BIQUATERNION VECTORFIELDS

OVER MINKOWSKI SPACE

By

J. LAMBEK

Submitted in partial fulfillment of
the requirements for the degree of
Doctor of Philosophy. April, 1950.

TABLE OF CONTENTS.

INTRODUCTION.

When Hamilton had invented quaternions, the question arose whether they could be used to advantage in mathematical physics. However, the world then only had three dimensions, and so the scalar part of the quaternion was suppressed, the resulting entity being called a vector. The relation between vector-analysis and quaternion algebra is well known, and need not be entered into here.

With the theory of relativity also came the fourth dimension. Although Minkowski himself rejected the quaternionic calculus as "too narrow and clumsy for the purpose in question", Silberstein has strongly advocated the cause of quaternions. He used quaternions with imaginary scalar parts to designate the position of points or events in space-time. This was necessary, since the metric in Minkowski space is not given by a positive definite quadratic form. We achieve the same result by making the vector part imaginary, in which case we obtain a Hermitian matrix representation of the position quaternion. Professor Dirac believes (as stated by him in conversation) that, some day, Hamiltonian quaternions, as opposed to Hermitian quaternions, will re-assert themselves in relativity theory; but I do not see how this can be.

To each point P of Minkowski space we assign a Hermitian quaternion S(P). The function S is called a special coordinate system. Given S, a physical entity will in general be measured by a function F_S which associated a biquaternion $F_S(P)$ with each

point P. However, it is the correspondence $S \rightarrow F_S$, here called a vectorfield, which really represents the physical entity, independently of any particular coordinate system (within the frame-work of special relativity). It must be postulated that, in some way, F_S does not loose its identity as we transform the coordinate system. That is, as we pass from S to S', the transformation carrying F_S into $F_{S'}$ shall be of a rather simple nature. In the language of Van der Waerden: The latter transformation shall be a representation of the former.

In a relativistic treatment of classical physics, such as Silberstein's (see bibliography), we find five types of vectorfields: One invariant, two four-vectors, and two six-vectors. If we also use the quaternionic method in discussing relativistic quantum-mechanics, four more vectorfields arise, which have been called wave-vectors (or spinors). Thus nine types of vectorfields can be found a posteriori. In part II, I shall define the concept of vectorfield, and try to discover whether other types than the above nine exist. I shall also extend this notion from functions to operators.

Unfortunately, when working on this problem, I was not aware of the theory of representations of the Lorentz Group, as developed by Van der Waerden, whose work is based on that of G. Frobenius, I. Schur, and H. Weyl. Roughly speaking, the results of part II are implied by those of Van der Waerden (see bibliography)

To be more specific: He determines all differentiable representations of the Lorentz Group as linear transformations of n-dimensional vector space, but states in a footnote (§17) that it would suffice to demand continuity instead of differentiability In this thesis, I have determined all continuous representations (here called transformation schemes) of the Lorentz Group as orthogonal transformations of biquaternion space, i.e. Euclidian space with four complex dimensions. The problem solved here is therefore not quite the same as the problem solved by Van der Waerden, but loses in importance owing to his work. His method is more elegant and more general than mine, but less elementary, as he makes use of infinitesimal transformations.

Part I contains a discussion of special coordinate transformations in Minkowski space and biquaternion space. Vectors are regarded as labels of points, as in the theory of relativity, and not as points, as is customary in algebra. It is proved that Lorentz transformations can be expressed by quaternions, a fact which is assumed, or rather postulated, by Silberstein. The proof given here is long; but an attempt to shorten it (for instance by using an exponential parametrization) might sacrifice its elementary character.

In part III, I have obtained explicitly all continuous homomorphisms of the Lorentz group on itself or part of itself, but I do not know whether there exist discontinuous ones. Dieudonné, in his book on classical groups (see bibliography), proves that the Lorentz group is simple, i.e. that all such

homomorphisms map the whole group either into itself or into unity. The result of part III would be implied by this, if it were known that all automorphisms of the Lorentz group are little more than inner automorphisms.

In this introduction reference has been made to physical theory. Nonetheless, the problem will be posed and solved in purely mathematical terms. I had intended to add an appendix with physical applications, especially to Dirac's relativistic equation of the electron, but was dissuaded from doing so, owing to the recent publication of a paper by A.W. Conway on this very subject (see bibliography).

On the whole, the present paper has been worked out in the spirit of Silberstein and Conway, and no attempt has been made to assimilate its results to the terminology of Van der Waerden and Dieudonné.

Part I.

It is customary to define quaternions through their rules of multiplication, and to show that the resulting system is an algebra. By means of a well known theorem, a matrix representation of quaternions is then obtained. To simplify matters we shall reverse the usual procedure here. A matrix

$$a = \begin{pmatrix} a_0 & -a_1 & -a_2 & -a_3 \\ a_1 & a_0 & -a_3 & a_2 \\ a_2 & a_3 & a_0 & -a_1 \\ a_3 & -a_2 & a_1 & a_0 \end{pmatrix} \qquad \ldots \quad (i)$$

with complex elements a_n $(n=0,1,2,3)$ will be called a biquaternion. We can write $a=\sum_{n=0}^{3} a_n i_n$, where the i_n are certain matrices all of whose elements are 0, 1, or -1. In particular, i_0 is the identity matrix, and we shall write $i_0=1$. Furthermore $i_n{}^2=-1$ for $n=1,2,3$, and $i_1 i_2 = i_3 = -i_2 i_1$, etc. Since matrix multiplication is associative and distributive, and since 0 and 1 are biquaternions, the latter will form an algebra. This is however not commutative, for $i_1 i_2 - i_2 i_1 = 2i_3 \neq 0$; nor is it a division algebra, as $(1+ii_1)(1-ii_1)=0$.

The transpose a^T of a is formed by interchanging rows and columns in (i), thus $a^T = a_0 - a_1 i_1 - a_2 i_2 - a_3 i_3$. The conjugate a^C

of a is formed from (i) by replacing each element by its complex conjugate. Among other obvious relationships we have:

$$a^{TT}=a=a^{CC},\ a^{TC}=a^{CT},\ (ab)^{C}=a^{C}b^{C},\ (ab)^{T}=b^{T}a^{T}.$$

If $a^T=a$ or $-a$, we speak of scalars or vectors respectively. a will be real or purely imaginary, according as $a^C=a$ or $-a$. Real biquaternions are of course Hamiltonian quaternions. As in the theory of matrices, we shall consider Hermitian or skew-Hermitian matrices, for which $a^{CT}=a$ or $-a$ respectively.

If a and b are two biquaternions, their inner product $(a,b)=\sum_{n=0}^{3}a_n b_n=(1/2)(ab^T+ba^T)$ is clearly a scalar. Inner multiplication is distributive and commutative, but not associative. $N(a)=(a,a)=aa^T=a^Ta$ is called the norm of a, its square being the determinant of (i). We have $N(ab)=N(a)N(b)$, but

$$N(a\pm b)-N(a)-N(b) = \pm 2(a,b)\ . \qquad \text{... (ii)}$$

Let B be a set of points, over which there is defined a function J, associating a complex number $J(Q,Q')$ with each pair of points Q, Q' of B. Furthermore let there exist a one to one correspondence $S:\ Q\leftrightarrow y$, between the set B and the set of biquaternions y such that $J(Q,Q')^2=N(y-y')$. Then B will be called a biquaternion space. It is of course merely a four-dimensional space with Euclidian metric and complex coordinates. S is called a special coordinate system of B. We shall write $y=S(Q)$. If S' is another special coordinate system, and if $y'=S'(Q)$, then we can write $y'=T(y)$, where $T=S'S^{-1}$.

T is called a special coordinate transformation. It is characterized by the existence of $T^{-1}=SS'^{-1}$ and by the equation:

$$N(T(y)-T(y'))=N(y-y') \,, \qquad \ldots \qquad \text{(iii)}$$

which must hold for all biquaternions y, y'. Since with S, S', and S'' the correspondence $SS'^{-1}S''$ will also be a special coordinate system, it follows that the set of all special coordinate systems of B is a group, its unit element being the identical transformation $I=SS^{-1}$.

<u>Theorem 1</u> :If T is any special coordinate transformation of biquaternion space, then either $T(y)=pyq+b$ or $T(y)=py^Tq+b$ for all biquaternionsy y, where b, p, q are biquaternions, and $N(p)=N(q)=1$.

<u>Proof</u> : We shall call that point O for which $S(O)=0$ the origin of S. Now any special coordinate transformation can be expressed as the product of a translation $T(y)=y+b$, and an orthogonal transformation which leaves the origin invariant.

If T is an orthogonal transformation of B, we thus have $T(0)=0$. Hence by (iii) $NT(y)=N(y)$. In view of (ii) therefore

$$(T(y),T(y'))=(y,y')$$

for all biquaternions y, y'. Hence

$$(T(y),i_n)=(y,T^{-1}(i_n))=\sum_{m=0}^{3}y_m(i_m,T^{-1}(i_n))=\sum_{m=0}^{3}y_m(T(i_m),i_n)=(\sum_{m=0}^{3}y_mT(i_m),i_n$$

for n=0,1,2,3. Thus $T(y)=\sum_{m=0}^{3}y_mT(i_m)$, as the n-th components of both sides are equal.

Instead of $y'=T(y)$ we may therefore write $y'_m=\sum_{n=0}^{3} t_{mn} y_n$, $(m=0,1,2,3)$. In view of the invariance of $N(y)=\sum_{n=0}^{3} y_n^{\ 2}$, it follows that $\sum_{m=0}^{3} t_{mk} t_{mn}=\delta_{kn}$ $(k,n=0,1,2,3)$. If we introduce the matrix $t=(t_{mn})$, this can be written $t^T t=1$. Hence $\det t = \det t^T = 1$ or -1. Now the orthogonal transformation $T(y)=y^T$ is easily seen to have determinant -1, and any orthogonal transformation with negative determinant can be written as the product of this particular transformation and one whose determinant is positive. We may thus limit our enquiry to the case $\det t = 1$.

