r_2(\theta)$ are orthogonal functions. With suitable restrictions on the functions $r_s(\theta)$, one can regard (58) as the equation of a hyper-sphere in the functional space in which the r's are the coordinate vectors.

Turning from this very general problem, consider the lines satisfying (57) which pass through a particular point (x_0, y_0, z_0) of the ray. Since, as we have seen, the wave velocity at a point is a unique vector (56), and as in the neighbourhood of (x_0, y_0, z_0) the set (α) of lines L of equal phase on the ray determines a plane, namely, the tangent-plane to the ray at (x_0, y_0, z_0), it is natural to expect a connexion between the velocity vector and this plane. The vector is found to be normal to the plane. Each line L of the set α determines at (x_0, y_0, z_0) a plane element containing the normal to the ray at that point. There are ∞^1 such plane elements through (x_0, y_0, z_0) all intersecting in the normal. Let $f(x, y, z) = 0$ and $g(x, y, z) = 0$ be any two surfaces which at (x_0, y_0, z_0) each contain one of the plane elements. The direction cosines of their intersection at (x_0, y_0, z_0) are proportional respectively to

$$\frac{\partial(f, g)}{\partial(y, z)}, \quad \frac{\partial(f, g)}{\partial(z, x)}, \quad \frac{\partial(f, g)}{\partial(x, y)}.$$

Thus, although the tangent plane of a single surface such as f is not fixed by the wave velocity vector, yet the direction of the intersection of any pair of the set of surfaces to which f and g belong is determined. The set of surfaces f, g, etc., will correspond to a single wave surface in ordinary optics provided that they are loci of equal phase. Whereas in ordinary optics the normal to the wave surface at any point gives the direction of the ray (provided that the wave surface is a surface of equal amplitude and equal phase), in the system of functional optics which we are now considering, the normal to the ray surface at a particular point determines an element of the intersection of an infinite set of surfaces of equal phase. To prove this statement we require to specify how the phase difference for any two lines whatever (not merely on the same ray) is to be calculated. Let X, Y, Z be a solenoidal vector of unit length, whose value is given at each point (x, y, z). We have

$$X^2 + Y^2 + Z^2 = 1, \qquad \frac{\partial X}{\partial x} + \frac{\partial Y}{\partial y} + \frac{\partial Z}{\partial z} = 0, \quad \ldots \quad (60)$$

so that X, Y, Z can be considered as the direction cosines of the normal to a minimal surface. The specification of X, Y, Z throughout space is equivalent to giving a family of minimal surfaces, which we shall think of as a family of rays in the optical problem. In prescribing (X, Y, Z), which is, in fact, the wave velocity vector, throughout *xyz*-space and independent of the variables u, v, we have in mind the analogue of a steady progressive wave-motion.* The phase difference between L_1 and L_2 at the same epoch is the functional

$$\Phi = \iint X\, dy\, dz + Y\, dz\, dx + Z\, dx\, dy, \qquad (61)$$

integrated between L_1 and L_2. A locus of equal phase is therefore to be found by solving the extended Pfaff's problem†

$$d\Phi = X\, dy\, dz + Y\, dz\, dx + Z\, dx\, dy = 0. \qquad (62)$$

If $f(x, y, z) = 0$ is an integral manifold of (62) then f must satisfy the differential equation

$$X\frac{\partial f}{\partial x} + Y\frac{\partial f}{\partial y} + Z\frac{\partial f}{\partial z} = 0, \qquad (d\Phi df = 0), \qquad (63)$$

which, of course, shows that the surface f is everywhere orthogonal to the family of minimal surfaces which are the rays of the optical problem. The general solution of (63) involves an arbitrary function. If y_1, y_2 are characteristic variables of the symbolic form $d\Phi$, and therefore integrals of the characteristic system entering the solution of (63) by LAGRANGE's method,

$$d\Phi = dy_1\, dy_2,$$

and by the theory of symbolic differential forms, (62) requires $y_1 = \phi(y_2)$ where ϕ is an arbitrary functional symbol ; this is the general solution of (63). The infinity of solutions passing through a given point of space has the property that any two solutions determine a direction in space whose direction cosines are proportional to

$$\frac{\partial(y_1, y_2)}{\partial(y, z)}, \qquad \frac{\partial(y_1, y_2)}{\partial(z, x)}, \qquad \frac{\partial(y_1, y_2)}{\partial(x, y)}.$$

For, if we write

$$f(x, y, z) = y_1 - \phi(y_2) = 0, \qquad g(x, y, z) = y_1 - \psi(y_2) = 0,$$

we have

$$\frac{\partial(f, g)}{\partial(y, z)} = \left(\frac{d\phi}{dy_2} - \frac{d\psi}{dy_2}\right)\frac{\partial(y_1, y_2)}{\partial(y, z)}.$$

But

$$\frac{\partial(y_1, y_2)}{\partial(y, z)} = \frac{d\Phi}{d(yz)} = X, \qquad (64)$$

so the unique direction at each point (x, y, z) marked out by the family of surfaces there turns out to be (X, Y, Z) the direction normal to the ray surface at (x, y, z).

* *Cf.* SCHRÖDINGER, 'Ann. Physik,' vol. 79, §1 (1926).

† GOURSAT, "Le Problème de Pfaff," p. 111, *et seq.*

From the mathematical point of view, the propagation of the wave is the transformation of the point (x, y, z), together with the vector (X, Y, Z) into the point (x', y', z'), together with its corresponding vector (X′, Y′, Z′). These transformations are related to the extended PFAFF's problem in much the same way that contact transformations are related to PFAFF's problem,* and, for a reason which will soon appear, might be called "functional contact transformations". It will be recalled that in the transformation theory of ordinary optics, for a system of media through which passes a system of rays, HAMILTON's characteristic function determines the behaviour of the rays, and that contact transformations are the mathematical expression of HUYGENS's Principle. The time taken by light to travel through the medium from an arbitrary point (x, y, z) to another arbitrary point (x', y', z') depends only on the six quantities (x, y, z, x', y', z'); and if it is represented by $V(x, y, z, x', y', z')$, this is HAMILTON's characteristic function for the medium in question. In functional optics imagine that the disturbance has reached the arbitrary line L at the epoch λ and will reach another arbitrary line L′ at the epoch λ': we are going to require that the "distance" between the lines λ and λ' in the uv-plane is a function of the lines L and L′ only, and shall call this the characteristic functional for the system under consideration. Expressed formally,

$$W(L, L') = W[x(s), y(s), z(s), x'(t), y'(t), z'(t)] = \iint du\, dv, \quad . \quad (65)$$

integration being between λ and λ'. Suppose that the line L is varied on the locus of equal phase σ corresponding to the epoch λ. When λ' is given, (65) defines a set of loci, one member of the set for each L, and this set corresponds to a family of secondary waves from a wave-front in ordinary optics. The new locus of equal phase for the epoch λ', which we shall call σ', must be in some sense the envelope of the set of loci defined by (65) when L is varied on σ. We have, therefore, to find the geometrical meaning of contact of two loci which are defined by means of functional relations of the type,

$$G(L) = \text{constant, and } J(L) = \text{constant.}$$

The element of the "tangent plane" to G at the point (x, y, z) is given by

$$\frac{dG}{d(yz)}\, dy\, dz + \frac{dG}{d(zx)}\, dz\, dx + \frac{dG}{d(xy)}\, dx\, dy = 0. \quad . \quad . \quad . \quad . \quad . \quad (66)$$

We have seen that the solution of this equation (66) is a system of surfaces which at (x, y, z) have a common element of intersection whose direction cosines are proportional respectively to

$$\frac{dG}{d(yz)}, \quad \frac{dG}{d(zx)}, \quad \frac{dG}{d(xy)}.$$

* *Cf.* CAMPBELL, "Theory of Continuous Groups," chaps. xiv–xix. We cannot, however, refer to them as "extended" contact transformations as this term has already another meaning.

