Sgríbhinní Institiúid Árd-Léinn Bhaile Átha Cliath
Sraith A, Uimh. 15

Communications of the Dublin Institute for
Advanced Studies. Series A, No. 15

THE PETROV CLASSIFICATION OF GRAVITATIONAL FIELDS

BY

J. L. SYNGE

INSTITIÚID ÁRD-LÉINN BHAILE ÁTHA CLIATH
64-65 CEARNÓG MHUIRFEANN, BAILE ÁTHA CLIATH

DUBLIN INSTITUTE FOR ADVANCED STUDIES
64-65 MERRION SQUARE, DUBLIN
1964

THE PETROV CLASSIFICATION OF GRAVITATIONAL FIELDS

by

J. L. SYNGE

TABLE OF CONTENTS

E R R A T A

p. 14: Theorem II. After "non-zero vectors" insert ", no two of which are collinear," .

p. 19: Line 2 after equation (7.31). Delete "are".

p. 29: Equations (8.22). For " $- i\, a_\rho\, \underline{P}_\rho$ " read " $i\, a_\rho\, \underline{P}_\rho$ " .

p. 30: Equation (8.27). For " $- a_\rho\, a_4\, P_\rho^{\,2}$ " read " $a_\rho\, a_4\, P_\rho^{\,2}$ " .

Equation (8.31). For " $A_1^{\,2} + A_2^{\,2} + A_3^{\,2}$ " read " $A_1 + A_2 + A_3$ " .

p. 31: Equation (8.34). For " + " read " − " (twice).

Equation (8.35). For " $i\, a_4$ " read " $- i\, a_4$ " .

Equations (8.36), second line. For " $\underline{P}_\rho^{\,2}$ " read " $P_\rho^{\,2}$ " , and for " + " read " − " .

Line 3 from end. For " (7.52) " read " (7.72)" .

p. 40: Line 1. For " (11.11) " read " (11.12) " .

p. 42: Line after equation (12.6). For " (11.19) " read " (11.20) " .

p. 43: Last line of equation (12.14), expression " $\frac{1}{2}\, h\, (- 1 + h^{-2} - \frac{1}{4}\, b^2 h^{-4})$". For " $\frac{1}{2}$ " read " $\frac{i}{2}$ " .

p. 45: Class 2a, reduced matrix. For " λ"' " read " λ" ".

The Petrov classification of gravitational fields

J. L. SYNGE

1. Introduction.

The Petrov classification of gravitational fields is rightly regarded as of great interest in modern relativity theory, particularly in connection with gravitational radiation, and the relevant literature is extensive. For a comprehensive survey, see Pirani (1962) [the references are at the end of this paper], and for a recent statement by the originator, see Petrov (1962).

There are several ways of approaching the problem of classification. They may be divided broadly into algebraic and geometric, and of the latter one method is due to Géhéniau (1957) and another to Debever (1964). Of these the method of Géhéniau appeals to me most, and I thought it might be of interest to go through the whole business in some detail without any claim to originality. Physicists may be expected to have some geometrical appreciation of Minkowskian space-time, but it is unlikely that they are familiar with complex Euclidean 3-space. Since this is central in the method of Géhéniau, I have gone into this geometry rather fully.

For many valuable discussions I am much indebted to Drs. F. Öktem and H. Yeh, Scholars at the Dublin Institute for Advanced Studies.

2. Notation and preliminaties.

In general the notation is that of my books (Synge, 1956, 1960). The ranges for literal indices are as follows:

small Latin:	1, 2, 3, 4
small Greek:	1, 2, 3
capital Latin:	1, 2, 3, 4, 5, 6

The summation convention is used, except where the contrary is indicated. The signature of space-time is +2 for real coordinates. At the beginning of the paper the formulae are valid for real coordinates or for coordinates of which the fourth is pure imaginary, but later, as an essential simplifying device, imaginary time will be used.

As minor variations in notation exist in the literature, it is necessary for clarity to write out certain well known formulae as follows.