We shall assume at first that $\det(1+t)\neq 0$. Consider

$$s=(1+t)^{-1}(1-t)=(1-t)(1+t)^{-1} , \qquad \ldots \qquad \text{(iv)}$$

the two expressions on the right being equal, since

$$(1-t)(1+t)=1-t^2=(1+t)(1-t) .$$

We find that $s^T=-s$, so that s is skew-symmetric, whence

$$s = \begin{pmatrix} 0 & -s_1 & -s_2 & -s_3 \\ s_1 & 0 & -s'_3 & s'_2 \\ s_2 & s'_3 & 0 & -s'_1 \\ s_3 & -s'_2 & s'_1 & 0 \end{pmatrix} \qquad \ldots \qquad \text{(v)}$$

Let s' be obtained from s by interchanging the s_n and the s'_n in (v), then $s's=-\sum_{n=1}^{3} s_n s'_n$ is a scalar. Now $(1+s)t=1-s$. Multiplying both sides of this by $(1+s')$ on the left, we get

$$(1+s's+s'+s)t = (1-s's+s'-s). \qquad \ldots \qquad \text{(vi)}$$

Let $u_0=1+s's$, $v_0=1-s's$; $u_n=s_n+s'_n$, $v_n=s_n-s'_n$, $(n=1,2,3)$.

If $y'_m=\sum_{n=0}^{3} t_{mn} y_n$, $(m=0,1,2,3)$, the above implies that

$$\begin{pmatrix} u_0 & -u_1 & -u_2 & -u_3 \\ u_1 & u_0 & -u_3 & u_2 \\ u_2 & u_3 & u_0 & -u_1 \\ u_3 & -u_2 & u_1 & u_0 \end{pmatrix} \begin{pmatrix} y'_0 \\ y'_1 \\ y'_2 \\ y'_3 \end{pmatrix} = \begin{pmatrix} v_0 & -v_1 & -v_2 & -v_3 \\ v_1 & v_0 & v_3 & -v_2 \\ v_2 & -v_3 & v_0 & v_1 \\ v_3 & v_2 & -v_1 & v_0 \end{pmatrix} \begin{pmatrix} y_0 \\ y_1 \\ y_2 \\ y_3 \end{pmatrix} \quad \ldots \quad \text{(vii)}$$

which, in quaternion notation, becomes simply $uy'=yv$. A calculation will show that s and s' have the same characteristic equation. Now $1+s=2(1+t)^{-1}$, so that $\det(1+s')=\det(1+s)\neq 0$. We may therefore write $y'=u^{-1}yv$. Since $N(y)$ is invarient, $N(u)=N(v)$, and we have without loss of generality $y'=pyq$, where both p and q are of norm 1.

We have assumed above that $\det(1+t)\neq 0$. If $\det(1-t)\neq 0$, we could have obtained the same result, by writing instead of (iv)
$$s=(1-t)^{-1}(1+t)=(1+t)(1-t)^{-1},$$
and proceeding in a similar fashion. It remains to consider the exceptional case: $\det(1+t)=\det(1-t)=0$.

Let x be a scalar variable, then $f(x)=\det(x-t)$ is a scalar polynomial in x of degree 4, with leading coefficient 1. Now $x^4 f(1/x)=\det(1-xt)=\det(t)\det(t^T-x)=\det(t-x)=f(x)$. By elementary algebra it follows that the equation $f(x)=0$ has roots $a,1/a,b,1/b$. But 1 and -1 are given roots, so that $f(x)=(x^2-1)^2$. Since every matrix satisfies its own characteristic equation, $f(t)=(t^2-1)^2=0$, whence $(t-t^T)^2=0$, and consequently $(t-t^T)^2=(t-t^T)^2+2(tt^T+t^Tt)=4$.

Let us put $r=(1/2)(t+t^T)$, $s=(1/4)(t+t^T)(t-t^T)$. Clearly r is symmetric, and s is skew-symmetric. Moreover it follows from the foregoing that

$$t=r(1+s),\quad r^2=1,\quad s^2=0\ . \qquad \ldots\ldots \quad \text{(viii)}$$

If we can show that t yields a transformation $y'=pyq$, $N(p)=N(q)=1$, in the two special cases $t=r$ and $t=1+s$, it will follow that $t=r(1+s)$ also gives such a transformation, since the latter clearly form a group.

First assume that $t=1+s$. If s is expressed as in (v), the condition $s^2=0$ implies that $\sum_{n=1}^{3} s_n{}^2=0$ and $s_n=\pm s'_n$. If s' is defined as in (vi), this means that $s's=0$ and $s=\pm s'$. When $s=s'$, $t=1+s$ is a biquaternion of norm 1, and our transformation becomes $y'=py$, $N(p)=1$. Similarly, when $s'=-s$, a glance at (vii) will convince us that we have the transformation $y'=yq$, $N(q)=1$.

We may thus confine attention to the case $t=r$, i.e. we may assume that t is symmetric and, in view of (viii), that $t^2=1$. We also still have that the determinant of t is 1, since the omitted factor (1+s) clearly has determinant 1. We may exclude from consideration the trivial cases $t=1$ and $t=-1$. Under those circumstances it can then be shown by algebra that there exist four column vectors z_n (n=0,1,2,3) such that

$$tz_0=z_0,\ tz_1=-z_1,\ tz_2=z_2,\ tz_3=-z_3$$

and furthermore $z_m{}^T z_n=1$ when $m=n$, $=0$ when $m\neq n$. If we put $s=(z_1z_3{}^T-z_3z_1{}^T)$, it follows that s is skew-symmetric. A simple calculation will show that, in the notation of (vi), $s'=(z_0z_2{}^T-z_2z_0{}^T)$. By (ix), $st=-s$, $s't=s'$, whence $(s'+s)t=s'-s$. Let $u_0=v_0=0$; $u_n=s_n+s'_n$, $v_n=s_n-s'_n$, (n=1,2,3), then as in (vii) we obtain the transformation $uy'=yv$. Now $u=s'+s$ so that $N(u)=-(s'+s)^2=-s'^2-s^2=\sum_{n=0}^{3} z_nz_n{}^T=1$. Hence $y'=u^{-1}yv=pyq$, $N(p)=N(q)=1$.

Q.E.D.

Let M be a set of points over which there is defined a function J associating a complex number J(P,P') with each pair of points P, P' of M. Furthermore let there exist a one to one correspondence S: $P \leftrightarrow x$, between M and the set of Hermitian biquaternions x , such that

$$J(P,P')^2 = N(x-x') .$$

Then M will be called a Minkowski space. We notice that $J(P,P')^2$ is always real, but it can be negative or zero. As in the case of biquaternion space we define special coordinate systems and transformations.

As above, any special coordinate transformation T can be factored, and we need only consider the translation $x'=x+a$, where a must now be Hermitian, the transformation $x'=x^T$, and the transformation $x'=pxq$, where $N(p)=N(q)=1$. However now p and q will have to be restricted in some way, so that pxq will be Hermitian for any Hermitian x. Thus $x^{CT}=x$ should imply that $q^{CT}xp^{CT}=(pxq)^{CT}=pxq$. Putting $u=q^Cp$, we thus obtain for all Hermitian x, $xu^{CT}=ux$. Since $x=1$ is Hermitian, $u^{CT}=u$, whence $xu=ux$. Since $x=ii_n$ (n=1,2,3) is Hermitian, $i_nu=ui_n$. But this implies that u is a scalar, as is easily shown. Clearly $N(u)=1$, whence $u=\pm 1$ and $q^C=\pm p^T=r^T$, say. Therefore $pxq= rxr^{CT}$. We thus have the following

<u>Theorem 2</u> : If T is any special coordinate transformation of Minkowski space, then $T(x)=\pm rx^{(T)}r^{CT}+a$, $N(r)=1$, a and r being biquaternions, a Hermitian.

Part II.

Let S_o, S'_o be two given special coordinate systems of M and B respectively. From now on we shall confine ourselves to coordinate systems $S=TS_o$ and $S'=T'S'_o$, where $T(x)=rxr^{CT}$ and $T'(y)=pyq$, p,q, and r being biquaternions of norm 1. We shall speak of restricted coordinate systems and transformations.

Consider a correspondence F from M to B: $Q=F(P)$. If $x=S(P)$ and $y=S'(Q)$, this can be written: $y=S'FS^{-1}(x)$. Let a correspondence $S \to S'$ be given. For any given S we may then consider the quaternion function $S'FS^{-1}$ instead of the point function F. Here however we shall be interested in the function $F_S=S'F$, which associates a biquaternion y with any point P of Minkowski space. We then have $F_{TS}=T'F_S$, where $T'=(TS)'S'^{-1}$ is some restricted coordinate transformation of B, depending on T, and perhaps on S.

Suppose now that $F_{TS}(P)$ can be calculated if only $F_S(P)$ and T are known. We shall not considerably strengthen this condition if we stipulate that T' is independent of S. Then

$$(T_1T_2)'=(T_1T_2S)'S'^{-1}=(T_1T_2S)'(T_2S)'^{-1}(T_2S)'S'^{-1}=T_1'T_2' .$$

A correspondence $T \to T'$ which is multiplicative in this sense will be called a transformation scheme from M to B. We shall attempt to find all such transformation schemes.

Let $T \to T'$ be a given transformation scheme. We may write $T(x)=rxr^{CT}$, $T'(y)=pyq$, where p, q, and r are biquaternions of norm 1, and $p=P(r)$, $q=Q(r^{CT})$. In virtue of the multiplicative property

$$P(r)P(r')yQ(r'^{CT})Q(r^{CT})=P(rr')yQ(r'^{CT}r^{CT})$$

for all biquaternions y. If we abbreviate

$$a=P(rr')^{-1}P(r)P(r'),\quad b=Q(r'^{CT}r^{CT})Q(r^{CT})^{-1}Q(r'^{CT})^{-1},$$

this becomes: $ay=yb$, for all y. Putting $y=1$, we find that $a=b$, and putting $y=i_n$ $(n=1,2,3)$, we show that their common value is a scalar. Now $N(a)=1$, hence $a=b=1$ or -1. We have thus proved

<u>Theorem 3</u> : If $T \to T'$ is a transformation scheme from M to B, with $T(x)=rxr^{CT}$, then there exist functions P and Q, such that $T'(y)=P(r)yQ(r^{CT})$, with the property that $P(rr')=\pm P(r)P(r')$ and $Q(r'^{CT}r^{CT})=\pm Q(r'^{CT})Q(r^{CT})$, the sign being the same in both cases, but perhaps depending on r, r'.