The loci G and J will be said to touch whenever the element of intersection at (x, y, z) belonging to G coincides with that belonging to J. This involves that the functional derivates of G and J with respect to the coordinate planes at (x, y, z) are proportional to each other. We now return to consider σ' as the envelope (in the sense consistent with the above conception of contact) of the loci defined by (65) when L lies on σ. This requires that, in the first place, if L is deformed in any way on σ, L′ being fixed, W (L, L′) will be unaltered. Hence if (dx, dy, dz) and $(\delta x, \delta y, \delta z)$ be two independent linear elements on σ at (x, y, z)*

$$\frac{d\mathrm{W}}{d\,(yz)}\,(dy\,\delta z - \delta y\,dz) + \frac{d\mathrm{W}}{d\,(zx)}\,(dz\,\delta x - \delta z\,dx) + \frac{d\mathrm{W}}{d\,(xy)}\,(dx\,\delta y - \delta x\,dy) = 0,$$

or in symbolic differential notation,

$$\frac{d\mathrm{W}}{d\,(yx)}\,dy\,dz + \frac{d\mathrm{W}}{d\,(zx)}\,dz\,dx + \frac{d\mathrm{W}}{d\,(xy)}\,dx\,dy = 0 \qquad (67)$$

But since any pair of linear elements of σ at (x, y, z) fulfil the condition

$$\mathrm{X}\,dy\,dz + \mathrm{Y}\,dz\,dx + \mathrm{Z}\,dx\,dy = 0 \qquad (68)$$

we must have

$$\frac{1}{\mathrm{X}}\frac{d\mathrm{W}}{d\,(yz)} = \frac{1}{\mathrm{Y}}\frac{d\mathrm{W}}{d\,(zx)} = \frac{1}{\mathrm{Z}}\frac{d\mathrm{W}}{d\,(xy)}. \qquad (69)$$

In the second place, since W (L, L′) is unaltered when L′ is deformed arbitrarily on σ', L being fixed, we have

$$\frac{d\mathrm{W}}{d\,(y'z')}\,dy'\,dz' + \frac{d\mathrm{W}}{d\,(z'x')}\,dz'\,dx' + \frac{d\mathrm{W}}{d\,(x'y')}\,dx'\,dy' = 0, \qquad (70)$$

for any pair of linear elements belonging to σ' at (x', y', z') and therefore satisfying

$$\mathrm{X}'\,dy'\,dz' + \mathrm{Y}'\,dz'\,dx' + \mathrm{Z}'\,dx'\,dy' = 0. \qquad (71)$$

On account of (70) and (71)

$$\frac{1}{\mathrm{X}'}\frac{d\mathrm{W}}{d\,(y'z')} = \frac{1}{\mathrm{Y}'}\frac{d\mathrm{W}}{d\,(z'x')} = \frac{1}{\mathrm{Z}'}\frac{d\mathrm{W}}{d\,(x'y')}. \qquad (72)$$

The equations (69) and (72), together with

$$\mathrm{X}^2 + \mathrm{Y}^2 + \mathrm{Z}^2 = 1 \text{ and } \mathrm{X}'^2 + \mathrm{Y}'^2 + \mathrm{Z}'^2 = 1, \qquad (73)$$

form a system of six relations from which one can determine the six quantities $x', y', z', \mathrm{X}', \mathrm{Y}', \mathrm{Z}'$ in terms of $x, y, z, \mathrm{X}, \mathrm{Y}, \mathrm{Z}$, so that once one is given the single functional W (L, L′), one can deduce the equations of transformation. Since X, Y, Z

* *Cf.* Volterra, "Theory of Functionals," pp. 76–77.

is the direction of the normal at (x, y, z), knowledge of the transformation equations enables one to construct the corresponding rays of the optical problem.

The equations (67)–(72) express the principle that in this system of functional optics is to act as the equivalent of HUYGENS's Principle. If we had employed a functional notation throughout, writing, for example, $dy\,dz = d\xi$ and for $\frac{dW}{d\,(yz)}, \frac{dW^*}{d\xi}$, then the equations would have exactly the same form as those in ordinary optical theory expressing that the wave-front at a particular time is the envelope of the secondary waves from each element of the front at a previous time. The infinitesimal transformations of the system would then be given by equations of the form for infinitesimal contact transformations but with functional elements (symbolic differentials) replacing ordinary differentials.

From the equations (69) and (72), we see that we can write

$$dW = \kappa\,(X\,dy\,dz + Y\,dz\,dx + Z\,dx\,dy) + \mu\,(X'\,dy'\,dz' + Y'\,dz'\,dx' + Z'\,dx'\,dy'), \quad (74)$$

where κ and μ are factors of proportionality not yet determined. If we write

$$\kappa X = -P, \qquad \kappa Y = -Q, \qquad \kappa Z = -R,$$
$$\mu X' = P', \qquad \mu Y' = Q', \qquad \mu Z' = R',$$

we have

$$dW = P'\,dy'\,dz' + Q'\,dz'\,dx' + R'\,dx'\,dy' - P\,dy\,dz - Q\,dz\,dx - R\,dx\,dy. \quad (75)$$

Since W is a function of two lines only, if (x', y', z', P', Q', R') are expressed in terms of (x, y, z, P, Q, R) in conformity with an optical transformation, then the right-hand member of (75) must be a total symbolic differential. The infinitesimal transformations of this system were discussed in I† and found to conform to the Hamiltonian system (41). They are given by

$$\left.\begin{aligned} x' &= x + a\,\Delta u + a'\,\Delta v, & P' &= P + p\,\Delta u + p'\,\Delta v \\ y' &= y + b\,\Delta u + b'\,\Delta v, & Q' &= Q + q\,\Delta u + q'\,\Delta v \\ z' &= z + c\,\Delta u + c'\,\Delta v, & R' &= R + r\,\Delta u + r'\,\Delta v \end{aligned}\right\}, \quad . \; . \quad (76)$$

where $a, b, c, a', b', c', p, q, r, p', q', r'$ are functions of x, y, z, P, Q, R, and satisfy the following equation system,

$$bc' - cb' = \frac{\partial K}{\partial P}, \qquad (rb' - r'b) - (qc' - q'c) = -\frac{\partial K}{\partial x}, \text{ etc.,}$$

K being a function of x, y, z, P, Q, R.

Now consider the line L on the locus σ corresponding to the epoch λ. If λ be deformed infinitesimally at the point (u, v) in such a way that the area of the uv-plane

* *See* VOLTERRA, "Theory of Functionals," p. 93.
† P. 20.

bounded by λ is increased by the fixed quantity $[\Delta u \Delta v]$, the corresponding transformation of L in accordance with (76) and, therefore with the propagation of the functional wave in the interval $[\Delta u \Delta v]$, will have the property now to be deduced.

