XVI.

APPLICATION OF QUATERNIONS TO LORENTZ TRANSFORMATIONS.

By P. A. M. DIRAC.[1]

[Received 30 July. Published 27 November, 1945.]

Linear Transformations of Quaternions.

The fact that quaternions have four components and the physical world has four dimensions has led people to think that there ought to be a close correspondence between them. However, a quaternion

$$q = q_0 + iq_1 + jq_2 + kq_3$$

has a modulus whose square is

$$q_0^2 + q_1^2 + q_2^2 + q_3^2 \tag{1}$$

and a vector A_μ in space-time has a length whose square is

$$A_0^2 - A_1^2 - A_2^2 - A_3^2, \tag{2}$$

and the difference in the signs of (1) and (2) prevents one from identifying a quaternion directly with a vector in space-time. People have tried to get over the difficulty by introducing another square root of -1, independent of i, j and k, and with its help setting up a connexion between quaternions and vectors in space-time such that expression (1) becomes equal to (2). In this way the physical world is put into correspondence with the scheme of bi-quaternions, instead of with the scheme of quaternions. Now the scheme of bi-quaternions is not of any special interest in mathematical theory. Quaternions themselves occupy a unique place in mathematics in that they are the most general quantities that satisfy the division axiom—that the product of two factors cannot vanish without either factor vanishing. Bi-quaternions do not satisfy this axiom, and do not have any fundamental property which distinguishes them from other hyper-complex numbers. Also, they have eight components, which is rather too many for a simple scheme for describing quantities in space-time.

It would seem to be a more fruitful line of investigation to keep to the ordinary quaternions and to try to connect them with the physical

[1] Lecture given at the Dublin Institute for Advanced Studies on July 12th, 1945.

world by subjecting them to a group of transformations equivalent to the Lorentz group. Since quaternions satisfy the division axiom, it is natural to consider transforming them by a linear transformation with a denominator, i.e. by a transformation which corresponds to the general linear transformation

$$z^* = \frac{az + b}{cz + d} \tag{3}$$

for complex numbers. We shall find that the transformations so obtained are too general, but that a simple way of restricting them will suggest itself, which will give us just what we want.

We cannot immediately take over the transformation (3) to quaternions, because we do not know the order in which the multiplication processes and the division process on the right-hand side of (3) are to be carried out. To settle this point, we must first express the quaternion as a ratio of two quaternions and then transform the two quaternions by a homogeneous linear transformation.

Express the quaternion q as the ratio of u to v, i.e.

$$q = uv^{-1}. \tag{4}$$

With λ any quaternion, we now have

$$q = u\lambda\lambda^{-1} v^{-1} = u\lambda\, (v\lambda)^{-1},$$

showing that q can be equally well expressed as the ratio of $u\lambda$ to $v\lambda$. However, q is not equal to the ratio of λu to λv. Thus, when we express a quaternion as a ratio of two quaternions, these two quaternions can be multiplied by an arbitrary factor only on one side,—only on the right if we use the form (4).

To transform q, we transform u and v by the linear homogeneous equations

$$\left.\begin{aligned} u^* &= au + bv \\ v^* &= cu + dv \end{aligned}\right\} \tag{5}$$

with a, b, c, d arbitrary quaternions, and take

$$q^* = u^* v^{*-1}.$$

It is necessary to put the coefficients a, b, c, d in (5) all on the left, since we can replace u, v by $u\lambda$, $v\lambda$ and we do not want q^* to be altered thereby. We now have q^* determined by q alone,

$$\begin{aligned} q^* &= (au + bv)\,(cu + dv)^{-1} = (au + bv)\, v^{-1}\, v\, (cu + dv)^{-1} \\ &= (au + bv)\, v^{-1} \left[(cu + dv)\, v^{-1}\right]^{-1} = (aq + b)\,(cq + d)^{-1}. \end{aligned} \tag{6}$$

This is the generalization of the transformation (3) to quaternions.