We may of course replace the P and Q of theorem 3 by P' and Q', where $P'(r)=D(r)P(r)$, $Q'(r^{CT})=D(r)Q(r^{CT})$, and $D(r)=\pm 1$. To indicate the dependence of the sign in theorem 3 on r and r', let us call it $E(r,r')$. The question arises, can we select the function D in such a way that $D(rr')=E(r,r')D(r)D(r')$? For then, if the equations of theorem 3 are restated in terms of P' and Q', the ambiguity in sign will disappear, in fact P' and Q' will be multiplicative functions, and our problem will be greatly simplified.

A calculation shows that P and Q can be replaced by multiplicative functions if and only if $P(s)^2=P(s')^2P(s'')^2$ for all biquaternions s,s',s'' of norm 1 such that $s^2=s'^2s''^2$. It is difficult to imagine what can be done with this condition. Perhaps it is safer to realize at this stage that our original aim, namely to find all transformation schemes, was too ambitious.

If it so happens that the functions P and Q of theorem 3 can be chosen to be continuous functions (i.e. continuous in the four complex components of r) , we shall call $T\rightarrow T'$ a continuous transformation scheme. We shall limit our enquiry to continuous transformation schemes from M to B.

We may then assume that P(r) is a continuous function of r. Hence for given r', $E(r,r')=P(rr')/(P(r)P(r'))$ is a continuous function of r. Similarly, for given r, it is a continuous function of r'. Its value being always 1 or -1, it is therefore a constant, namely 1 or -1. Without loss of generality, we need only consider the positive case, whence $P(rr')=P(r)P(r')$, a similar equation holding for Q. For if we had the minus sign, we could write $P'=-P$, $Q'=-Q$, and both P' and Q' would be multiplicative. Thus

<u>Theorem 3'</u> : If $T\rightarrow T'$ is a continuous transformation scheme from M to B, with $T(x)=rxr^{CT}$, then there exist continuous multiplicative functions P and Q, which map the set of all biquaternions of norm 1 on itself or part of itself, such that $T'(y)=P(r)yQ(r^{CT})$.

To go on from here we need a lemma, which, for the sake of continuity, will only be proved in part III.

Lemma : If F is any continuous and multiplicative function which maps the set of all biquaternions of norm 1 on itself or part of itself, then F is one of the three functions:

$$F(r)=1 \ , \ F(r)=uru^T \ , \ F(r)=ur^C u^T \ ,$$

u being a constant biquaternion of norm 1.

In view of this lemma and theorem 3', we immediately have

Theorem 4 : If $T \to T'$ is a continuous transformation scheme from M to B, with $T(x)=rxr^{CT}$, then T' is one of the following nine types of functions:

$$T'(y)=y \ , \ =ur^{(C)}u^T y \ , \ =yvr^{(C)T}v^T \ , \ =ur^{(C)}u^T yvr^{(C)T}v^T \ ,$$

where u and v are constant biquaternions of norm 1.

Of particular interest will be those cases in which the u and v of the above theorem are 1, and then we shall say that the transformation scheme is principal. Hence any principal continuous transformation scheme is determined by the fact that T'(y) is one of the following nine expressions:

$$y \ , \ ry \ , \ r^C y \ , \ yr^T \ , \ yr^{CT} \ , \ ryr^T \ , \ r^C yr^T \ , \ ryr^{CT} \ , \ r^C yr^{CT} \ .$$

Any continuous transformation scheme can be reduced to a principal one, by means of a fixed orthogonal transformation of determinant 1 of biquaternion space. For by theorem 4 we have: $T'(y)=uP(r)u^T yvQ(r^T)v^T$, where $P(r)=1, r$, or r^C, and $Q(r^T)=1, r^T$, or r^{CT}. Hence $u^T T'(y)v=P(r)(u^T yv)Q(r^T)$.

If we write $y'=U(y)=u^T yv$, this becomes: $UT'U^{-1}(y')=P(r)y'Q(r^T)$. Here U is an orthogonal transformation of determinant 1 of B, and the transformation scheme $T \to UT'U^{-1}$ is principal.

Let S be a restricted coordinate system of M. To each such S let there correspond a function F_S which assigns a biquaternion $y=F_S(P)$ to each point P of M. The correspondence $S \to F_S$ will be called a vectorfield, provided there exists a continuous transformation scheme $T \to T'$ such that $F_{TS}=T'F_S$. In virtue of theorem 4, we thus have

<u>Theorem 4'</u> : The correspondence $S \to F_S$ is a vectorfield, provided, when $T(x)=rxr^{CT}$, then $F_{TS}(P)=P(r)F_S(P)Q(r^{CT})$, where $P(r)=1$ or $=ur^{(C)}u^T$, and $Q(r^{CT})=1$ or $=vr^{(C)T}v^T$, u and v being fixed biquaternions of norm 1.

If the transformation scheme belonging to the vectorfield is principal, we shall also say that the vectorfield is principal. We note in particular that the identical correspondence $S \to S$ is a principal vectorfield with transformation scheme $T \to T'$.

We shall state and prove a number of simple properties of vectorfields. For convenience, the transformation scheme of theorem 4' will be denoted by $(P(r);Q(r^{CT}))$.

(1) If $S \to F_S$ is a vectorfield, its conjugate $S \to F_S{}^C$ and transpose $S \to F_S{}^T$ are also vectorfields, with respective transformation schemes $(P(r)^C; Q(r^{CT})^C)$ and $(Q(r^{CT})^T; P(r)^T)$.

(2) If $S \to F_S$ and $S \to G_S$ are two vectorfields with the same transformation scheme, then their sum $S \to (F_S + G_S)$ is also a vectorfield, still having the same transformation scheme.

(3) If $S \to F_S$ and $S \to F'_S$ are two vectorfields, such that $P'(r) = Q(r^{CT})^T$, then their product $S \to (F_S F'_S)$ is also a vectorfield, with transformation scheme $(P(r); Q'(r^{CT}))$.

(4) A vectorfield is called an invariant if its transformation scheme is (1;1). Each component $S \to (F_S)_n$ (n=0,1,2,3) of an invariant vectorfield is itself an invariant.

(5) Every vectorfield can be expressed as a principal vectorfield, multiplied on both sides by constant invariants of norm 1. For suppose $F_{TS}(P) = upu^T F_S(P) vqv^T$, where $p = 1, r, r^C$, and $q = 1, r^T, r^{CT}$. Put $F_S(P) = uG_S(P)v^T$, then the above becomes $G_{TS}(P) = pG_S(P)q$, hence $S \to G_S$ is a principal vectorfield. Moreover $S \to u$ and $S \to v^T$ are constant invariants.

(6) If a vectorfield is transformed by $(p; p^{CT})$, its hermitian and skew-hermitian parts are also vectorfields with the same transformation schemes. This follows from (1) and (2). The two parts are called four-vectors in physics. Of particular interest are the principal vectorfields transformed by $(r; r^{CT})$ or $(r^C; r^T)$.

(7) If a vectorfield is transformed by $(p;p^T)$, then its scalar and vector parts are also vectorfields with the same transformation scheme. Again this follows from (1) and (2). The scalar part is easily seen to be an invariant, and the vector part is what the physicist calls a six-vector. Of particular interest are the principal vectorfields transforming by $(r;r^T)$ or $(r^C;r^{CT})$.

Presumably the converses of (4), (6), and (7) are also true; but it would be beyond the scope of our present exposition to attempt their proofs. At any rate, they would read as follows: A scalar vectorfield is an invariant; a Hermitian vectorfield is a four-vector or an invariant; a vector vectorfield is a six-vector or an invariant. It is difficult to say to what extent converses of (2) and (3) might hold.

In a way, a vectorfield is a function plus the manner in which it transforms. We shall now try to generalize the concept of vectorfield from functions to operators. But first we need the quaternion analogue of linear operators.

Consider a correspondence A: (F,G)→H, where F,G,H are functions that assign biquaternions to all points of Minkowski space. We shall write H=(FAG). A will be called an operator biquaternion in M, provided the following three postulates are satisfied:

(I) (FAG)=(AFG) whenever F is a scalar function.

(II) ((F+G)AH)=(FAH)+(GAH) ; (FA(G+H))=(FAG)+(FAH) .

(III) (aFAGb)=a(FAG)b, for all constant a and b.

It is easily seen that the above three postulates constitute necessary and sufficient conditions in order that it should be possible to write $A=\sum_{n=0}^{3}A_n i_n$, the A_n being linear operators in the usual sense. Thus, in a way, an operator biquaternion is merely a biquaternion whose components are linear operators.

Suppose for each restricted coordinate system S of M we have an operator biquaternion A_S. Such a correspondence $S\to A_S$ will be called an operator vectorfield, provided, whenever $S\to F_S$ and $S\to G_S$ are any two vectorfields with given transformation schemes, then $S\to(F_SA_SG_S)$ is again a vectorfield.