Let $\Delta_1 u$, $\Delta_1 v$ and $\Delta_2 u$, $\Delta_2 v$ be a pair of linear elements of the uv-plane such that

$$\Delta_1 u \Delta_2 v - \Delta_2 u \Delta_1 v = [\Delta u \Delta v] = k. \quad \quad (77)$$

Each element will determine a transformation in accordance with (76), and if $\Delta_1 y$ denotes

$$y'_1 - y = b \Delta_1 u + b' \Delta_1 v,$$

and $\Delta_2 y$ denotes

$$y'_2 - y = b \Delta_2 u + b' \Delta_2 v, \text{ etc.},$$

then we have

$$[\Delta y, \Delta z] = \Delta_1 y \Delta_2 z - \Delta_2 y \Delta_1 z = (bc' - cb')\, k,$$

and similarly

$$[\Delta z, \Delta x] = (ca' - ac')\, k$$
$$[\Delta x, \Delta y] = (ab' - ba')\, k. \quad \quad (78)$$

These three equations determine an element of ray surface at (x, y, z) corresponding to the given function K, or, if we view the matter from the dynamical side, the coefficients of k in (78) are the components v_{yz}, v_{zx}, v_{xy} of the velocity at (x, y, z) of a line whose dynamical trajectory is the ray in question. Now the system of possible rays emanating from a point is geometrically equivalent to the system of possible dynamical trajectories through that point and belonging to a particular value of the Hamiltonian function K. Since for an isotropic and homogeneous medium (*cf.* (42) and (52)) we have

$$v_{yz}^2 + v_{zx}^2 + v_{xy}^2 = A^2 \text{ (a constant)},$$

the equation for the secondary wave from (x, y, z) corresponding to the elapse of the element k on the uv-plane is

$$\Sigma\, [\Delta y, \Delta z]^2 = A^2 k^2. \quad \quad (79)$$

This is the functional analogue of the MONGE equation which defines the propagation of light waves in ordinary optics.*

Up to the present we have dealt only with the phase of the functional waves and have not undertaken the full specification of the magnitude whose variation as a function of a line in xyz-space and of a line in uv-space constitutes the wave motion which we have in mind. We shall call this magnitude the wave displacement in analogy with the elastic solid theory of light, although from the point of view of general wave theory it ought to be called the wave functional. We now require some means of distinguishing briefly between a line regarded as a geometrical locus only and the dynamical entity whose motion was discussed in § 2, and shall do so by referring to the latter always as a p-line. A possible physical process which can be imagined as taking place in the propagation of the wave is that each p-line of the

* *Cf.* CARTAN, " Invariants Intégraux," p. 196.

medium filling *xyz*-space is caused to move by the wave and is under the dynamical action of neighbouring *p*-lines. In order that the process be a wave at all, the displacement Ψ must be a periodic function of the phase W. This involves the idea of the dependence of one functional on another *by a non-linear relation* and raises a mathematical difficulty which has now to be investigated.* If we reduce the matter to its simplest terms, we are concerned with the dependence of the functional Ω on the linear functional $A = \iint du\, dv$ which is a function of the line λ bounding the area by an equation of the type

$$\Omega = f(A). \qquad (80)$$

If λ is deformed infinitesimally in the neighbourhood of the point (u_1, v_1) on the line λ, the limiting value of the ratio

$$\frac{\text{increment of } \Omega}{\text{increment of the area A}}, \qquad (81)$$

is a function of the point (u_1, v_1) and of the line λ. That is, the limit in question has no longer the property which is essential to the notation $\frac{d\Omega}{d\,(uv)}$, for the functional differentiation of a linear functional. If, however, we confine ourselves to considering the class of those curves λ passing through (u_1, v_1) which are solutions of the functional equation (*cf.* (57) above),

$$\iint du\, dv = A_0, \qquad (82)$$

where A_0 is a constant, then the ratio (81) is independent of λ. We can therefore write the limit of the ratio as $\left[\frac{d\Omega}{d\,(uv)}\right]_{A_0}$ and, of course, its value is $f'(A_0)$.

This result is readily extended to the case

$$\Omega = f(B),$$

where

$$B = \iint \phi\,(u, v)\, du\, dv.$$

The derivative of Ω with respect to the *uv*-plane at the point (u_1, v_1) and $B = B_0$ is

$$\left[\frac{d\Omega}{d\,(uv)}\right]_{B_0} = f'(B_0)\, \phi\,(u_1, v_1). \qquad (83)$$

Once the restriction on λ in the calculation of the functional derivative is understood, there is no need to use a special notation, it will therefore be given up. The derivative is a function of the line λ and of the point (u_1, v_1) and, with the same understanding

* This difficulty was recognized by Born, 'Proc. Roy. Soc.,' A, vol. 143, pp. 410–423 (1934), and alluded to in I, p. 25.

as to the calculation of the derivative, it is possible to repeat the operation of differentiation, thus obtaining in general

$$\frac{d^r\Omega}{d\,(uv)^r} = f^{(r)}\,(\mathrm{B}_0) \prod_{s=1}^{r} \phi\,(u_s, v_s), \quad \dots\dots\dots\dots \quad (84)$$

where (u_s, v_s) is the point of the uv-plane at which the sth derivative is taken.

It is out of place to attempt to treat the subject of functional differentiation at greater length here. The functional Ω may be a function of any number (n) of linear functionals which are all functions of the line λ; passing to the limit when n is increased indefinitely, one may have Ω depend on a linear functional by a functional relation; also there is the question of partial differentiation. In all such cases, whatever special problems arise in the definitions on account of ambiguity of some kind or another, the general mathematical procedure must be to introduce the restrictions required to remove ambiguity and to enable one to retain the simple notation which resembles as closely as possible that of the differential calculus. There is, however, a case of functional differentiation which calls for particular attention. Suppose that Q is the linear functional $\iint dy\,dz$ integrated over the region of the yz-plane bounded by the line L, and that the uv-plane is mapped on the yz-plane by means of the equations

$$y = y\,(u, v), \qquad z = z\,(u, v) \qquad \text{with} \qquad (y, z) = \phi\,(u, v). \quad \dots\dots (85)$$

Hence L corresponds to a particular line λ of the uv-plane. In virtue of (85), Q, regarded as a function of the line λ, is the functional

$$\mathrm{Q} = \iint \phi\,(u, v)\,du dv, \quad \dots\dots\dots\dots \quad (86)$$

and its derivative with respect to the uv-plane is an ordinary function, so that Q cannot have functional derivatives of higher order. In what way must the equations (85) and (86) be altered in order to achieve the possibility of higher order derivatives? It seems necessary to introduce some degree of indefiniteness into the mapping of uv on yz. Let α be a parameter which is a function of the line λ, and let

$$y = y\,(u, v, \alpha) \qquad z = z\,(u, v, \alpha). \quad \dots\dots\dots\dots \quad (87)$$

If α is kept constant

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

Write

$$\mathrm{Q} = \iint \phi\,(u, v, \alpha)\,du\,dv,$$

then

$$\frac{d\mathrm{Q}}{d\,(uv)} = \frac{d\,(yz)}{d\,(uv)} = \phi\,(u, v, \alpha),$$

and since α is a function of the line λ, it is possible to repeat the operation of functional differentiation. The geometrical significance of (87) is that we are now dealing with a family of mappings of the uv-plane on the yz-plane. A particular mapping corresponds to a definite value of α, but is independent of the line λ provided that λ belongs to the class which is defined by this particular value of α. Usually we shall have $\alpha = \iint dudv$ integrated over the region of the uv-plane bounded by the line λ.