For metric tensor g_{ab}, the Riemann tensor is

$$R_{abcd} = \tfrac{1}{2}(g_{ad,bc} + g_{bc,ad} - g_{ac,bd} - g_{bd,ac}) + \\ + g^{mn}([ad,m][bc,n] - [ac,m][bd,n]) . \qquad (2.1)$$

The Ricci tensor and the scalar curvature are

$$R_{bc} = g^{ad} R_{abcd} , \qquad R = g^{bc} R_{bc} , \qquad (2.2)$$

and the Einstein tensor is

$$G_{bc} = R_{bc} - \tfrac{1}{2} g_{bc} R . \qquad (2.3)$$

The Weyl tensor is

$$W_{abcd} = R_{abcd} - \tfrac{1}{2}(g_{ad} R_{bc} + g_{bc} R_{ad} - g_{ac} R_{bd} - g_{bd} R_{ac}) \\ + \tfrac{1}{6}(g_{ad} g_{bc} - g_{ac} g_{bd}) R . \qquad (2.4)$$

We note the symmetries

$$W_{abcd} = -W_{bacd} = W_{cdab} , \qquad W_{abcd} + W_{acdb} + W_{adbc} = 0 . \qquad (2.5)$$

The Weyl tensor satisfies the ten identities

$$g^{ad} W_{abcd} = 0 . \qquad (2.6)$$

These identities play an essential role in the classification problem. In fact, for the classification of W_{abcd} all essential information is contained in (2.5) and (2.6); the structure displayed in (2.4) is not needed.

The Riemann tensor R_{abcd} also has the symmetries shown in (2.5), but it does not in general satisfy (2.6). We make no attempt to classify

R_{abcd} in general, but only in vacuo where we have $R_{bc} = 0$ and consequently equations of the form (2.6) with W_{abcd} replaced by R_{abcd}. The two problems, (i) classification of W_{abcd} in general, and (ii) classification of R_{abcd} in vacuo, are mathematically the same. We shall carry out the classification of W_{abcd} in general. It is perhaps important to emphasise that we thus make a classification throughout a whole universe, even inside matter (or in an electromagnetic field) with energy tensor given by

$$-\kappa T_{ab} = R_{ab} - \tfrac{1}{2} g_{ab} R , \qquad (\kappa = 8\pi) , \tag{2.7}$$

and not merely in vacuo, where of course $W_{abcd} = R_{abcd}$.

3. What does classification mean?

The general theory of relativity is unique among physical theories in respect to the arbitrariness of coordinate systems. A single universe may be described by a metric tensor field $g_{ab}(x)$ for one coordinate system and by $\bar{g}_{ab}(\bar{x})$ for another coordinate system, the first ten functions being quite different from the second ten functions, but related to them by

$$\bar{g}_{ab}(\bar{x}) = g_{cd}(x) X^c_a X^d_b , \tag{3.1}$$

the X's indicating partial derivatives.

But suppose we have what appear to be two different universes, U and $\bar{U}$, and the question arises whether they are in fact the same universe. This will be the case if there exists a transformation $x^a \to \bar{x}^a$ between the coordinates of U and those of $\bar{U}$ such that the relation (3.1) connects the metric g_{ab} of U and the metric $\bar{g}_{ab}$ of $\bar{U}$. It becomes then a question of the integrability of the ten non-linear first-order partial differential equations (3.1).

We shall not attempt to discuss this very difficult question. Instead, we ask a less general question: Could an event E of U be the same event as an event $\bar{E}$ of $\bar{U}$? Again the answer turns on (3.1), but this is no longer a set of partial differential equations but a set of algebraic equations, the X's being constants. The question is then whether the sixteen X's can be chosen so that (3.1) is satisfied, $\bar{g}_{ab}$ and g_{cd} being given numbers (not fields). The answer is well known: it is possible to do this provided the signatures of $\bar{g}_{ab}$ and g_{cd} are the same (for signature + 2, each can be reduced to diag (1, 1, 1, − 1)). We may sum up by saying that all events are identical (this is sometimes expressed by saying that space-time is elementarily flat).

But this identity of events is lost when we examine their neighbourhoods, meaning thereby not merely the metric tensor but also its partial derivatives. If we include only first-order derivatives, the identity of events remains (we can make all the first derivatives vanish). If we go to second-order derivatives, the identity of events no longer holds. In fact, it is possible to classify events in terms of the Weyl tensor in the following sense.