Let (p';q') , (p'';q'') be the two given transformation schemes, and (p;q) the resulting one. Then

$$(p'F_Sq'A_{TS}p''G_Sq'') = p(F_SA_SG_S)q \; . \qquad \ldots \qquad \text{(i)}$$

If we replace F_S by aF_S and G_S by G_Sb in (i), then in virtue of (III):

$$p'ap'^{-1}(p'F_Sq'A_{TS}p''G_Sq'')q''^{-1}bq''=pa(F_SA_SG_S)bq \; . \ldots \text{(ii)}$$

From (i) and (ii) we obtain

$$p'ap'^{-1}p(F_SA_SG_S)qq''^{-1}bq'' = pa(F_SA_SG_S)bq \; . \quad \ldots \qquad \text{(iii)}$$

We shall assume that the norm of $(F_SA_SG_S)$ is not identically zero. Taking b=1, we therefore find that $p'ap'^{-1}p=pa$, i.e. $p^{-1}p'a=ap^{-1}p'$. Hence $p^{-1}p'$ commutes with all biquaternions, it is therefore a scalar. But its norm is 1, hence p'=p or -p. Now p' and p are both functions of r,

and if we take $r=1$, both will reduce to 1. Thus the minus sign is excluded, hence $p'=p$, and similarly $q''=q$. In view of (III), we may therefore cancel p' against p on the left, and q'' against q on the right of (i). Thus $(F_S q' A_{TS} p'' G_S) = (F_S A_S G_S)$, or in operator notation $q' A_{TS} p'' = A_S$, that is $A_{TS} = q'^T A_S p''^T$. We might say that $S \to A_S$ transforms by the transformation scheme $(q'^T; p''^T)$.

If $N(FAG)(P)=0$ for all points P of M and all functions F, G, for which A is defined, we shall say that A is singular. Singular operator biquaternions are not of sufficient interest to us to deserve special consideration here. We shall therefore be satisfied with

<u>Theorem 5</u> : If $S \to A_S$ is a non-singular operator vectorfield, then $T(x)=rxr^{CT}$ implies $A_{TS}=P(r)A_S Q(r^{CT})$, where $P(r)=1$ or $=ur^{(C)}u^T$, and $Q(r^{CT})=1$ or $=vr^{(C)T}v^T$, u and v being biquaternions of norm 1.

A function F which assigns a biquaternion $y=F(P)$ to each point P of Minkowski space may itself be regarded as an operator biquaternion over M, if we write $(GFH)(P)=G(P)F(P)H(P)$. Under this interpretation, a vectorfield is an operator vectorfield, and instead of $F_{TS}(P)=pF_S(P)q$ we may write, in operator notation, $F_{TS}=pF_S q$ (which is what we have done anyway in the proof of theorem 5, for reasons of conciseness).

Conjugate, transpose, and sum of operator biquaternions may be defined in the obvious manner, thus:

$$(FA^C G)=(F^C A G^C)^C, \quad (FA^T G)=(G^T A F^T)^T, \quad (F(A+B)G)=(FAG)+(FBG) .$$

The above definitions clearly become identities if the operators are replaced by functions. Moreover it is easily verified that A^C, A^T, $A+B$ satisfy postulates I to III, and are therefore in fact operator biquaternions.

There seems to be no obvious way of defining the product AB of any two operator biquaternions A and B. However, in some special cases no difficulties are met. For instance, if we abbreviate (1AF) as (AF), we can also write (ABF)=(A(BF)).

The results which were stated about vectorfields above, with the exception of (3), also hold about operator vectorfields. (3) must of course be omitted here, owing to the absence of a product of operator biquaternions.

Part III.

Lemma : If F is any continuous and multiplicative function which maps the set of all biquaternions of norm 1 on itself or part of itself, then F is one of the three functions:

$$F(r)=1 \ , \ F(r)=uru^T \ , \ F(r)=ur^Cu^T \ ,$$

u being a constant biquaternion of norm 1.

Proof : We want to find all functions F such that

(i) F(r) is defined for all biquaternions r of norm 1,

(ii) F(r) is a biquaternion of norm 1,

(iii) F is multiplicative, i.e. $F(rr')=F(r)F(r')$ for all biquaternions r, r' of norm 1,

(iv) F(r) is a continuous function of the four complex components of r.

We shall derive a number of propositions from these four assumptions, without referring explicitly to the latter.

(1) $F(1)=1$. For $F(1)F(r)=F(1.r)=F(r)$.

(2) $F(r)^T=F(r^T)$. For $F(r)F(r)^T=NF(r)=1=F(1)=FN(r)=F(rr^T)=F(r)F(r^T)$.

(3) $F(-1)=1$ or -1 . For $F(-1)^T=F(-1)$ by (2), so that F(-1) is a scalar. Moreover $F(-1)^2=F(1)=1$, whence (3).

According to (3) we shall consider two cases.

Case A : $F(-1)=1$.

(4) If r is a vector, then $F(r)=1$ or -1 .
For let $r^T=-r$, then $F(r)^T=F(r^T)=F(-r)=F(-1)F(r)=F(r)$, so that F(r) is a scalar. But $NF(r)=1$, whence (4).

(5) If r is any biquaternion of norm 1, then there exist three complex scalars a_n (n=1,2,3) such that

$$r=\prod_{n=1}^{3}(\cos a_n+i_n\sin a_n) .$$

For let $\tan 2a_3=2(r_0r_3-r_1r_2)/(r_0^2+r_1^2-r_2^2-r_3^2)$, and put $s=r(\cos a_3-i_3\sin a_3)$. Then

$$s_0s_3-s_1s_2=\cos 2a_3(r_0r_3-r_1r_2)-\sin 2a_3(r_0^2+r_1^2-r_2^2-r_3^2)/2 = 0.$$

Let $\tan a_1=s_1/s_0$, $\tan a_2=s_2/s_0$, then $\tan a_1\tan a_2=s_1s_2/s_0^2=s_3/s_0$. Hence $\sec^2a_1\sec^2a_2=N(s)/s_0^2=N(r)/s_0^2=1/s_0^2$, so that, without loss of generality, $\cos a_1\cos a_2=s_0$. Thus

$$s=\cos a_1\cos a_2(1+i_1\tan a_1+i_2\tan a_2+i_3\tan a_1\tan a_2)=\prod_{n=1}^{2}(\cos a_n-i_n\sin a_n) .$$

Hence $r=s(\cos a_3+i_3\sin a_3)$ can be factored as desired.

(6) $F(\cos a_n+i_n\sin a_n)=1$ or -1 . This follows from (4), since $\cos a_n+i_n\sin a_n$ is the product of two vectors of norm 1; for instance $\cos a_1+i_1\sin a_1=i_2(i_3\sin a_1-i_2\cos a_1)$.

(7) $F(\cos a_n+i_n\sin a_n)=1$. We merely apply (6) to $\cos(a_n/2)+i_n\sin(a_n/2)$, and note that the square of this expression is $\cos a_n+i_n\sin a_n$.

(8) If $F(-1)=+1$, then $F(r)=1$ identically. This is a direct consequence of (5), (7), and the multiplicative property.

Having completed case A, we shall now consider

<u>Case B</u> : $F(-1)=-1$.

(9) If r is a vector, so is F(r) .

For let $r^T=-r$, then $F(r)^T=F(r^T)=F(-r)=F(-1)F(r)=-F(r)$.

(10) If $F(i_n)=j_n$, (n=1,2,3) , then $j_n^2=-1$, $j_1j_2=j_3=-j_2j_1$ etc.

For by (9) j_n is a vector, so that $j_n^T=-j_n$, whence $j_n^2=-N(j_n)=-1$. Also $j_1j_2=F(i_1)F(i_2)=F(i_1i_2)=F(i_3)=j_3$, etc.

(11) There exist three functions $b=f_n(a)$ modulo 2π, $(n=1,2,3)$, such that $F(\cos a+i_n\sin a)=\cos b+j_n\sin b$.

For let $r=\cos a+i_n\sin a$, $m\neq 0,n$, then $(r,i_m)=0$, that is $i_mr=r^Ti_m$. But then $j_mF(r)=F(r)^Tj_m$, that is $(F(r),j_m)=0$, whence $F(r)=k+j_nl$, k and l being scalars. Now $k^2+l^2=NF(r)=1$, hence we may write $k=\cos b$, $l=\sin b$, which proves (11) .

(12) $f_n(a+b)=f_n(a)+f_n(b)$ and $f_n(\pi/2)=\pi/2$, modulo 2π.

The first of these equations follows from the multiplicativity of F and de Moivre's theorem, the second one from the fact that $F(i_n)=j_n$. As an obvious corollary we have $f_n(ka)=kf_n(a)$ for integral k.

(13) $f_1(a)=f_2(a)=f_3(a)=f(a)$, say.

Let c be any complex number, $a_1=c/2$, $a_2=(\pi-c)/2$, $a_3=\pi/4$. If $r=\prod_{n=1}^{3}(\cos a_n+i_n\sin a_n)$, then $F(r)=\prod_{n=1}^{3}(\cos b_n+j_n\sin b_n)$, where $b_n=f_n(a_n)$, by (11). Now $r_0=\prod_{n=1}^{3}\cos a_n-\prod_{n=1}^{3}\sin a_n=(1/\sqrt{2})\cos(a_1+a_2)=0$. Hence r is a vector, and therefore F(r) is a vector, by (9). Thus $\prod_{n=1}^{3}\cos b_n-\prod_{n=1}^{3}\sin b_n=0$. But $2b_3=f_3(2a_3)=f_3(\pi/2)=\pi/2$, modulo 2π, so that $\sin b_3=\cos b_3=\pm 1/\sqrt{2}$. Therefore $\cos(b_1+b_2)=0$, whence $f_1(c)+f_2(\pi-c)=2f(a_1)+2f(a_2)=2(b_1+b_2)=\pi$, modulo 2π, by (12). But $f_2(c)+f_2(\pi-c)=f_2(\pi)=2f_2(\pi/2)=\pi$, which, together with the above, gives $f_1(c)=f_2(c)$. Similarly this equals $f_3(c)$, hence (13).

(14) If $f(a)=0$, modulo 2π, then $a=0$, modulo 2π.

Since $F(r)$ is a continuous function of r, it follows that $f(a)$ is a continuous function of a, modulo 2π. Thus, for every $\epsilon > 0$, there exists a $\delta > 0$, such that $|f(a)-\pi| = |f(a)-f(\pi)| \leq \epsilon$, modulo 2π, whenever $|a-\pi| \leq \delta$ modulo 2π.