Alternatively, we could express q as a ratio $\alpha^{-1}\beta$, and carrying through the corresponding work we should get the transformation of q in the form

$$q^* = (qa + b)^{-1}(qc + d). \tag{7}$$

Equation (6) may be written

$$\begin{aligned} q^* &= a(q + a^{-1}b)\,[c(q + c^{-1}d)]^{-1} \\ &= a(q + a^{-1}b)(q + c^{-1}d)^{-1}c^{-1}. \end{aligned}$$

It is now clear that we must have

$$a^{-1}b \neq c^{-1}d, \tag{8}$$

otherwise the transformation (6) would be singular.

Two transformations of the form (6), (8) applied in succession give another transformation of this form. To get the result of the transformation (6) itself followed by the transformation

$$\tilde{q} = (a'q^* + b')(c'q^* + d')^{-1}, \tag{9}$$

we must take the transformation (5) and follow it by

$$\begin{aligned} \tilde{u} &= a'u^* + b'v^* \\ \tilde{v} &= c'u^* + d'v^*. \end{aligned}$$

The result of these two transformations is

$$\begin{aligned} \tilde{u} &= (a'a + b'c)u + (a'b + b'd)v \\ \tilde{v} &= (c'a + d'c)u + (c'b + d'd)v. \end{aligned}$$

Hence

$$\begin{aligned} \tilde{q} = \tilde{u}\tilde{v}^{-1} &= [(a'a + b'c)u + (a'b + b'd)v][(c'a + d'c)u + (c'b + d'd)v]^{-1} \\ &= [(a'a + b'c)q + a'b + b'd][(c'a + d'c)q + c'b + d'd]^{-1}. \end{aligned} \tag{10}$$

The transformation (6), (8) has a reciprocal transformation of the same form, as may be seen in the following way. We shall try to choose the coefficients a', b', c', d' in (9) so that $\tilde{q}$ is equal to the original q. From (10), we see that the conditions for this are

$$\left.\begin{aligned} a'b + b'd = 0 \qquad c'a + d'c = 0 \\ a'a + b'c = c'b + d'd = m, \end{aligned}\right\} \tag{11}$$

where m is a real number, not zero. These conditions are satisfied by

$$\left.\begin{aligned} a' &= \frac{m}{b^{-1}a - d^{-1}c}\,b^{-1} = \frac{m}{a - bd^{-1}c} \\ b' &= \frac{-m}{b^{-1}a - d^{-1}c}\,d^{-1} = \frac{m}{c - d\,b^{-1}a} \\ c' &= \frac{m}{a^{-1}b - c^{-1}d}\,a^{-1} = \frac{m}{b - ac^{-1}d} \\ d' &= \frac{-m}{a^{-1}b - c^{-1}d}\,c^{-1} = \frac{m}{d - ca^{-1}b}. \end{aligned}\right\} \tag{12}$$

The forms for a', b', c', d' given first are the convenient ones to use for checking that the conditions (11) are satisfied, but these forms must be replaced by those given second in the event of b, d, a or c vanishing. The denominators in the forms given second cannot vanish on account of (8), so that a', b', c', d' can always be chosen so as to satisfy the required conditions. Thus, the reciprocal transformation always exists and the transformations (6), (8) form a group.

Restriction to the Lorentz Group.

With the quaternion q expressed as the ratio (4), we introduce the three quantities

$$Q_1 = u\bar{v}, \quad Q_2 = u\bar{u}, \quad Q_3 = v\bar{v}, \tag{13}$$

where the bar denotes the conjugate complex of a quaternion. The three Q's have the property that if we replace u, v by $u\lambda$, $v\lambda$, they get multiplied by the real number $\lambda\bar{\lambda}$, since, for example, Q_1 gets changed to

$$u\lambda\,\overline{(v\lambda)} = u\lambda\bar{\lambda}\bar{v} = Q_1\lambda\bar{\lambda}.$$