Suppose we are given a Weyl tensor W_{abcd} at an event E and a Weyl tensor $\bar{W}_{abcd}$ at an event $\bar{E}$. If we are dealing with two different coordinate systems in a single universe, then

$$\bar{W}_{abcd} = W_{pqrs} X^p_{\bar{a}} X^q_{\bar{b}} X^r_{\bar{c}} X^s_{\bar{d}} . \qquad (3.2)$$

We ask: given $\bar{W}_{abcd}$ and W_{pqrs}, can we find sixteen numbers, the X's, so that (3.2) holds? The answer is that we cannot do this in general.

We can do it only if $\bar{W}_{abcd}$ and W_{pqrs} are of the same <u>class</u>. That is in fact what class means in this connection.

4. <u>The eigenvalue problem</u>.

Define

$$g_{abcd} = - g_{ad}\, g_{bc} + g_{ac}\, g_{bd} . \tag{4.1}$$

These quantities have the symmetries shown in (2.5) - replace W by g. Let F^{ab} be any skew-symmetric contravariant tensor $(F^{ba} = - F^{ab})$. Consider the equations

$$W_{abcd}\, F^{cd} = \lambda\ g_{abcd}\, F^{cd} , \tag{4.2}$$

where λ is a scalar. There are six equations here, and the consistency condition is the satisfaction of a 6×6 determinantal equation, yielding six eigenvalues λ, independent of the coordinates used. We shall be much concerned with these eigenvalues. They need not all be distinct. If complex, they occur in conjugate pairs.

It is easy to see that these eigenvalues remain unchanged if we replace the quantities in (4.2) by their transforms under a transformation with Jacobian matrix $X^p_{\bar{a}}$. Thus in order that two Weyl tensors should be of the same class, it is <u>necessary</u> that the eigenvalues of W_{abcd} relative to g_{abcd} should be the same as those of $\bar{W}_{abcd}$ relative to $\bar{g}_{abcd}$. But this condition is by no means <u>sufficient</u>; if it were, the problem of classification would be rather trivial.

At this point we introduce the Petrov notation, correlating number-pairs in the range 1, 2, 3, 4 to single numbers in the range 1, 2, 3, 4, 5, 6 according to the following scheme:

$$(23) \leftrightarrow 1, \quad (31) \leftrightarrow 2, \quad (12) \leftrightarrow 3, \quad (14) \leftrightarrow 4, \quad (24) \leftrightarrow 5, \quad (34) \leftrightarrow 6. \tag{4.3}$$

Thus we write $W_{2323} = W_{11}$, $W_{2331} = W_{12}$, ... $W_{3434} = W_{66}$.

In fact, with capital Latin suffixes in the range 1, ..., 6, the whole set of components W_{abcd} may be exhibited, except for reversals of sign, as W_{AB}, where

$$W_{BA} = W_{AB} . \tag{4.4}$$

We have in fact a symmetric 6 x 6 matrix, with 21 elements. But these are reduced to 20 by the last of (2.5), which gives

$$W_{14} + W_{25} + W_{36} = 0 . \tag{4.5}$$

In this notation, the six components F^{ab} of a skew-symmetric tensor may be written F^A, and the six equations (4.2) may be written

$$W_{AB} F^B = \lambda g_{AB} F^B ; \tag{4.6}$$

the characteristic equation reads

$$\det (W_{AB} - \lambda g_{AB}) = 0 . \tag{4.7}$$

The contents of the present section are true without recourse to (2.6); what has been said here would be true if W were replaced by R throughout.

5. Restriction to imaginary time.

For notational reasons we shall now use coordinates x_a for which x_4 is a pure imaginary ($x_4 = it$). Further, at the event we are concerned with, we use coordinates for which $g_{ab} = \delta_{ab}$.

the Kronecker delta. This leaves the coordinates free only to within a Lorentz transformation

$$x'_a = L_{ab}\, x_b \,, \qquad L_{ac}\, L_{bc} = \delta_{ab} = L_{ca}\, L_{cb} \,. \tag{5.1}$$