Let us assume that $f(a)=0$. Unless a is of the form $a=2\pi p/q$, p integral, q odd, we can find an integer k such that $|ka-\pi| \leq \delta$, modulo 2π. (If a is a rational multiple of π, this is trivial: otherwise we refer to Kronecker's theorem.) Hence, by continuity, $|f(ka)-\pi| \leq \epsilon$. This contradicts the assumption that $f(ka)=kf(a)=0$. Therefore $f(a)=0$ implies that $a=2\pi p/q$, p integral, q odd.

Without loss of generality we may assume that p and q are relatively prime, hence there exist integers p' and q' such that $p'p-q'q=1$. Then $f(2\pi/q)=f(p'a-2\pi q')=f(p'a)=p'f(a)=0$, mod 2π.

Suppose there is an increasing sequence of odd numbers $q_n=2p_n-1$, such that $f(2\pi/q_n)=0$ for $n=1,2,\ldots$. Then $2\pi p_n/q_n \to \pi$. Hence, by continuity, $f(\pi)=0$, which contradicts the fact that $f(\pi)=2f(\pi/2)=\pi$. Thus there exists a largest q for which $f(2\pi/q)=0$; and it is an easy matter to show that, whenever $f(a)=0$, then $a=2\pi p/q$, modulo 2π.

We shall assume that $q\neq 1$. Let a_1 be any irrational multiple of π, then $f(4a_1)\neq 0$ mod 2π. Let $a_2=\pi/q$, then $f(2a_2)=0$ mod 2π. Let $\cot a_3=\tan a_1 \tan a_2$, then $\prod_{n=1}^{3}\cos a_n-\prod_{n=1}^{3}\sin a_n=0$, so that $r=\prod_{n=1}^{3}(\cos a_n+i_n\sin a_n)$ is a vector. But then, by (9), $F(r)$ is also

a vector, whence $\prod_{n=1}^{3}\cos f(a_n)-\prod_{n=1}^{3}\sin f(a_n)=0$. Since $4f(a_1)=f(4a_1)\neq 0$, therefore $\cos f(a_1)\neq 0$. Since $2f(a_2)=f(2a_2)=0$, therefore $\sin f(a_2)=0$. It follows that $\cos f(a_3)=0$, so that $f(\pi-2a_3)=\pi-2f(a_3)=0 \bmod 2\pi$. Hence $\pi-2a_3$ is an integral multiple of $2\pi/q$, so that $a_3=\pi/2-\pi p/q$. Thus

$$\tan a_1=\cot a_3/\tan a_2=\tan(\pi p/q)/\tan(\pi/q) \;.$$

If $q\neq 1$, this assumes at most denumerably many values, contrary to the fact that a_1 is an arbitrary irrational multiple of π. Thus $q=1$.

Hence $f(a)=0$ implies that $a=2\pi p$, p integral. This proves (14).

(15) If $g(b)=\tan f(\tan^{-1}b)$, b being any complex number, then g is a single valued function.

For let $b=\tan a$, so that $\tan^{-1}b=a+n\pi$, n integral. Then $2f(a+n\pi)=f(2a+2n\pi)=f(2a)+f(2n\pi)=2f(a)$ modulo 2π, and therefore $g(b)=\tan f(a+n\pi)=\tan f(a)$.

(16) If $g(b)=0$, then $b=0$.

For suppose $g(b)=0$. Then $f(2a)=2f(a)=0 \bmod 2\pi$, and therefore, by (14), $2a=0 \bmod 2\pi$. Hence $b=\tan a=0$.

(17) For finite b, $g(b)$ is finite.

For suppose $g(b)=\tan f(a)$ is not finite. Then $f(2a)=2f(a)=\pi=f(\pi)$, modulo 2π, so that $f(2a-\pi)=0 \bmod 2\pi$, and therefore, by (14), $2a=\pi$, modulo 2π. Hence $b=\tan a$ would not be finite either, contrary to hypothesis.

(18) $g(b_1 b_2)=g(b_1)g(b_2)$.

Since $g(0)=\tan f(0)=\tan 0=0$, this is trivially true if either b_1 or b_2 vanishes. Hence we may assume that $b_1, b_2 \neq 0$. Let $b_1=\tan a_1$, $b_2=\tan a_2$, $b_1 b_2=\cot a_3$. Then $\prod_{n=1}^{3} \sin a_n = \prod_{n=1}^{3} \cos a_n$. But, by (9), if r is a vector, so is F(r), hence

$$\prod_{n=1}^{3} \sin f(a_n) = \prod_{n=1}^{3} \cos f(a_n) \; . \qquad \ldots \qquad \text{(i)}$$

Since $b_n=\tan a_n \neq 0$ $(n=1,2)$, therefore $2a_n \neq 0$, mod 2π. In view of (14), $2f(a_n)=f(2a_n)\neq 0$, modulo 2π, whence $\cos f(a_n)\neq 0$ $(n=1,2)$. Similarly $\sin f(a_3)\neq 0$, whence, by (i), $\tan f(a_1)\tan f(a_2)=\cot f(a_3)$. Thus $g(b_1 b_2)=g(\cot a_3)=g(\tan((\pi/2)-a_3))=\tan f((\pi/2)-a_3)=$ $\tan((\pi/2)-f(a_3))=\cot f(a_3)=\tan f(a_1)\tan f(a_2)=g(b_1)g(b_2)$, as was to be shown.

(19) $g(b_1+b_2)=(g(b_1)+g(b_2))g(1-b_1 b_2)/(1-g(b_1 b_2))$.

For $g((b_1+b_2)/(1-b_1 b_2)) = g(\tan(a_1+a_2))$

$$= \tan f(a_1+a_2) =(g(b_1)+g(b_2))/(1-g(b_1)g(b_2$$

Multiply both sides of this equation by $g(1-b_1 b_2)$, then (19) follows in virtue of (18).

(20) $g(1)=1$. For by (16) there exists a number b such that $g(b)\neq 0$. But, by (18), $g(b)=g(b)g(1)$, whence (20).

(21) $g(-1)=-1$. For, by (18) and (19), we have

$$g(2)g(i)=g(2i)=g(i+i)=2g(i)g(1-i^2)/(1-g(i^2))$$
$$=2g(i)g(2)/(1-g(-1)) \; .$$

In virtue of (16), we may cancel, therefore $1-g(-1)=2$, whence (21).

(22) There exists a constant k, such that $g(b+c)=k(g(b)+g(c))$, whenever $b,c\neq 0$.

Put $d^2=bc$, so that $d\neq 0$, and therefore $g(d)\neq 0$, by (16). Hence, by (18) and (19), $g(1-d^2)/(1-g(d^2))=g(2d)/(2g(d))=g(2)/2=k$, say. Again by (19)

$$g(b+c)=(g(b)+g(c))g(1-d^2)/(1-g(d^2))=k(g(b)+g(c)) ,$$

as was to be shown.

(23) $g(b+c) = g(b)+g(c)$.

If either b or c vanishes, this is trivially true. Otherwise we must show that k in (22) is 1. Let $b,c\neq 0$. We can find a number $d\neq 0$, such that neither $b+d$ nor $d+c$ will vanish. Then by (22)

$$g((b+d)+c)=k^2g(b)+k^2g(d)+kg(c) ,$$

$$g(b+(d+c))=kg(b)+k^2g(d)+k^2g(c) .$$

The associative law of addition therefore gives

$$k(k-1)(g(b)-g(c))=0 .$$

If we take for instance $b=1$, $c=-1$, it follows that $k(k-1)=0$. By (16), $k=g(2)/2\neq 0$, hence $k=1$. This completes the proof of (23).

(24) Either $g(b)=b$, or $g(b)=b^C$ identically.

It is well known (and can easily be proved) that these are the only complex functions, not identically zero, which satisfy the functional equations: $g(bc)=g(b)g(c)$ and $g(b+c)=g(b)+g(c)$, proved in (18) and (23) respectively.

(25) Either $f(a)=a$, or $f(a)=a^C$ modulo 2π.

For $\tan f(a/2)=g(\tan(a/2))=\tan(a/2)$ or $=\tan(a^C/2)$, by (24). Hence $f(a)=2f(a/2)=a$ or $=a^C$, modulo 2π.

(26) If $F(-1)=-1$, then $F(r)=uru^T$ for all biquaternions r of norm 1, or $F(r)=ur^C u^T$, u being a constant biquaternion of norm 1.

It follows from (5), (11), and (13) that

$$r = \prod_{n=1}^{3}(\cos a_n + i_n \sin a_n) \ , \quad F(r) = \prod_{n=1}^{3}(\cos f(a_n) + j_n \sin f(a_n)) \ .$$

Hence by (25) $F(r) = \prod_{n=1}^{3}(\cos a_n^{(C)} + j_n \sin a_n^{(C)})$.

Upon multiplying this out, we have, in virtue of (10) :

$$r = \sum_{n=0}^{3} i_n r_n \ , \ F(r) = \sum_{n=0}^{3} j_n r_n^{(C)} \ ,$$

where $i_0=j_0=1$. Corresponding to $y=\sum_{n=0}^{3} y_n i_n$, let us put $G(y)=\sum_{n=0}^{3} y_n j_n$. The function G is thus defined for all biquaternions y. Obviously $G(y-y')=G(y)-G(y')$, and $NG(y)=N(y)$. Hence $G(y)$ is an orthogonal transformation, and, by theorem 1, $G(y)=py^{(T)}q$, where $N(p)=N(q)=1$. Since $F(r)=G(r^{(C)})$, it follows that $G(yy')=G(y)G(y')$, whenever $N(y)=N(y')=1$. In this case therefore $\pm p(yy')^{(T)}q=py^{(T)}qpy'^{(T)}q$, i.e. $\pm(yy')^{(T)}=y^{(T)} \ qpy'^{(T)}$. If we put x=1, this shows that $qp=\pm 1$. Hence $q=\pm p^T$, so that $G(y)=py^{(T)}p^T$, and $(yy')^{(T)}=y^{(T)}y'^{(T)}$, whenever y and y' have norm 1. However $(i_1 i_2)^T \neq i_1^T i_2^T$, hence $G(y)=pyp^T$, and therefore $F(r)=G(r^{(C)})=py^{(C)}p^T$, as was to be shown.