Put
$$Q_1 = X_0 + iX_1 + jX_2 + kX_3, \tag{14}$$

where X_0, X_1, X_2, X_3 are real, and put

$$Q_2 = X_4 - X_5, \quad Q_3 = X_4 + X_5. \tag{15}$$

We now have the three Q's determining six real numbers,

$$X_0,\ X_1,\ X_2,\ X_3,\ X_4,\ X_5.$$

If u, v are replaced by $u\lambda$, $v\lambda$, the six X's all get multiplied by $\lambda\bar{\lambda}$ and their ratios are unchanged. Thus the ratios of the X's are determined by q. From (13)

$$Q_1\bar{Q}_1 = u\bar{v}v\bar{u} = uQ_3\bar{u} = Q_2Q_3$$

Hence, from (14) and (15),

$$X_0^2 + X_1^2 + X_2^2 + X_3^2 = X_4^2 - X_5^2. \tag{16}$$

If q is transformed according to (6), corresponding to u, v being transformed according to (5), the new Q's are linear functions of the old ones, for example,

$$Q_1^* = u^* \bar{v}^* = (au + bv)(\bar{u}\bar{c} + \bar{v}\bar{d})$$
$$= aQ_2\bar{c} + bQ_3\bar{d} + aQ_1\bar{d} + b\bar{Q}_1\bar{c}.$$

This will result in the new X's being linear functions of the old ones, with real coefficients determined by a, b, c, d. Thus the transformation (6) of q generates a linear homogeneous transformation of the six variables X. Equation (16) must, of course, hold also for the new X's.

We may describe our results geometrically by considering the six X's as the homogeneous coordinates of a point in five-dimensional projective space. Equation (16) is the equation of a quadric in the five-dimensional space. We then have the result that any transformation (6) of q generates in the five-dimensional space a projective transformation which leaves the quadric (16) invariant. The transformations do not leave anything else invariant in the five-dimensional space, since they contain fifteen independent real parameters (namely, the ratios of the sixteen real numbers which determine the four quaternions a, b, c, d), and the general projective transformation in five dimensions which leaves a quadric invariant also contains fifteen independent parameters.

It is now clear that the group of transformations (6) is more general than the Lorentz group, and that we must impose certain restrictions on the coefficients a, b, c, d in order to get the Lorentz group. These restrictions may conveniently be taken to be those corresponding to the planes

$$X_0 = 0, \quad X_5 = 0, \tag{17}$$

remaining invariant as well as the quadric (16). Equations (17) give, with the help of (13), (14) and (15),

$$u\bar{v} + v\bar{u} = 0, \quad u\bar{u} - v\bar{v} = 0. \tag{18}$$

The left-hand side of the first of equations (18) transforms under the transformation (5) into

$$u^*\bar{v}^* + v^*\bar{u}^* = (au + bv)(\bar{u}\bar{c} + \bar{v}\bar{d}) + (cu + dv)(\bar{u}\bar{a} + \bar{v}\bar{b})$$
$$= u\bar{u}(a\bar{c} + c\bar{a}) + v\bar{v}(b\bar{d} + d\bar{b}) + au\bar{v}\bar{d} + dv\bar{u}\bar{a} + cu\bar{v}\bar{b} + bv\bar{u}\bar{c}.$$

Now, using $R(\alpha)$ to denote the real part of the quaternion α and remembering that $R(\alpha\beta) = R(\beta\alpha)$, we have

$$au\bar{v}\bar{d} + dv\bar{u}\bar{a} = 2R(au\bar{v}\bar{d}) = 2R(u\bar{v}\bar{d}a) = u\bar{v}\bar{d}a + \bar{a}dv\bar{u}.$$

Similarly

$$cu\bar{v}\bar{b} + bv\bar{u}\bar{c} = u\bar{v}\bar{b}c + \bar{c}bv\bar{u}.$$

Hence

$$u^*\bar{v}^* + v^*\bar{u}^* = u\bar{u}(a\bar{c} + c\bar{a}) + v\bar{v}(b\bar{d} + d\bar{b}) + u\bar{v}(\bar{d}a + \bar{b}c) + (\bar{a}d + \bar{c}b)v\bar{u}. \quad (19)$$