Simplifications result. First, (4.1) becomes, in the Petrov notation,

$$g_{AB} = \mathrm{diag}\,(1,\ 1,\ 1,\ 1,\ 1,\ 1)\,. \tag{5.2}$$

Secondly, (2.6) reads $W_{abca} = 0$, and hence, in the Petrov notation,

$$\begin{aligned}
W_{22} + W_{33} + W_{44} &= 0\,,\\
W_{33} + W_{11} + W_{55} &= 0\,,\\
W_{11} + W_{22} + W_{66} &= 0\,,\\
W_{44} + W_{55} + W_{66} &= 0\,,\\
W_{12} = W_{45}\,, \quad W_{23} = W_{56}\,, &\quad W_{31} = W_{64}\,,\\
W_{15} = W_{24}\,, \quad W_{26} = W_{35}\,, &\quad W_{34} = W_{16}\,.
\end{aligned} \tag{5.3}$$

We had 20 components W_{AB} without the above relations; they reduce the number to 10, and it is easy to see that the 6×6 matrix W may be exhibited in the form

$$W = \begin{pmatrix} M & N \\ N & M \end{pmatrix} \tag{5.4}$$

where M and N are symmetric 3×3 matrices with zero traces:

$$M = \begin{pmatrix} W_{11} & W_{12} & W_{13} \\ W_{21} & W_{22} & W_{23} \\ W_{31} & W_{32} & W_{33} \end{pmatrix} = \begin{pmatrix} W_{44} & W_{45} & W_{46} \\ W_{54} & W_{55} & W_{56} \\ W_{64} & W_{65} & W_{66} \end{pmatrix}, \quad N = \begin{pmatrix} W_{14} & W_{15} & W_{16} \\ W_{24} & W_{25} & W_{26} \\ W_{34} & W_{35} & W_{36} \end{pmatrix}, \tag{5.5}$$

Denoting the transpose by a tilde, and trace by tr, we have

$$M = \tilde{M}, \qquad N = \tilde{N}, \qquad \mathrm{tr}\, M = 0, \qquad \mathrm{tr}\, N = 0. \tag{5.6}$$

For tr $N = 0$, we invoke (4.5); tr $M = 0$ follows from the first four of (5.3).

Putting F for the column matrix $(F_1, F_2, F_3, F_4, F_5, F_6)$ — since $g_{ab} = \delta_{ab}$ there is no difference between contravariant and covariant — the equations (4.6) now read

$$WF = \lambda F, \tag{5.7}$$

and the characteristic equation (4.7) reads

$$\det (W - \lambda I) = 0, \tag{5.8}$$

where I is the unit 6×6 matrix.

6. Passage to complex E_3.

When x_4 is a pure imaginary, and we so assume henceforth, there is a parity rule to the effect that a tensor component is real if it contains an even number of 4's and imaginary if it contains an odd number of 4's. Thus F_{23} $(= F_1)$ is real, and F_{14} $(= F_4)$ a pure imaginary. This parity rule, be it noted, is for indices in the range 1, 2, 3, 4, and not for those in the range 1, ..., 6. But it can be modified so as to apply to the latter, in this form: a quantity is real if it contains 4, 5, 6 evenly and imaginary if it contains them oddly. Thus M is real and N imaginary.

Let us write the column matrices

$$G = \text{col}\,(F_1, F_2, F_3)\,, \qquad H = \text{col}\,(F_4, F_5, F_6)\,, \tag{6.1}$$

so that G is real and H imaginary. By virtue of (5.4), we can write (5.7) in the form

$$\begin{pmatrix} M & N \\ N & M \end{pmatrix} \begin{pmatrix} G \\ H \end{pmatrix} = \lambda \begin{pmatrix} G \\ H \end{pmatrix} \tag{6.2}$$

This is equivalent to

$$MG + NH = \lambda G\,, \qquad NG + MH = \lambda H\,. \tag{6.3}$$

Define

$$K = M + N\,, \qquad J = G + H\,, \tag{6.4}$$

so that K is a symmetric complex 3 x 3 matrix with zero trace, and J a complex 3 x 1 matrix. Addition of the two equations (6.3) gives

$$KJ = \lambda J\,. \tag{6.5}$$

Thus for λ we have the cubic equation

$$\det\,(K - \lambda I) = 0\,, \tag{6.6}$$

where I is now the unit 3 x 3 matrix.