The possibility of the above extension of the scope of functional differentiation throws new light on the dynamics of § 2, where we considered only the particular case of linear relationships between the functional variables (*cf.* equation (86)). We shall now try to adopt in dynamics the complete functional notation adumbrated in our study of functional contact transformations. Before reconsidering dynamics briefly from this point of view in § 4, we ought to deal with possible methods of treating some of the many physical questions that obviously suggest themselves from ordinary optics, *e.g.*, the formation of shadows which, of course, involves the diffraction and interference of waves. These matters, however, will be treated in a subsequent paper.

4—Reconsideration of Dynamics

The conception of a system of dynamics in which the dynamical variables are functionals is by no means new. Although he did not use any of the terms of the modern functional calculus, Maxwell,* in his dynamical theory of electric circuits, used as variables quantities which are functions of the whole circuit, and in what follows the reader ought to have Maxwell's writings in mind.

Let u and v be taken as independent variables, and let α denote the area of the uv-plane bounded by the line λ. Following the procedure adopted in equation (87), let us propose

$$x = x\,(u, v, \alpha) \qquad y = y\,(u, v, \alpha) \qquad z = z\,(u, v, \alpha). \quad . \quad . \quad . \quad . \quad (88)$$

This means that for the class of lines λ that give α the particular value α_1, (x, y, z) will lie on a certain surface S_1 (the tube trajectory of §1) and for another value α_2 of α, (x, y, z) will in general lie on another surface S_2. We are going to require, however, that the surfaces S_1, S_2, . . . are the same geometrical entity S and that when we assign different particular values to α, we give different laws for the mapping of the uv-plane on S. This is equivalent to

$$\frac{\partial\,(x, y, z)}{\partial\,(u, v, \alpha)} = 0, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (89)$$

or to stating that

$$x = \mathrm{X}\,(\mathrm{U}, \mathrm{V})\ , \quad y = \mathrm{Y}\,(\mathrm{U}, \mathrm{V})\ , \quad z = \mathrm{Z}\,(\mathrm{U}, \mathrm{V}).$$

where X, Y, Z are symbols for ordinary functions and U, V are functions of u, v, α. If we were to choose U and V as independent variables, we should return to the system of §§ 1 and 2 (apart from the particular equations of motion adopted there)

* "Treatise," vol. II, chaps. v–viii.

but we shall assume that U and V are not physically accessible and that, therefore, we are not free to make this transformation. Let

$$(y, z)_\alpha = \phi_1(u, v, \alpha), \qquad (z, x)_\alpha = \phi_2(u, v, \alpha), \qquad (x, y)_\alpha = \phi_3(u, v, \alpha), \quad . \quad (90)$$

then

$$\left.\begin{aligned} q_1 &= \iint dy\, dz = \iint \phi_1(u, v, \alpha)\, du\, dv \\ q_2 &= \iint dz\, dx = \iint \phi_2(u, v, \alpha)\, du\, dv \\ q_3 &= \iint dx\, dy = \iint \phi_3(u, v, \alpha)\, du\, dv \end{aligned}\right\}, \quad . \quad . \quad . \quad . \quad . \quad . \quad (91)$$

are the functional displacements of the p-line which we imagine to move on the surface S from the position L_0 to the position L, these lines being the boundaries of the region of integration in the first column of integrals in (91). The corresponding integrals on the uv-plane are over the region bounded by the lines λ_0 and λ which fix the epochs at which the p-plane has the positions L_0 and L respectively in xyz-space.

The components of velocity are

$$\frac{dq_1}{d(uv)} = \frac{d(yz)}{d(uv)}, \qquad \frac{dq_2}{d(uv)} = \frac{d(zx)}{d(uv)}, \qquad \frac{dq_3}{d(uv)} = \frac{d(xy)}{d(uv)}, \quad . \quad . \quad . \quad (92)$$

which reduce to the Jacobians of §§ 1 and 2 when x, y, z do not depend on α. At first, as in (39), the components p_1, p_2, p_3 of momentum are assumed to be κ times the corresponding components of velocity but later their definition will not be so restricted (*cf.* ordinary dynamics). We have to propose a suitable form for the equations of motion, and are guided in this by aiming to conserve as much as possible of the transformation theory previously discussed. We therefore retain the equations

$$p_1 = \frac{dW}{d(yz)}, \qquad p_2 = \frac{dW}{d(zx)}, \qquad p_3 = \frac{dW}{d(xy)}, \text{ and } H = -\frac{dW}{d(uv)}, \quad . \quad (93)$$

where H, the Hamiltonian, is a function of $q_1, q_2, q_3, p_1, p_2, p_3$, while W is the functional which in the transformation theory plays the part of the characteristic functional in optics. Accordingly, it is not unnatural to write the equations of motion in the form

$$\frac{dp_1}{d(uv)} = -\frac{dH}{d(yz)}, \qquad \frac{d(yz)}{d(uv)} = \frac{dH}{dp_1}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (94)$$

Corresponding to (41),

$$H = \frac{1}{2\kappa}(p_1^2 + p_2^2 + p_3^2) + V, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (95)$$

the potential V, being a function of q_1, q_2, q_3, is now a functional (function of the line L) and the forces derived from it are given by functional differentiation. When we consider (94) in detail, we find the system has great generality. Suppose, for example, that V = 0, then

$$\frac{dp_1}{d\,(uv)} = 0\ , \qquad \frac{dp_2}{d\,(uv)} = 0\ , \qquad \frac{dp_3}{d\,(uv)} = 0.$$

Consequently,

$$p_1 = \kappa f_1\,(u, v), \qquad p_2 = \kappa f_2\,(u, v), \qquad p_3 = \kappa f_3\,(u, v),$$

where f_1, f_2, f_3, are arbitrary functions. It follows that

$$q_1 = \iint f_1\,(u, v)\,du\,dv, \text{ etc.} \qquad (96)$$

On account of the arbitrary nature of f_1, f_2, f_3, the equations of motion in the coordinates used are not adequate to determine the surface on which the p-line moves. The case V = 0 in this system includes all the possibilities of § 2. On the other hand, we may regard f_1, f_2, f_3 as analogous to the constants of integration which appear in ordinary differential equations, so that these functions would be no longer arbitrary in a given case but would be data of the problem. The equations (96) then determine (y, z), (z, x), and (x, y) as functions of u and v.