This completes case B. In virtue of (3), (8), and (26), we have now proved the lemma.

SUMMARY.

A biquaternion is of the form $a=\sum_{n=0}^{3} a_n i_n$, where $i_0=1$, $i_n^2=-1$ $(n=1,2,3)$, $i_1 i_2=i_3=-i_2 i_1$ etc., the a_n being complex numbers. By the conjugate a^C of a we mean the complex conjugate, whereas the transpose is $a^T=a_0-a_1 i_1-a_2 i_2-a_3 i_3$. The norm of a is $N(a)=aa^T$. a is Hermitian whenever $a^C=a^T$.

Biquaternion space B requires (a) a set of points Q, (b) a complex distance function $J(Q,Q')$, (c) a one to one correspondence S with the set of biquaternions $y=S(Q)$ such that $J(Q,Q')^2=N(y-y')$. S is called a special coordinate system, and if S' is another, $T=S'S^{-1}$ is called a special coordinate transformation. It is always of the form $T(y)=py^{(T)}q+b$, where p,q, and b are biquaternions, the first two of norm 1. In particular, $T(y)=pyq$ is called restricted.

Minkowski space M is defined in the same way except that we require a one to one correspondence $S: P\leftrightarrow x$ between all points P of M and all Hermitian biquaternions $x=S(P)$. Any special coordinate transformation T of M is of the form $T(x)=\pm rx^{(T)}r^{CT}+a$, where r is a biquaternion of norm 1, and a is a Hermitian biquaternion. In particular, $T(x)=rxr^{CT}$ is called restricted.

If T and T' are restricted coordinate transformations of M and B respectively, the correspondence $T \to T'$ is called a transformation scheme whenever $(T_1T_2)'=T_1'T_2'$. It follows that $T(x)=rxr^{CT}$ induces $T'(y)=P(r)yQ(r^{CT})$, where $P(rr')=\pm P(r)P(r')$, a similar equation holding for Q. If P and Q can be taken to be continuous functions, we also call the transformation scheme continuous.

If $T \to T'$ is a continuous transformation scheme from M to B, then $T(x)=rxr^{CT}$ induces $T'(y)=pyq$, where $p=1$ or $ur^{(C)}u^T$, and $q=1$ or $vr^{(C)T}v^T$, u and v being biquaternions of norm 1. This transformation scheme is denoted by $(p;q)$. It is called principal if $u=v=1$. For all practical purposes only principal continuous transformation schemes need be considered.

Let S be obtained from a given coordinate system of M by means of a restricted coordinate transformation. We define a vectorfield as a correspondence $S \to F_S$, where F_S is a function which assigns a biquaternion $F_S(P)$ to each point P of M, such that $F_{TS}=T'F_S$, $T \to T'$ being a continuous transformation scheme from M to B. If $S \to F_S$ is a vectorfield with transformation scheme $(p;q)$ then $F_{TS}(P)=pF_S(P)q$.

The conjugate and transpose of a vectorfield with transformation scheme (p;q) give vectorfields with transformation schemes $(p^C;q^C)$ and $(q^T;p^T)$ respectively. Two vectorfields with identical transformation schemes can be added to give a vectorfield. The product of two vectorfields with transformation schemes (p;s) and $(s^T;q)$ is a vectorfield with transformation scheme (p;q). Corresponding to the principal transformation schemes (1;1) ; $(r;r^{CT})$, $(r^C;r^T)$; $(r;r^T)$, $(r^C;r^{CT})$; $(r;1)$,$(r^C;1)$,$(1;r^T)$, and $(1;r^{CT})$ we have nine principal vectorfields , namely one invarient, two four-vectors, two six-vectors, and four wave-vectors.

Roughly speaking, an operator biquaternion is a biquaternion whose components are linear operators in the usual sense. We say that $S \to A_S$ is an operator vectorfield if $S \to (F_S A_S G_S)$ is a vectorfield for all vectorfields $S \to F_S$ and $S \to G_S$ of given transformation schemes. Unless $N(F_S A_S G_S)(P)=0$ identically, we can write $A_{TS}=pA_S q$, (p;q) being a continuous transformation scheme. With the exception of the rule for multiplication, the above results for vectorfields still hold for operator vectorfields.

BIBLIOGRAPHY.

A. W. CONWAY : Quaternions and Quantum Mechanics.
Pontificia Academia Scientiarum.
Acta Vol.XII-N.23, 1948.

J. DIEUDONNE : Sur les Groupes Classiques.
Publications de l'Institut de Mathématique de l'Université de Strasbourg. Paris, Hermann, 1948.

L. SILBERSTEIN : Theory of Relativity.
London, Macmillan, 1924.

B.L. Van der WAERDEN : Die Gruppentheoretische Methode in der Quantenmechanik.
Grundlehren. Berlin, Springer, 1932.

THE IMMERSIBILITY OF A SEMIGROUP INTO A GROUP.

By

J. LAMBEK

Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy. April, 1950

TABLE OF CONTENTS.

ACKNOWLEDGEMENTS.

I wish to thank my good friend and teacher, Professor Hans Zassenhaus of McGill University for his encouragement and criticism.

I am also indebted to Professor H.S.M.Coxeter of the University of Toronto for his kind and valuable suggestions.

J.L.

Montreal, April 1950.

1. Introduction.

A semigroup is a set of elements which is closed under an associative operation, usually called multiplication. When can a semigroup be embedded in a group, i.e. under what condition is it isomorphic to a subset of a group? A necessary condition for immersibility is clearly the so called cancellation law:

(C) If ax=ay or xb=yb, then x=y.

It is well known that a finite semigroup with cancellation law is a group, also that an Abelian semigroup (one in which multiplication is commutative) can be embedded in a group if and only if the cancellation law holds. In general however the cancellation law is not sufficient for immersibility, as was shown by A. Malcev in 1936. He also discovered necessary and sufficient conditions, which he maintained were too complicated for publication.

It is in fact not difficult to find such conditions. They state that certain systems of equations are not independent. This means that if all but one of the equations are given, the remaining equation can be deduced. However, as soon as we wish to give verbal utterance to these conditions, it becomes desirable to label the equations and the variables contained in them. This is where things begin to get involved. I have tried to overcome this complication by using parts of a polyhedron, rather than natural numbers, for labelling the equations and variables of any such system. In defining the term polyhedron I shall follow the book on topology by Seifert and Threlfall.

A <u>face</u> (abstract polygon) is a division of the unit circle into two or more arcs, called <u>sides</u>, by an equal number of points, which we will term <u>angles</u>. An abstract polyhedron is a system of F faces, containing together 2E sides, such that every side is mapped topologically on exactly one other. We may assume that the midpoint of one side is always mapped into the midpoint of another. A pair of sides thus mapped into each other is called an <u>edge</u>. Hence every edge has two sides. We may speak about the <u>midpoint</u> of an edge, which divides the edge into two <u>half-edges</u>. A set of angles which correspond to one another under the mapping is called a <u>vertex</u>. Every edge has two vertices. To every edge there belong four angles, which may be classified by pairs in two different ways: angles at the same vertex of the edge, and angles on the same side of the edge. Every half-edge has one vertex, two sides and two angles. The polyhedron is called <u>Eulerian</u> if the total number of vertices is V such that V+F-E=2. This is also a necessary and sufficient condition for the surface defined by the polyhedron to be homeomorphic (i.e. topologically equivalent) to the sphere. Throughout the present paper we shall always mean abstract Eulerian polyhedron when we say "polyhedron".

Given a semigroup H, we shall understand by <u>polyhedral condition</u> the following statement:

(P) If the elements of H are assigned to all sides and angles of any Eulerian polyhedron, such that to each half-edge there corresponds an equation xa=yb, where x and y have been assigned

to the sides, a and b to the corresponding angles of the half-edge, then these 2E equations are interdependent, i.e. any one of them can be derived from the totality of all others. (See figure 1.)

The application of this condition to two polyhedra which are topologically equivalent will of course give the same result. We shall prove that (P) is a necessary and sufficient condition for immersibility of a semigroup H with cancellation law into a group. This will establish the following

THEOREM: A semigroup can be embedded in a group if and only if the cancellation law (C) and the polyhedral condition (P) are satisfied.

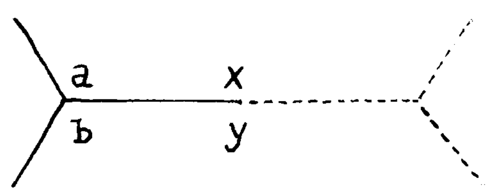


Fig 1

2. Necessity of polyhedral condition.

Let the semigroup H be contained in a group G, so that elements of H possess inverses in G. Assign elements of H to the sides and angles of a given polyhedron. Assume that the equations belonging to all but one half-edge are true, the remaining equation is to be deduced.

An oriented triangulation of the polyhedron is obtained as follows: Directed radii are drawn from the center of each face to the angles of the face. (It will be remembered that faces are circles, and that by angles we understand points on the circumference.) Directed radii are drawn from the midpoints of all sides to the center of the face. Each half-edge of the original polyhedron is given a direction from the midpoint of the edge towards the vertex. The half-edges thus oriented as well as the directed radii will be the oriented edges of the triangulation

The equation xa=yb, corresponding to any half-edge of the polyhedron, can be replaced by the two equations xa=p and yb=p, corresponding to two triangles of the triangulation. The variables occuring in these equations are assigned to the edges of the triangulation: namely x and y to the radii from the midpoint, a and b to the radii towards the vertex, and p to the half-edge itself. (See figure 2.)