Since a study of the eigen problem (5.7) is essential in the problem of classification, it will be realised that what has here been achieved is a major step. From an equation of degree 6 in (5.8), we have passed to one of degree 3 in (6.6). But what has become of three of the roots of (5.8)? The answer is that they are the complex conjugates of the three roots of (6.6), as may be seen by subtracting (6.4) instead of adding. Note that since K is complex, the cubic (6.6) does not necessarily possess a real root nor a complex conjugate pair of roots.

The 3-vector J is interesting; its components are, in space-time notation,

$$J_1 = F_{23} + F_{14} , \quad J_2 = F_{31} + F_{24} , \quad J_3 = F_{12} + F_{34} . \tag{6.7}$$

We can give the argument a physical complexion by introducing "electromagnetic vectors" $\underline{E}$ and $\underline{H}$ by

$$\begin{aligned} H_1 &= F_{23} , \quad H_2 = F_{31} , \quad H_3 = F_{12} , \\ E_1 &= i F_{14} , \quad E_2 = i F_{24} , \quad E_3 = i F_{34} . \end{aligned} \tag{6.8}$$

Then we have

$$\underline{J} = \underline{H} - i \underline{E} = - i (\underline{E} + i \underline{H}) , \tag{6.9}$$

a complex vector familiar in electromagnetic theory. (Note: in this work we indicate 3-vectors by an underline, corresponding to the printed heavy type.)

As we have now reached a critical point in the problem of classification of the Weyl tensor, let us sum up what has been done:

As a preliminary to classification, we have reduced the 6-dimensional problem for eigenvalues and eigen 6-vectors exhibited in (4.2) or (4.6) to a 3-dimensional problem exhibited in (6.5). We have passed from 4-dimensional space-time to 3-dimensional complex space.

As far as guidance from intuition is concerned, we might think that we had made a great simplification. We no longer have inhibitions arising from 4-dimensionality and indefinite metric. The reduction to 3-dimensions is good. But we have to pay a price. It is a complex 3-space, and many familiar intuitions may not be used. But some can, and that is a great help.

No reason has so far been given for considering this 3-space to be Euclidean. This will be discussed in the next section.

7. Complex Euclidean 3-space E_3 .

Complex Euclidean 3-space E_3 is a natural formal generalisation of the real Euclidean 3-space with which we are all familiar. A vector now has three complex components instead of three real ones. This means that the complex E_3 is actually a space of 6 dimensions, since there are 6 real numbers needed to define a vector (or, equivalently, a point). In view of this doubling of dimensionality, it is remarkable that we find so many of the familiar facts of real geometry in three dimensions reproduced in the complex E_3 .

But we must of course be careful not to allow ourselves to carry over elements of real geometry without careful scrutiny. Everything must be redefined, and, if some of these redefinitions are understood below, it is because the reader may be assumed to fill in the gaps for himself.

A word of warning to the reader accustomed to linear vector spaces with a scalar product of the Hilbert type. The complex E_3 is something quite different in that the conjugate complex plays no part at all in it. In Hilbert space a vector has a positive norm, the square root of the scalar product of the vector by itself. In complex E_3 there is indeed a scalar product (defined below), but no norm; the scalar product is complex, and to take its square root would be foolish. There is no "distance between points" in complex E_3 .

Scalar products and types of vectors.

A vector in E_3 is an ordered triad of complex numbers. We shall use several different notations for vectors to suit various occasions:

(i) The vector is represented explicitly by its components (V_1, V_2, V_3) or briefly by V_σ where $\sigma = 1, 2, 3$.

(ii) The vector is represented by the symbol $\underline{V}$ (the underline in typescript corresponds to the heavy type used in printing).

(iii) The vector is represented by the symbol V to indicate a column matrix with elements (V_1, V_2, V_3), or by its transpose $\tilde{V}$ (a row matrix with the same elements).

The sum of two vectors is given by adding the components in the usual way.

The scalar product of two vectors is defined to be the complex number $V_1 V_1' + V_2 V_2' + V_3 V_3'$: it may be written in any one of the following forms:

$$V_\sigma V_\sigma' = V_\sigma' V_\sigma = \underline{V} \cdot \underline{V}' = \underline{V}' \cdot \underline{V} = \tilde{V} V . \qquad (7.1)$$

Two vectors are orthogonal if their scalar product vanishes.