We shall mention only two other forms of V. In the first place when V is quadratic in the q's, these coordinates are periodic in α, and the motion corresponds to free vibration. (In optics we have to deal with the forced vibration and this would require the forces to be periodic in α.) In the second place, if V is the line-integral round the line L

$$\mathrm{V} = \int \mathrm{A}_x\,dx + \mathrm{A}_y\,dy + \mathrm{A}_z\,dz\,, \qquad \text{with div } \mathbf{A} = 0,$$

then

$$-\frac{dp_1}{d\,(uv)} = \frac{\partial \mathrm{A}_z}{\partial y} - \frac{\partial\,\mathrm{A}_y}{\partial z}. \text{ etc.,} \qquad (97)$$

the form of which brings to mind Maxwell's equations. Suppose that the vector **A** represents the magnetic field, and that therefore the derivatives of p_1, p_2, p_3, with respect to the uv-plane are equal respectively to the three components of the electric current density $\times\ 4\pi$, then (97) expresses the well-known electromagnetic law, but what is more striking is that the constancy of the Hamiltonian H expresses the law of conservation of current. This is shown as follows. Let (i, j, k) represent the total current density, then

$$i\,\frac{dq_1}{d\,(uv)} + j\,\frac{dq_2}{d\,(uv)} + k\,\frac{dq_3}{d\,(uv)} = \frac{1}{4\pi\kappa}\left(p_1\,\frac{dp_1}{d\,(uv)} + p_2\,\frac{dp_2}{d\,(uv)} + p_3\,\frac{dp_3}{d\,(uv)}\right).$$

Hence, integrating

$$\left|\frac{1}{2\kappa}\,(p_1{}^2 + p_2{}^2 + p_3{}^2\right|_{\mathrm{L}_0}^{\mathrm{L}} = 4\pi \iint i\,dy\,dz + j dz\,dx + k\,dx\,dy$$

$$= -\left|\int \mathrm{A}_x\,dx + \mathrm{A}_y\,dy + \mathrm{A}_z\,dz\right|_{\mathrm{L}_0}^{\mathrm{L}}. \qquad (98)$$

The limits of integration of the double integral over the surface S are the lines L_0 and L. This possibility of connecting the functional forces (97) with electric current should be compared with the system (43)–(49), but it should hardly be necessary to add that the system (97) has been considered only to illustrate the possibilities of the method.

The above formulation of functional dynamics is put forward tentatively, for it still remains to be seen that all the difficulties of notation have been removed. There is at least one attractive possibility. In I (p. 21) the generalization of the POISSON brackets appropriate to the system of § 2 was introduced, but on account of the fact that these new brackets involve three variables, it is difficult to understand how they could possibly be used as the basis of a quantum algebra (of electromagnetism connected with the atomicity of electric charge). On the other hand, with the system here proposed, using functional variables, this problem would be removed and the uncertainty relations obtained in I (p. 26) could be expressed by means of brackets involving only two variables, and therefore suitable for the definition of the algebraic operation of multiplication.

With a view to physical understanding of the independent functional variable, it is desirable to remove the vagueness inherent in the above formalism by referring once again to the ideas which prompted the whole investigation.

5—ELECTRICAL INTERPRETATIONS

For simplicity of discussion, consider the motion of a line in space of 3 dimensions only and imagine that one is given a certain form H of the line Hamiltonian function, say (41). The corresponding extended HAMILTON-JACOBI equation is

$$\left[\frac{dW}{d\,(yz)}\right]^2 + \left[\frac{dW}{d\,(zx)}\right]^2 + \left[\frac{dW}{d\,(xy)}\right]^2 + V + \frac{dW}{d\,(uv)} = 0, \quad . \ . \ . (99)$$

a solution of which has the form

$$W = \iint X\,dy\,dz + Y\,dz\,dx + Z\,dx\,dy, \quad . \ . \ . \ . \ . \ . \ . \quad (100)$$

where we must have div (X, Y, Z) = 0 if W is to be single-valued. (We have left out the part of W depending on *dudv*.) The specification of X, Y, Z at each point of space corresponds to an electrostatic field throughout space. Another solution yields another possible electrostatic field related to the same function H. The fundamental hypothesis is advanced in I that the transformation from the former field to the latter corresponds to a possible physical process, and that the structure of the system of possible transformations (and therefore of fields) should be governed by the magnitude e, the electronic charge, in accordance with certain rules analogous to those of quantum mechanics in which h (PLANCK's constant) appears.

An electrostatic field corresponding to the Hamiltonian function

$$H = p_{yz}^2 + p_{zx}^2 + p_{xy}^2$$

is one in which the electropotential surfaces of WHITTAKER are minimal surfaces. Since, on identifying E_x with p_{yz}, etc., we get

$$E_x\,(y, z) + E_y\,(z, x) + E_z\,(x, y) = \text{constant},$$

or

$$\frac{d}{d\,(uv)}\iint E_x\,dy\,dz + E_y\,dz\,dx + E_z\,dx\,dy = \text{constant}, \quad . \quad . \quad . \quad (101)$$

the property secured by our equations of motion (41) for the moving line on an equipotential surface is that the surface charge embraced between two curves on the surface is in constant ratio to the area of the uv-plane between the two corresponding epochs. It may be said that the line motion in the present instance dramatizes the measurement of the flux of electric induction over an equipotential surface. An electrical transformation in the sense of the last paragraph corresponds to the introduction or creation of charge somewhere in the system.

In attempting to form a physical idea of the process which has been referred to above as a dramatization of the measurement of charge, one is brought face to face with the question "what is the physical significance of u and v?" and this we shall now try to answer.

In the motion of a closed circuit, the trajectory in 3-space is the surface made up by successive positions of the line where the circuit is at successive instants. If we wish to have a picture which conforms to the special theory of relativity, we construct in 4-dimensional space the 2-dimensional manifold S made up of the successive spatial lines lying in the series of hyperplanes which corresponds to the elapse of time; the intrinsic geometrical form of S is unaffected by rotation of the axes of coordinates. Viewed from another set of axes (x', y', z', t'), the surface S can be regarded as the motion of a circuit whose successive appearances are sections of S by hyperplanes perpendicular to the t' axis. This second series of curves on S is on the same level of physical importance as the first and indeed as any other series obtained by change of the time axis through a Lorentz transformation. But whereas to the first observer each member of his own system of sections of S corresponds to a *value of the variable* t, to the same observer any section of S appropriate to the time axis of another observer will correspond in general to an *equation between the variables* x, y, z, t. For example, in the case of the Lorentz transformation

$$x' = \beta\,(x - vt), \qquad y' = y, \qquad z' = z, \qquad t' = \beta\left(t - \frac{vx}{c^2}\right),$$

let the curve C be the intersection of S with $t' = a$, where a is a constant. The curve C is not a spatial curve for the first observer (system $xyzt$), nevertheless, it would have perfect sense for this observer to refer to C as the "position of the moving line *at the time* $a = \beta\,(t - vx/c^2)$". Instead of giving a value to the variable representing the time in his own system of coordinates, he would have given the equation connecting two variables equivalent to assigning a value to the time of some other frame of reference.

Now let us consider the tube S as the locus of a point whose coordinates x, y, z, t are functions of two parameters u, v. The curve C is that line on S for which, say, $f(u, v) = 0$. Such a relation between u and v corresponds to a value of the time in some system of coordinates (provided, of course, that C is a time-like section). What we have previously called the epoch, namely, a particular relation between u and v, such as $f(u, v) = 0$, has in fact, to correspond to the time in some system of coordinates.