In order that the first of equations (18) shall remain invariant, we must have

$$u^*\bar{v}^* + v^*\bar{u}^* = r\,(u\bar{v} + v\bar{u}) \quad (20)$$

where r is a real number, not zero. Comparing the right-hand sides of (19) and (20), we get

$$a\bar{c} + c\bar{a} = 0, \quad b\bar{d} + d\bar{b} = 0, \quad (21)$$

$$\bar{d}a + \bar{b}c = \bar{a}d + \bar{c}b = r. \quad (22)$$

Similarly, the left-hand side of the second of equations (18) transforms under the transformation (5) into

$$\begin{aligned} u^*\bar{u}^* - v^*\bar{v}^* &= (au + bv)(\bar{u}\bar{a} + \bar{v}\bar{b}) - (cu + dv)(\bar{u}\bar{c} + \bar{v}\bar{d}) \\ &= u\bar{u}(a\bar{a} - c\bar{c}) + v\bar{v}(b\bar{b} - d\bar{d}) + au\bar{v}\bar{b} + bv\bar{u}\bar{a} - cu\bar{v}\bar{d} - dv\bar{u}\bar{c} \\ &= u\bar{u}(a\bar{a} - c\bar{c}) + v\bar{v}(b\bar{b} - d\bar{d}) + u\bar{v}(\bar{b}a - \bar{d}c) + (\bar{a}b - \bar{c}d)v\bar{u}. \end{aligned}$$

Thus, in order that the second of equations (18) shall remain invariant, we must have

$$\bar{b}a - \bar{d}c = 0, \quad \bar{a}b - \bar{c}d = 0, \quad (23)$$

$$a\bar{a} - c\bar{c} = d\bar{d} - b\bar{b} = s, \quad (24)$$

where s is a real number, not zero.

To solve equations (21), (22), (23), (24), assume $a \neq o$ and put

$$ca^{-1} = -\mu, \quad \text{or} \quad c = -\mu a. \quad (25)$$

Substituting for c in the first of equations (21), we get

$$a\bar{a}\bar{\mu} + \mu a\bar{a} = 0,$$

showing that μ is pure imaginary. Substituting for c in the second of equations (23), we get

$$b = -\bar{\mu}d = \mu d. \quad (26)$$

Substituting for b and c in (22) and (24), we get

$$\bar{d}a + \bar{d}\mu^2 a = \bar{a}d + \bar{a}\mu^2 d = r,$$

$$a\bar{a}(1 + \mu^2) = (1 + \mu^2)d\bar{d} = s.$$

Now $1 + \mu^2$ cannot vanish, as r and s would then vanish, and hence

$$\bar{d}a = \bar{a}d, \quad a\bar{a} = d\bar{d}.$$

This gives

$$(\bar{d} + \bar{a})(d - a) = 0,$$

showing that

$$d = \pm a. \tag{27}$$

Equations (25), (26) and (27) with μ pure imaginary are sufficient to ensure that all the equations (21), (22), (23), (24) are satisfied.

If $a = 0$, we find from (21), (22), (23), (24) that $d = 0$ and $b = \pm c$. This solution may be looked upon as a special case of the solution provided by (25), (26) and (27) with μ infinitely great.

We now have the result that those of the transformations (6) that correspond to Lorentz transformations are the ones of the form

$$q^* = (aq \pm \mu a)(-\mu a q \pm a)^{-1}, \tag{28}$$

where a is an arbitrary quaternion, μ is a pure imaginary quaternion, and one must take both upper or both lower signs. The transformations (28) must, of course, form a group.