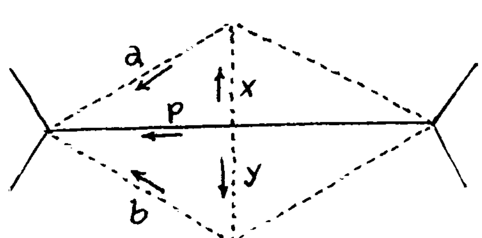


Fig 2

If the equation xa=yb was to be inferred, we may now add yb=p to the given equations, and leave only xa=p to be deduced.

Thus, corresponding to each oriented triangle, we have an equation; for instance $xa=p$ and $yb=p$ correspond to two of the four triangles on figure 2. Of these 4E equations all but one are given, and one is to be derived.

Consider any path made up entirely of edges of the triangulation. Corresponding to such a path we form a product in the following way: If the element x of H has been assigned to the n-th edge of the path, then the n-th term of the product is x or x^{-1}, depending on whether this edge has been traversed in the right or in the wrong direction. For instance, there are six different closed paths by means of which the perimeter of the upper left triangle of figure 2 can be once described. Correspondingly we may obtain one of the six products:

(i) xap^{-1}, $ap^{-1}x$, $p^{-1}xa$; $pa^{-1}x^{-1}$, $a^{-1}x^{-1}p$, $x^{-1}pa^{-1}$.

As long as $xa=p$, each of these products has the value 1; and conversely, if any one of the six products (i) is 1, then $xa=p$.

Consider now a closed path consisting of edges of the triangulation, and surrounding only triangles for which the corresponding equations are given. We prove by induction that the product corresponding to this path will be unity. We have shown above that this is indeed so, if the path surrounds only one triangle. Otherwise we may decompose the closed path into two paths Q and R traversed in succession, such that there will exist a path P, lying entirely inside the closed path, and joining the endpoint of Q to the endpoint of R. Let P' be the path P traversed in opposite direction. If f(P) denotes the product associated with P,

then $f(P)f(P')=1$. Hence $f(Q)f(R)=f(Q)f(P)f(P')f(R)=1$, since $f(Q)f(P)=f(P')f(R)=1$, by induction hypothesis.

Suppose now the upper left triangle of figure 2 is the one for which the corresponding equation is to be derived. Since the surface defined by our polyhedron is homeomorphic to the sphere, the perimeter of this triangle divides the whole surface into two simply connected regions. Let us describe a closed path along this perimeter, and call the outside of the triangle the inside of the path. Corresponding to this path we obtain one of the six products (i), whose value will be unity, in virtue of the above. Hence $xa=p$, as was to be deduced. We have thus shown the necessity of the polyhedral condition.

3. Ratios.

In preparation for the sufficiency proof of the polyhedral condition, we shall consider a semigroup H satisfying (C) and (P). It is our intention to introduce ratios, in more or less the same way as is usually done when H is the set of natural numbers.

Let a and b be any two elements of H. We shall designate by a/b the set of pairs of elements x,y such that xa=yb. If a/b is not the empty set we shall call it a ratio. Similarly we define I, (a), $(a)^{-1}$ as sets of pairs x,y such that x=y , xa=y , x= ya, respectively. With the help of the cancellation law, we can easily show that they are also ratios. In fact

(1) ... I=t/t , (a)=at/t , $(a)^{-1}$=t/at ,

where t is an arbitrary element of H. We also note that

(2) ... (a)=(b) if and only if a=b.

When can we say that two ratios are equal? We will prove:

(3) ... a/b = c/d if and only if there exist x and y belonging to H such that xa=yb and xc=yd . (It is assumed here that a/b is in fact a ratio, and therefore not empty.)

The necessity of this condition follows directly from the definition of ratios. To prove its sufficiency, let us assume that the condition of the theorem holds, and also that ua=vb. Let us now apply the polyhedral condition to a simple polyhedron consisting of two edges, two faces, and two vertices (figure 3).

The equations corresponding to three of the half-edges of this polyhedron are true by assumption, consequently the fourth must hold, viz. uc=vd . This argument works both ways, hence ua=vb if and only if uc=vd, and therefore a/b = c/d , by definition.

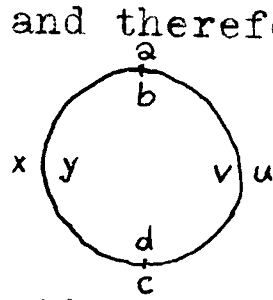


Fig 3

We define multiplication of ratios as follows:

(4) ... (a/d)(d/b) = a/b .

Thus two ratios may be multiplied to give another ratio, provided they can be written in the form a/d and d/b respectively such that a/b is a ratio. Is the product of two ratios unique, if it exists? To answer this in the affirmative we must show:

(5) ... If a/d = a'/d' and d/b = d'/b' , then a/b = a'/b' .

We may assume that xa=yd, xa'=yd' , zd=wb, zd'=wb' , ua'=vb' , and wish to prove that ua=vb. The result follows immediately if we consider the polyhedron consisting of two vertices, three face and three edges. (See figure 4.)

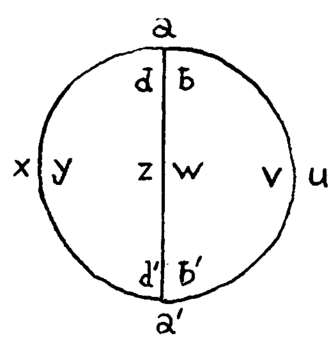


Fig 4

I is the unit element under multiplication. For

$$(a/b)I = (a/b)(b/b) = a/b$$

by (1) and (4). The same applies to multiplication by I on the left. We note the existence of inverses; thus, by (1) and (4),

$$(a/b)(b/a) = a/a = I \ .$$

In particular we see from (1) that $(a)^{-1}$ is the inverse of (a), as was anticipated by our notation. We find that

$$(a)\ (b)=(abt/bt)(bt/t) = abt/t = (ab)\ .$$

Hence, in view of (2), the correspondence $a \rightarrow (a)$ maps H isomorphically on a subset of the set of ratios.

We have embedded H in the set of ratios. The latter has all properties of a group, except that it is not closed under multiplication, and associativity has not yet been shown to hold. We shall embed it in a larger set, in which multiplication is always defined and associative. It may be worth noting that, if H is an Abelian semigroup, the ratios do form a group already.

4. Associativity.

The set of ratios almost form a group, except that it is not closed under multiplication, so that also the associative law, as usually stated, has no meaning. With some care however it is possible to enunciate an associative law even here. If we can bracket a sequence of ratios in such a way that they can be multiplied out to give a single ratio, then this "product" shall be unique. To be more precise: we shall say that a finite sequence of ratios $(\ldots,a/b,b/c,\ldots)$ contracts into the sequence $(\ldots,a/c,\ldots)$. If a sequence of ratios reduces to a single ratio by iterated contraction, we will call this ratio its product. The associative law then states:

(6) ... If a sequence of ratios has a product, then it is unique.

To prove (6), consider a sequence $S(0)$ of $n+1$ ratios. This contracts to $S(\pm 1)$, consisting of n ratios, which in turn contracts to $S(\pm 2)$, and so on, until we obtain a single ratio $S(\pm n)$. If the plus sign is chosen, we have one method of iterated contraction; if the minus sign is chosen, we have another. We must show that $S(+n)=S(-n)$.

If k is an integer between 0 and n, we write $i=\pm k$, and note that $S(i)$ has $n+1\mp k$ places or terms. Represent the j-th place of $S(i)$ by the point (i,j) in the Cartesian plane. If $k\neq 0$, all but two terms of $S(\pm k\mp 1)$ reappear in $S(\pm k)$: Join the corresponding points by straight lines. But two consecutive terms, say a_i/c_i and c_i/b_i are contracted into a_i/b_i :

Join the two former points by straight lines to the latter, which will be called a vertex. Also join the two points (or vertices) $(\pm n,1)$ to the point at infinity, along the line $y=1$. A broken line joining two vertices, even if it passes through the point at infinity, will be called an edge. There are three edges meeting at every vertex. The simply connected regions into which the edges divide the plane will be called faces. Since the plane can be mapped on a sphere by an inverse stereographic projection, we obtain a concrete representation of an Eulerian polyhedron. We may also verify independently that $V=2n$, $F=n+2$, and $E=3n$, so that $V+F-E=2$. A simple case, for which $n=2$, is illustrated by figure 5.

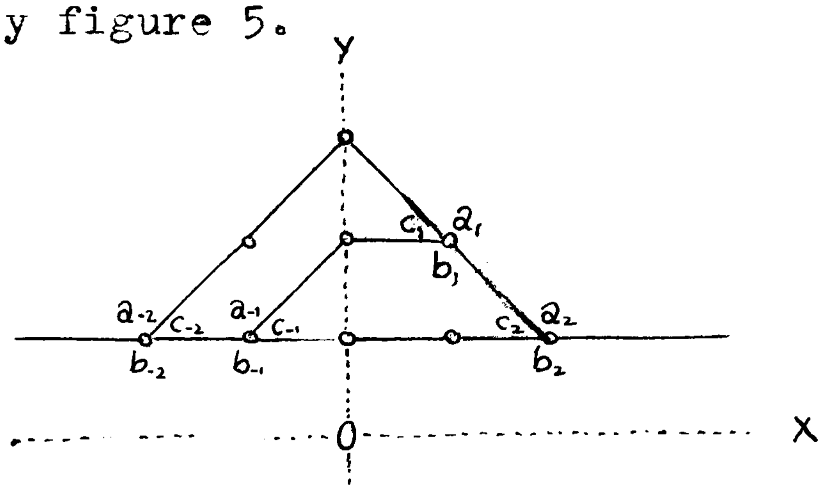


Fig 5

Consider the vertex corresponding to the contraction of $S(\pm k\mp 1)=(\ldots,a_i/c_i,c_i/b_i,\ldots)$ into $S(\pm k)=(\ldots,a_i/b_i,\ldots)$. To the three angles formed at this vertex we assign a_i, c_i, and b_i, in this order, going from top to bottom, as shown in figure 5. By our construction, if a and b have been assigned to the upper, respectively lower angle at one end of any finite edge, and if this edge passes through any integral lattice point (i,j), then the j-th term of $S(i)$ is a/b. But in the same way, we find

that this term is c/d, where c and d correspond to the two angles at the other end of the edge. Thus, for any finite edge, we have a "proportion" a/b=c/d. We will prove that such a proportion also holds for the edge passing through the point at infinity.