Let C_1 and C_2 be two curves on S corresponding to the epochs λ_1 and λ_2 specified by $f_1(u, v) = 0$ and $f_2(u, v) = 0$ respectively, which in turn correspond to the times t_1 and t_2 in two different systems of coordinates. If the parameters u and v are normalized so that $\iint du\, dv$ integrated over the uv-plane between the curves λ_1 and λ_2 is equal to the area of the tube S between the curves C_1 and C_2 in general, $\iint du\, dv$ is invariant with respect to Lorentz transformations and is a measure of the separation of λ_1 and λ_2 (*cf.* § 1), which corresponds to proper time in the motion of a particle. The time at a particular instant in any system of coordinates, at least of the Lorentz group, will be represented by a curve on the uv-plane. Thus, the order of times belonging to different systems of reference is a functional one, and what we are doing when we regard time in this way is to look on the substitution of a particular value for the time in one system of reference not as the mere naming of a particular instant but as one of the class of equations which are equivalent, on suitable Lorentz transformation, to naming a particular instant in some possible system of reference. It is evident that if $\iint du\, dv$ is invariant with respect to Lorentz transformation it cannot be used to measure time intervals in any one frame, for the latter are not Lorentz invariant.

Now it is the essence of the measurement of time that we compare the process to be dated with some standard process usually periodic which is taken as a clock, and in the case of the moving line it is likewise necessary to select one type of motion as standard and to refer other motions to this one. Electrically this corresponds to the measurement of electric charge by comparing fields with a standard field. Suppose, then, that we wish to describe the motion of line B using that of A as standard. We have to set up a one-to-one correspondence between the positions of B and the positions of A. One physical way of making this correspondence is to make the position of A and B belong to the same 3-space, that is, to the same time for one observer O. If A_1 is the position of A for time t_1 of the observer O, and B_1 is the curve occupied by B in the space of the observer O for the same time t_1, then B_1 corresponds to A_1. By making O variable we can complete the correspondence. We choose u and v as a pair of independent variables to describe the motion of B such that $\iint du\, dv$ gives the proper area of the tube S_A described by the motion of A between the epochs in question.

We have now to consider how two observers who do not employ the same motion A as standard are to compare observations, and, indeed, to investigate the general question of transformation of the functional coordinates. Let us first study the analogue of Newtonian Relativity, and for simplicity confine our attention to motion of a line on a plane. An observer O uses coordinates y, z, to specify position on the plane, and the pair of independent variables u, v, when dealing with the motion of a line on the plane. A particular case of the motion of a line on the yz-plane is mathematically equivalent to giving a law for mapping the uv-plane on the yz-plane, $\frac{\partial\,(y,\,z)}{\partial\,(u,\,v)}$ being specified at each point $(u,\,v)$. Now suppose the observer O′ uses the coordinates $y'z'$ and the same independent variables $(u\;v)$ and that at a particular epoch λ_1 the line L′ (arbitrary) of $y'z'$ coincides with the line L of yz. When λ is deformed in the neighbourhood of $(u_1,\,v_1)$ L′ differs from L in the neighbourhood of $(y'_1,\,z'_1)$ and $(y_1,\,z_1)$ which are coincident. That is, the line L′ is moving relatively to the yz plane. The relative velocity in the sense of § 1 of this paper is $\frac{\partial\,(y',\,z')}{\partial\,(u,\,v)} = w$, which must be constant over the $y'z'$ plane in order that it should refer to a relationship between the two systems of reference $(y'z')$ and (yz). (Otherwise it will be necessary to introduce some kind of tensor to express the relation of general relativity.) In general, then, when the observers O and O′ use the same independent variables u, v the velocities of any line motion will be connected by the equation

$$\frac{\partial\,(y',\,z')}{\partial\,(u,\,v)} = \frac{\partial\,(y,\,z)}{\partial\,(u,\,v)} + w. \qquad \qquad (102)$$

The analogue of Einsteinian (special) relativity must be based on the hypothesis that when the $y'z'$-plane is moving in the above sense with respect to the yz-plane, O and O′ do not have the same standard of reference for epoch. We now have to assign to O the new pair of independent variables $u'v'$, and to discover equations of transformation from (yz) and (uv) to $(y'z')$ and $(u'v')$ involving w. From equation (102) in the analogue of Newtonian Relativity

$$\iint dy'\,dz' = \iint dy\,dz + w \iint du\,dv. \qquad \qquad (103)$$

If we proceed in strict analogy with the special theory of relativity it is necessary to postulate some velocity which is invariant with respect to transformation of the functional coordinates corresponding to " uniform translation " of the $y'z'$-plane with respect to the yz-plane. In § 3 of this paper we have discussed a possible process of just the type required, and it seems not unnatural to assume that under the conditions we are considering the invariant velocity has the constant value a. Writing

$$\gamma^2 = \left(1 - \frac{w^2}{a^2}\right)^{-1}, \qquad \qquad (104)$$

we have

$$\iint dy'dz' = \gamma \iint dydz + w\gamma \iint dudv$$

$$\iint du'dv' = \gamma \iint dudv + \frac{w\gamma}{a^2} \iint dy\, dz, \qquad (105)$$

while the expression

$$a^2\,(du\,dv)^2 - (dy\,dz)^2 \qquad (106)$$

is invariant with respect to transformations (105).

In the case of a line traversing a tube S in the 4-dimensional space the invariant which corresponds to (106) will be

$$a^2\,(du\,dv)^2 - [(dy\,dz)^2 + (dz\,dx)^2 + (dx\,dy)^2 - c^2\,(dx\,dt)^2 - c^2\,(dy\,dt)^2 - c^2\,(dz\,dt)^2] = a^2\,(du\,dv)^2 - dS^2\,. \qquad (107)$$

When we identify the electromagnetic field components (at least in the case $\mathbf{E}\,.\,\mathbf{H} = 0$) with the momentum at the same point in the motion of a line on an electropotential surface in the given field, by means of the equations

$$E_x = \kappa\,(y, z), \qquad H_x = \kappa c\,(x, t), \qquad \text{etc.}, \qquad (108)$$

the pair of independent variables must be associated with a line motion taken as a standard of reference which has some electrical significance. In this way the representation of the field components by means of the Jacobians is a method which illustrates that the measurement of an electromagnetic field is essentially the comparison of the field to be measured with a given standard field and this is physically correct. Combining (107) and (108) and extracting the square root, we obtain

$$a\,du\,dv\,\sqrt{1 + \frac{1}{\kappa^2 a^2}\,(\mathbf{H}^2 - \mathbf{E}^2)}, \qquad (109)$$

invariant with respect to transformations of the type (105). Of course, it is necessary to give (105) an electrical interpretation, and this follows at once because

$$\frac{\partial\,(y', z')}{\partial\,(u', v')} = \left(\frac{\partial\,(y, z)}{\partial\,(u, v)} + w\right) \Big/ \left(1 + \frac{w}{a^2}\frac{\partial\,(y, z)}{\partial\,(u, v)}\right), \qquad (110)$$

and therefore if x be the same in both systems

$$E_x' = \Big(E_x + (E_x)_o\Big) \Big/ \left(1 + \frac{E_x\,(E_x)_o}{b^2}\right), \qquad (111)$$

where b is the limiting field κa, $(E_x)_o$ is a uniform field in respect to which the O′ frame differs from that of O. Thus, as might have been expected because Born in his first paper avowedly used the special theory of relativity as his guide, we have arrived at the conception of a limiting electric field together with the principle that electric fields are not linearly superposable. Further, in (109), we get as the multiplier of *du dv* the well-known Lagrangian density adopted by BORN in his first paper to which the form of his later theory reduces when the electric and magnetic fields are everywhere mutually perpendicular. The method of analysis followed in the present paper is significant in a special way, for it shows how BORN's theory is naturally connected with the possibility of the process analogous to light first postulated in I, and discussed in § 3 here under the name " functional waves ".*