We may establish a connexion between the quaternion q and a vector in space-time in the following way. The linear transformations of the six X's that leave the quadric (16) and the planes (17) invariant will cause X_1, X_2, X_3 and X_4 to transform among themselves in such a way that the quadratic form $X_1^2 + X_2^2 + X_3^2 - X_4^2$ gets multiplied by a numerical factor, and indeed by the same factor that X_0^2 and X_5^2 get multiplied by. Thus the four quantities

$$\xi_\nu = X_\nu / X_5 \qquad (\nu = 1,2,3,4)$$

will transform like the components of a vector in space-time, as will the four quantities

$$\eta_\nu = X_\nu / X_0 \qquad (\nu = 1,2,3,4).$$

It is more convenient to work with the ξ's. We can express the ξ's in terms of the components of u and v by using equations (13), (14) and (15). We know, however, that the ξ's depend only on the ratio $uv^{-1} = q$, and hence we may substitute $u = q$ and $v = 1$ in (13). We then get

$$\xi_1 = \frac{2q_1}{1 - |q|^2}, \quad \xi_2 = \frac{2q_2}{1 - |q|^2}, \quad \xi_3 = \frac{2q_3}{1 - |q|^2}, \quad \xi_4 = \frac{1 + |q|^2}{1 - |q|^2}. \tag{29}$$

These equations make every quaternion q correspond to a vector ξ in space-time in such a way that a transformation (28) applied to q corresponds to a Lorentz transformation applied to ξ. We have

$$\begin{aligned} \xi_4^2 - \xi_1^2 - \xi_2^2 - \xi_3^2 &= \frac{1 + 2|q|^2 + |q|^4 - 4(q_1^2 + q_2^2 + q_3^2)}{(1 - |q|^2)^2} \\ &= 1 + \frac{4\, q_0^2}{(1 - |q|^2)^2}, \end{aligned} \tag{30}$$

and hence the vector ξ lies within the light-cone and is of length greater than or equal to unity. It is easily seen that every such vector corresponds to two quaternions q, differing only in the sign of their real part.

The general transformation (28) may be decomposed into the two simple transformations

$$q^* = (q^\dagger + \mu)(-\mu q^\dagger + 1)^{-1}, \tag{31}$$

$$q^\dagger = \pm a q a^{-1}, \tag{32}$$

as is easily verified. The simple transformation (32) leaves $|q|$ invariant and therefore leaves ξ_4 invariant. It thus corresponds to a Lorentz transformation that leaves the time axis unchanged. One can easily see that if (32) contains the $+$ sign it corresponds to a pure rotation in three-dimensional space, and if it contains the $-$ sign it corresponds to a rotation and a reflexion in three-dimensional space.

Let us now examine the simple transformation (31) and omit the $\dagger$, so that it reads

$$q^* = (q + \mu)(1 - \mu q)^{-1} \tag{33}$$

We have, remembering that μ is pure imaginary,

$$1 - |q^*|^2 = 1 - \frac{|q + \mu|^2}{|1 - \mu q|^2} = \frac{1 + \mu^2}{|1 - \mu q|^2}(1 - |q|^2). \tag{34}$$

Again

$$\begin{aligned} q^* &= (q + \mu)(1 + \bar{q}\mu) / |1 - \mu q|^2 \\ &= (q + \mu + |q|^2 \mu + \mu \bar{q} \mu) / |1 - \mu q|^2, \end{aligned} \tag{35}$$

so that

$$\begin{aligned} \mu q^* - q^* \mu &= \{\mu q - q\mu + \mu(\mu \bar{q} - \bar{q}\mu)\mu\} / |1 - \mu q|^2, \\ &= \{\mu q - q\mu - \mu(\mu q - q\mu)\mu\} / |1 - \mu q|^2, \\ &= (\mu q - q\mu)(1 + \mu^2) / |1 - \mu q|^2. \end{aligned}$$

Combining this with (34), we get

$$\frac{\mu q^* - q^* \mu}{1 - |q^*|^2} = \frac{\mu q - q\mu}{1 - |q|^2}. \tag{36}$$

Equation (34) shows that if $1 + \mu^2 < 0$, $1 - |q^*|^2$ and $1 - |q|^2$ are of opposite sign, so that ξ_4 changes sign under the transformation. The corresponding Lorentz transformation must then be one that interchanges future and past. Conversely, if $1 + \mu^2 > 0$, the corresponding Lorentz transformation is of the physical kind that does not interchange future and past. Equation (36) shows that components of the three-dimensional vector ξ_1, ξ_2, ξ_3 perpendicular to the three-dimensional vector μ_1, μ_2, μ_3 are invariant. Thus the Lorentz transformation corresponding to (33) with $1 + \mu^2 > 0$ must consist simply in giving things a velocity in the direction of μ_1, μ_2, μ_3, without giving them a rotation in three-dimensional space.