In view of (3), the above proportion may be replaced by the two equations xa=yb and xc=yd. Here x and y may be conveniently assigned to the two sides of the edge (see figure 6), and the two equations may be said to correspond to the two half-edges. Consider now the edge joining (n,1) and (-n,1) through the point at infinity. An appropriate transformation will bring the point at infinity into the finite part of the plane. Since a_{-n}/b_{-n} is a ratio, by definition, there exist elements u and v of H such that $ua_{-n}=vb_{-n}$. We may assign u and v to the upper, respectively lower side of the edge depicted in figure 7, and the given equation will correspond to the left half of this edge. With the help of the polyhedral condition, we deduce the remaining equation $ua_{n}=vb_{n}$. It follows from (3) that $a_{-n}/b_{-n}=a_{n}/b_{n}$, i.e. S(-n)=S(+n). This concludes the proof of the associative law.

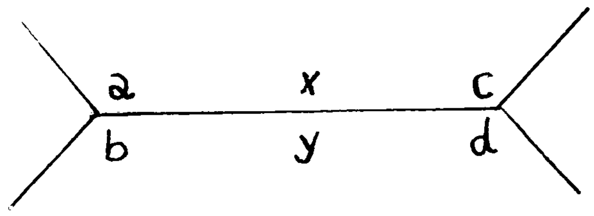


Fig 6

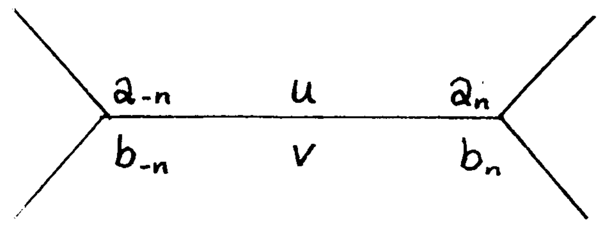


Fig 7

5. Sufficiency of polyhedral condition.

Two finite sequences of ratios, U and V, will be called <u>similar</u>, if there is a sequence W from which both can be obtained by repeated contraction. We will prove the following result :

(7) ... If both U and V reduce to S by iterated contraction, then they are similar.

First, suppose U contracts to S, so that U=(P,a/c,c/b,Q) and S=(P,a/b,Q), where P and Q may be empty sequences. Since V reduces to S, we may put V=(X,Y,Z), where X, Y, and Z reduce to P, a/b, and Q respectively, by iterated contraction. It is easily seen that W=(X,Y,b/c,c/b,Z) can be reduced to both U and V by repeated contraction, so that U and V are similar. Hence (7) holds when U reduces to S in one step.

Next, suppose U reduces to S in n steps, n>1. Then U contracts to U' which reduces to S in n-1 steps. By induction hypothesis, there exists a sequence W' which reduces to U' and V by iterated contraction. Since U reduces to U' in only one step, by the above, there exists a sequence W which reduces to both U and W' and therefore V. Hence U and V are similar, as was to be proved. (See figure 8 for an illustration of the second part of this proof.)

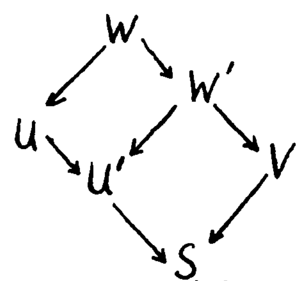
We are now in a position to show that similarity of sequences of ratios is an equivalence relation in the usual sense.

(8) ... Similarity is symmetric, reflexive, and transitive.

Symmetry is obvious. Reflexivity follows from the fact that (S,I) contracts to S. To prove transitivity, let us assume that R is

similar to S and S is similar to T. Hence there exists a sequence U which reduces to both R and S, and a sequence V which reduces to both S and T (see figure 9). Since both U and V reduce to S, by (7) they can both be obtained from a sequence W by repeated contraction. Now W reduces to R via U and to T via V, hence R is similar to T, as was to be proved.

Fig. 8



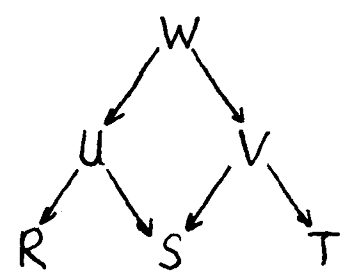


Fig 9

In this connection we may also state:

(9) ... If S is similar to S' and T is similar to T', then (S,T) is similar to (S',T').

For, by repeated contraction, we obtain S and S' from U, T and T' from V, hence (S,T) and (S',T') from (U,V).

A ratio may be regarded as a sequence of ratios with one term. When are two ratios similar?

(10) ... Two ratios are similar if and only if they are equal.

Because of reflexivity we know that equal ratios are similar. Conversely, let two ratios be similar. By definition, this means that both can be derived from the same sequence W by repeated contraction. From (6) we deduce that they are equal.

Let us denote by S^x the class of all sequences which are similar to S, so that $S^x=T^x$ if and only if S and T are similar, in view of (8). We define multiplication of similarity classes as follows:

(11) ... $S^x T^x = (S,T)^x$.

From (9) we know that the product thus defined is unique. Associativity becomes apparent if we write both $(S^\times T^\times)U^\times$ and $S^\times(T^\times U^\times)$ as $(S,T,U)^\times$. The unit element under multiplication is $I^\times$, since both (S,I) and (I,S) are similar to S; for both are obtained by contracting (I,S,I). If T contains the reciprocals of the ratios of S in reverse order, then both (S,T) and (T,S) reduce to and are therefore similar to I; thus $T^\times$ may be regarded as the inverse of $S^\times$ under multiplication. We have thus proved that the similarity classes form a group G, with multiplication defined by (11).

The correspondence $a/b \to (a/b)^\times$ is a homomorphic mapping of the set of ratios on a subset of G. For, by (11) and (4),

$$(a/b)^\times \ (b/c)^\times = (a/b, b/c)^\times = (a/c)^\times = ((a/b)(b/c))^\times .$$

More than this, the mapping is isomorphic. For if $(a/b)^\times = (c/d)^\times$ then a/b and c/d are similar, hence $a/b = c/d$, by (10). The correspondence $a/b \to (a/b)^\times$ therefore embeds the set of ratios in G. But the correspondence $a \to (a)$ embeds the semigroup H in the set of ratios, as we have shown in section 3. Hence the correspondence $a \to (a)^\times$ embeds H in G. This establishes the sufficiency of the polyhedral condition.

6. Application to Abelian semigroups.

Let H be an Abelian semigroup with cancellation law. Although it is not difficult to show directly that H is immersible in a group (namely the set of ratios), we shall test the usefulness of the polyhedral condition, by showing independently that the latter holds in H.

Let elements of H be assigned to all angles and sides of any given polyhedron. Of the equations corresponding to the half-edges we will assume that all but one hold, and we wish to deduce the remaining equation. As in the necessity proof of the polyhedral condition, we introduce a triangulation and replace each equation xa=yb by two equations xa=p and yb=p corresponding to triangles. We may assume then that all but two of these latter equations are given.

If the triangulation is regarded as a network, each vertex is seen to be an even node, i.e. has an even number of edges meeting at it. Hence there is an Euler line, i.e. the entire network can be traced in one single closed curve which does not pass through any point twice. Since the polyhedron was homeomorphic to the sphere, this Euler line will divide the triangles into two classes, so that triangles with a common edge do not belong to the same class. We will write the equation corresponding to a triangle of the first class as xa=p and the equation corresponding to a triangle of the second class as p=yb, making a careful distinction between the two sides of each equation. Now multiply all 4E-2 given equations together, after their sides have been thus arranged. It will be observed that the four variables belonging to the half-edge

whose equation is to be deduced occur once in the product equation. All other variables occur twice, once on each side of the product equation, and may therefore be cancelled, by (C). There results an equation containing four variables, and it is easily seen that this is in fact the equation we wished to deduce. Hence the polyhedral condition is satisfied, as was to be proved.

7. Summary and Conclusions.

A semigroup is a group without inverses under multiplication. It is proved that a semigroup H can be embedded in (or is isomorphic to a subset of) a group if and only if the cancellation law (C) and the polyhedral condition (P) hold. Here (C) means that ax=bx or xa=xb should always imply that a=b, for any elements a, b, x of H. (P) states: If elements of H are assigned to the angles and sides (i.e. sides of edges) of any Eulerian polyhedron (a polyhedron whose surface is homeomorphic to the sphere), so that to each half-edge there corresponds an equation xa=yb, where x and y have been assigned to the two sides, a and b to the corresponding two angles of the half-edge, then these equations are interdependent, i.e. any one of them can be deduced from the totality of all others.

We have spoken above of <u>the</u> polyhedral condition. We might equally well say that there are as many different conditions as there are topologically inequivalent polyhedra, namely an enumerable infinitude. Even worse, the number of equations entering any such condition, being twice the number of edges of the polyhedron, is unbounded as we vary the latter. The number of conditions can be reduced to some extent: it suffices to consider only such polyhedra as can be cut up into two trees; moreover we may restrict the number of edges meeting at any vertex to three. However it remains an open question whether a finite number of conditions may not do. I conjecture that this question is to be answered in the negative.

BIBLIOGRAPHY.

MALCEV, A. (Moscow): "On the immersion of an algebraic ring into a field". Received 1936, Mathematische Annalen, vol. 113, p. 687.

SEIFERT, H. & THRELFALL, W : Lehrbuch der Topologie. New York, Chelsea Publishing Company, 1947.