Neither of the electromagnetic 6-vectors can be represented immediately by means of the direction cosines of a 2-dimensional manifold in 4-dimensional space unless the electric and magnetic vectors are mutually orthogonal. By combining the two electromagnetic six-vectors linearly according to the proper γ-transformation,† one can, however, construct two new 6-vectors in which the condition of orthogonality holds, and it is then possible to connect the electric and magnetic vectors of the original field with the direction cosines of the surfaces belonging to the transformed field. The equations of transformation and the ensuing calculation can be most conveniently expressed by combining the electromagnetic field vectors into a bivector of GIBBS as has been done by BATEMAN, SILBERSTEIN, SCHRÖDINGER, and others.

Let

$$\mathrm{P}_x = \mathrm{E}_x + i\mathrm{H}_x \text{ etc.}, \qquad \mathrm{G} = \mathbf{E} \cdot \mathbf{H} \qquad \mathrm{F} = \tfrac{1}{2}(\mathbf{E}^2 - \mathbf{H}^2). \quad . \quad . \quad (112)$$

The transformed bivector is given by

$$\mathrm{E}'_x + i\mathrm{H}'_x = \mathrm{P}'_x = \mathrm{P}_x\, e^{i\phi} \qquad \text{where} \qquad \tan 2\phi = \mathrm{G}/\mathrm{F}. \quad . \quad . \quad . \quad (113)$$

Now let us introduce the electropotential surfaces of the field $\mathbf{E}'\mathbf{H}'$; the direction cosines are proportional to the field components and we can write

$$\frac{\partial\,(y, z)}{\partial\,(u, v)} = \frac{\mathrm{E}'_x}{\sqrt{\mathbf{E}'^2 - \mathbf{H}'^2}} \cdot \frac{d\mathrm{S}}{d\,(uv)}, \qquad c\,\frac{\partial\,(x, t)}{\partial\,(u, v)} = \frac{\mathrm{H}'_x}{\sqrt{\mathbf{E}'^2 - \mathbf{H}'^2}}\,\frac{d\mathrm{S}}{d\,(uv)}, \quad .(114)$$

where

$$\frac{d\mathrm{S}}{d\,(uv)} = \sqrt{\{(y, z)^2 + (z, x)^2 + (x, y)^2 - c^2\,(x, t)^2 - c^2\,(y, t)^2 - c^2\,(z, t)^2\}}, \quad (115)$$

dS being an element of area of the surface in question.

* It should be noted that functional waves enter the theory in two quite distinct ways ; first as the analogue of light in its role as the conventional means of connecting events in different frames of reference, and as the process whose velocity is invariant for linear transformations of the space and time coordinates ; secondly, as the analogue of de Broglie waves which are intimately connected with the transformation theory of dynamics and are not propagated with invariant velocity. In the former role of functional waves we are concerned with the correlation of different electrical frames of reference, in the latter with the transformation theory of electromagnetic fields and ultimately the existence of atoms of electric charge.

† This term was introduced by SCHRÖDINGER, ' Proc. Roy. Soc.,' A, vol. 150, p. 465 (1935). *See also* WATSON, ' Phys. Rev.,' vol. 48, p. 776 (1935).

On account of (114) we have

$$(y, z) + ic\,(x, t) = \frac{E'_x + iH'_x}{\sqrt{\mathbf{E}'^2 - \mathbf{H}'^2}} \cdot \frac{dS}{d\,(uv)} = \frac{P_x}{\sqrt{P_x^2 + P_y^2 + P_z^2}} \cdot \frac{dS}{d\,(uv)}. \quad (116)$$

If $\dfrac{dS}{d\,(uv)} = m$ is constant during the motion of the line on the electropotential surface, then we can treat the bivector

$$[(y, z) + ic\,(x, t), \qquad (z, x) + ic\,(y, t), \qquad (x, y) + ic\,(z, t)]$$

as the velocity vector and **P** as the corresponding momentum bivector with the Hamiltonian function

$$\mathscr{H} = m\sqrt{P_x^2 + P_y^2 + P_z^2} = m\,(J + iK) \quad . \; . \; . \; . \; . \; . \; . \quad (117)$$

where

$$J = \sqrt{2}\,(F^2 + G^2)^{\frac{1}{4}} \cos\phi = \sqrt{(F^2 + G^2)^{\frac{1}{2}} + F}$$

$$K = \sqrt{2}\,(F^2 + G^2)^{\frac{1}{4}} \sin\phi = \sqrt{(F^2 + G^2)^{\frac{1}{2}} - F}. \quad . \; . \; . \; . \quad (118)$$

The equations (116) are equivalent to the two systems

$$\text{(I)} \quad (y, z) = m\frac{\partial J}{\partial E_x}, \qquad c\,(x, t) = -\,m\frac{\partial J}{\partial H_x}, \quad . \; . \; . \; . \quad (119)$$

and

$$\text{(II)} \quad (y, z) = m\frac{\partial K}{\partial H_x}, \qquad c\,(x, t) = m\frac{\partial K}{\partial E_x}. \quad . \; . \; . \; . \; . \quad (120)$$

In this way the two distinct representations of the field used in I (p. 25) are combined into a single system and it appears possible that a single complex functional could be employed in the transformation theory of that paper.

On Lorentz transformation with velocity v parallel to x, **P** and $\mathscr{H}$ are transformed according to the equations

$$P'_x = P_x\,, \quad P'_y = \beta\left(P_y + \frac{iv}{c}P_z\right), \qquad P'_z = \beta\left(P_z - \frac{iv}{c}P_y\right), \; \mathscr{H}' = \mathscr{H}, \quad (121)$$

consequently we cannot treat the set (P_x, P_y, P_z, $\mathscr{H}$) as a complex four-vector. They may, however, be combined in a biquaternion, thus

$$p = J + iK + P_x\boldsymbol{a} + P_y\boldsymbol{b} + P_z\boldsymbol{c} = J + iK + \mathbf{P} \quad . \; . \; . \; . \; . \quad (122)$$

where

$$\boldsymbol{a}^2 = \boldsymbol{b}^2 = \boldsymbol{c}^2 = -1\,, \; \boldsymbol{ab} = \boldsymbol{c}, \boldsymbol{ca} = \boldsymbol{b}, \boldsymbol{bc} = \boldsymbol{a}. \quad . \; . \; . \; . \; . \; . \quad (123)$$

This particular biquaternion is a "nullifier," for $Tp = 0$. The use of quaternionic forms of representation in the special theory of relativity is not new. Johnston* and Larmor† pointed out the appropriate character of a linear associative algebra