THE ADDITION OF VELOCITIES.

Let us study the transformation (33) for the case when q is pure imaginary. The corresponding vector ξ is then, according to (30), of unit length. It must transform into a vector of unit length and hence q^* must also be pure imaginary, a result which could also have been inferred directly from equation (35).

Let us suppose further that $|q| < 1$, so that, according to (29), $\xi_4 > 0$. A vector ξ of unit length with $\xi_4 > 0$ may be looked upon as specifying a velocity in space-time. Thus equations (29) *connect each pure imaginary quaternion of modulus less than unity with a velocity.* The zero quaternion is connected with zero velocity. The transformation (33) with $|\mu| < 1$ now shows how a Lorentz transformation applied to a velocity changes the quaternion connected with that velocity. The quaternion μ is itself connected with a velocity, namely that velocity which the Lorentz transformation gives to a thing previously at rest, since μ is the value of q^* when $q = 0$. Equation (33) now appears as *the law of addition of velocities, expressed in terms of the quaternions connected with those velocities.*

The velocity w which results from adding two velocities u and v in the same direction is given by the well-known formula of the theory of relativity

$$w = (u + v)(1 + uv)^{-1}, \tag{37}$$

if the velocity of light is taken as unity. Equation (37) is of the same algebraic form as (33), if q^*, q and μ are considered as the analogues of iw, iu and iv. However, this does not give the correct way in which (33) should be applied when the velocities are in the same direction. Instead, one should look upon u, v and w as hyperbolic tangents of angles and introduce u', v' and w', the hyperbolic tangents of half the angles, and then consider q^*, q and μ as iw', iu' and iv'. With this way of making the connexion, equation (33) appears as the generalization to three-dimensional space of the usual formula for the addition of velocities. The non-commutation of quaternions now takes into account the fact that the order of the velocities is important. The quaternion formulation appears to be the most suitable one for expressing generally the law of addition of velocities.

As an example of the foregoing methods, let us work out the three-dimensional rotation produced by first giving to a thing at rest a velocity connected with the quaternion μ_1, then giving it a velocity connected with the quaternion μ_2 and finally giving it a velocity which annuls the velocity it acquired by the first two operations. The first operation corresponds to the quaternion transformation

$$q' = (q + \mu_1)(1 - \mu_1 q)^{-1}.$$

The second coresponds to the quaternion transformation

$$\begin{aligned} q'' &= (q' + \mu_2)(1 - \mu_2 q')^{-1} \\ &= [q'(1 - \mu_1 q) + \mu_2(1 - \mu_1 q)][1 - \mu_1 q - \mu_2 q'(1 - \mu_1 q)]^{-1} \\ &= [q + \mu_1 + \mu_2 - \mu_2\mu_1 q][1 - \mu_1 q - \mu_2 q - \mu_2\mu_1]. \end{aligned} \tag{38}$$

This equation is of the form (28) with the plus sign and with $a = 1 - \mu_2\mu_1$. Let us decompose the transformation (38) into two simple transformations, as we did with (28). The simple transformation of the form (31) so obtained gets cancelled by the final operation which annuls the velocity produced by the first two operations, and we are left with a simple transformation of the form (32) with the plus sign and with the a value given above, i.e.

$$q^{\dagger} = (1 - \mu_2\mu_1)\, q\, (1 - \mu_2\mu_1)^{-1}. \tag{39}$$

This transformation gives the required rotation.