* 'Proc. Roy. Soc.,' A, vol. 96, p. 331 (1919).
† *Ibid.*, vol. 96, p. 334 (1919).

based on CLIFFORD's ideas for the representations of electromagnetic relations, while SILBERSTEIN, in his "Theory of Relativity", has treated special relativity throughout in quaternion notation, which on account of its algebra must eventually be preferred to the Minkowski calculus. Before the advent of BORN's theory, however, it can hardly be said that biquaternions were necessary, for all the essential representation was made by bivectors, and it was certainly not recognized that the scalar part of the biquaternion would have to be connected with a natural unit for the measurement of electromagnetic fields. Before treating the general electromagnetic field, let us introduce the biquaternion whose bivector part is

$$\mathbf{V} = -i\,[(y, z) + ic\,(x, t), \qquad (z, x) + ic\,(y, t)\,, \qquad (x, y) + ic\,(z, t)],$$

where the Jacobians represent the component velocities of a moving line. It seems natural to combine with this bivector, the scalar a of (104), the invariant velocity of the functional waves analogous to light. We have, therefore, the biquaternion

$$a + \mathbf{V}.$$

If we multiply this by κ and use (108) we create in the case $\mathbf{E}\,.\,\mathbf{H} = 0$, the biquaternion, $q_0 = b + \mathbf{H} - i\mathbf{E}$ whose norm is $\mathrm{T}q_0 = \sqrt{b^2 - 2\,\mathrm{F}}$.

For the general field where $\mathbf{E}\,.\,\mathbf{H} \neq 0$, let

$$q = x + iy + \mathbf{H} - i\mathbf{E}.$$

Suppose that we require that $\mathrm{T}q$ shall be real. Then since

$$\mathrm{T}q^2 = q\,\mathrm{K}q = x^2 - y^2 + 2ixy - 2\mathrm{F} - 2i\mathrm{G},$$

we must have

$$xy = \mathrm{G}.$$

Hence

$$\mathrm{T}q = \sqrt{x^2 - 2\mathrm{F} - \mathrm{G}^2/x^2},$$

or, putting $x = b$ of BORN's theory,

$$b\,\mathrm{T}q = \sqrt{b^4 - 2b^2\mathrm{F} - \mathrm{G}^2}\ \left(= \sqrt{(b^2 - \mathrm{J}^2)\,(b^2 + \mathrm{K}^2)}\right),$$

which is BORN's Lagrangian ; q is then

$$q = b + i\mathrm{G}/b + \mathbf{H} - i\mathbf{E}.$$

If we operate on q with the biscalar $(b^2 - i\mathrm{G})/\mathrm{T}q$ so as to make a new biquaternion whose scalar is real we obtain

$$q_1 = \frac{b^2 - i\mathrm{G}}{\mathrm{T}q} \cdot q = \frac{b^4 + \mathrm{G}^2}{b\mathrm{T}q} + \mathbf{B} - i\mathbf{D}$$

where $\mathbf{B}$ and $\mathbf{D}$ make the other pair of field vectors required in BORN's theory.*

* The meanings of **B** and **H** must be interchanged for comparison with BORN's equations.

The relations which have just been pointed out serve in themselves as sufficient evidence of the appropriateness of biquaternions relative to BORN's theory. Other aspects of the use of this notation, which is particularly suggestive, have been discussed in detail elsewhere.*

Let us return to what may be called the relativity form of the Hamiltonian function for the motion of a line corresponding to the general electromagnetic field. We are at liberty to choose the value of $\frac{dS}{d\,(uv)}$ in (116). Let us assume

$$\frac{dS}{d\,(uv)} = \frac{J}{M} \quad \text{or} \quad \frac{K}{N},$$

where M is a function of J only, and N of K only, and M and N are possible Hamiltonians. We have, therefore,

$$(y, z) = \frac{\partial M}{\partial E_x}, \text{ etc.,} \qquad \text{or} \qquad (y, z) = \frac{\partial N}{\partial H_x}, \text{ etc.}$$

The result of the calculation is $M^2 = \pm J^2 + \text{const.}$, which suggests that what we must choose are

$$M = \sqrt{b^2 - J^2} = \sqrt{b^2 - F - \sqrt{F^2 + G^2}},$$

and

$$N = \sqrt{b^2 + K^2} = \sqrt{b^2 - F + \sqrt{F^2 + G^2}}.$$

In the case of mutually perpendicular electric and magnetic fields these give

$$M_1 = \sqrt{b^2 + \mathbf{H}^2 - \mathbf{E}^2}, \qquad N_1 = b.$$

This value M_1 of M agrees with our former value of the Hamiltonian expected by analogy with the special theory of relativity, whereas the value N_1 of N requires that the velocity of the moving line with this Hamiltonian is zero (all Jacobians of the type (y, z) vanish). The representation of the field by means of its magnetopotential surfaces is achieved only if the correspondence of **E** and **H** with the components of line velocity in M_1 is inverted.

CONCLUSION

The ideas outlined in the last section appear adequate to connect BORN's theory of a limiting electric field strength and of a non-linear law of composition of fields with the conception of functional waves as the physical process which in its dual role is analogous on the one hand to light and on the other hand to the de Broglie waves of so-called material particles. It is evident that once one accepts a non-linear law of composition of fields one has to give up the MAXWELL-LORENTZ method of describing a field in vacuum by means of only two vectors. The consequences of

* "Trans. Roy. Soc. Can.," vol. 30, Sect. III, pp. 105–113 (1936).

this have not been worked out by the present writer, but it seems proper to point out that BORN's electromagnetic equations are of the form (12) of § 1, so that it is still possible to regard any given electromagnetic field as derived by functional differentiation from a solution of an extended HAMILTON-JACOBI equation (actually a pair of such equations would be necessary if twelve independent quantities, the components of four vectors, were necessary to describe the field). Thus, one can combine a transformation theory of electromagnetic fields with a theory of the same type as BORN's. One must retain the idea that the field vectors can be represented by functional derivatives as in I, but one is allowed a fairly wide choice of form for the Hamiltonian function to determine the actual functional derivative equation which plays the part analogous to that of the wave equation in quantum mechanics. The restriction imposed by BORN's theory on the form of Hamiltonian function is, roughly speaking, analogous to the restriction required by special relativity in the dynamics of a particle.

SUMMARY

This is an extended and more detailed study of the system, due to VOLTERRA, discussed by the writer in a previous paper as the mathematical basis of a transformation theory of electromagnetism. The following topics are dealt with: the functional distance between two closed lines in space: motion of a line regarded as the motion of a point depending on two independent variables; definition of velocity and acceleration by functional differentiation; minimal surface as analogue of the straight line: possible physical meaning of force as defined by VOLTERRA's equations: system of two lines and analogue of the conservation of momentum: geometrical optics with 2-dimensional rays corresponding to the dynamical trajectories: definition of phase, wave velocity vector: connexion between loci of equal phase and the rays in 3-dimensions: transformation theory: question of the complete specification of the wave functional—non-linear relations between linear functionals: attempt to construct dynamics with the possibility of such non-linear relations; functional time; relativity of functional coordinate systems; representation of general electromagnetic field in functional dynamics; BORN's hypothesis of limiting field intensity and the functional analogue of light; non-linear superposition of fields; use of biquaternions.