Putting $V' = V$, we get the self scalar product:

$$V_\sigma V_\sigma = \underline{V} \cdot \underline{V} = \tilde{V} V . \qquad (7.2)$$

We now define three types of vectors:

(a) Zero vector 0 or $\underline{0}$ for which all the components vanish $(V = 0)$.

(b) Unit vector for which the self scalar product is unity ($\tilde{V} V = 1$; equivalently $V_\sigma V_\sigma = 1$, or $\underline{V} \cdot \underline{V} = 1$).

(c) Null (or isotropic) vector for which the self scalar product is zero, but which is not the zero vector ($\tilde{V} V = 0$; equivalently $V_\sigma V_\sigma = 0$, or $\underline{V} \cdot \underline{V} = 0$).

For example, $(1, 0, 0)$ is a unit vector; $(1, 0, i)$ is a null vector.

Vector product and orthonormal triad.

The vector product of two vectors $\underline{A}$ and $\underline{B}$ is defined in the usual way:

$$\underline{A} \times \underline{B} = (A_2B_3 - A_3B_2 ,\ A_3B_1 - A_1B_3 ,\ A_1B_2 - A_2B_1) = -\underline{B} \times \underline{A} . \quad (7.3)$$

It is immediately seen that $\underline{A} \times \underline{B}$ is orthogonal to $\underline{A}$ and to $\underline{B}$.

An ordered set of three unit vectors, mutually orthogonal, is called an orthonormal triad. If $\underline{A}$, $\underline{B}$, $\underline{C}$ are the three vectors, the defining conditions are

$$\underline{A} \cdot \underline{A} = \underline{B} \cdot \underline{B} = \underline{C} \cdot \underline{C} = 1 , \qquad \underline{B} \cdot \underline{C} = \underline{C} \cdot \underline{A} = \underline{A} \cdot \underline{B} = 0 , \quad (7.4)$$

or equivalently in matrix notation

$$\tilde{A} A = \tilde{B} B = \tilde{C} C = 1 , \qquad \tilde{B} C = \tilde{C} A = \tilde{A} B = 0 . \quad (7.5)$$

For an orthonormal triad, consider the determinant

$$D = \begin{vmatrix} A_1 & A_2 & A_3 \\ B_1 & B_2 & B_3 \\ C_1 & C_2 & C_3 \end{vmatrix} . \quad (7.6)$$

Squaring it, we find $D^2 = 1$, and so $D = \epsilon = \pm 1$. If $D = 1$, we say that the triad is proper; if $D = -1$, improper. It is then easy to show that

$$\underline{A} = \epsilon \underline{B} \times \underline{C} , \qquad \underline{B} = \epsilon \underline{C} \times \underline{A} , \qquad \underline{C} = \epsilon \underline{A} \times \underline{B} , \quad (7.7)$$

where $\epsilon = +1$ or -1 according as the triad is proper or improper.

Collinearity and coplanarity.

Two vectors $\underline{A}$ and $\underline{B}$ are said to be collinear if there exist complex numbers a and b (not both zero) such that

$$a \underline{A} + b \underline{B} = \underline{0} . \quad (7.8)$$

If not collinear, the vectors are said to be linearly independent.

Three vectors, $\underline{A}$, $\underline{B}$, $\underline{C}$, are said to be coplanar if there exist complex numbers a, b, c (not all zero) such that

$$a\,\underline{A} + b\,\underline{B} + c\,\underline{C} = \underline{0}\,. \qquad (7.9)$$

If not coplanar, the three vectors are said to be linearly independent.

If $\underline{A}$ is a given vector and a an arbitrary complex number, then the set of vectors (or points) $a\,\underline{A}$ may be called a line of $\underline{A}$. But it must be realised that, since a is complex, this is not a one-dimensional continuum like an ordinary line. Likewise, for given $\underline{A}$, $\underline{B}$ and arbitrary a, b, the expression $a\,\underline{A} + b\,\underline{B}$ defines the plane containing $\underline{A}$ and $\underline{B}$.

Two theorems about null vectors.

All that has been said above is somewhat dull, for it is almost entirely the transcription into complex E_3 of known properties of real 3-space. Novelty appears when we consider null vectors, and we shall now prove two theorems:

Theorem I: Two non-collinear (or linearly independent) null vectors cannot be orthogonal; equivalently, if two null vectors are orthogonal, they are collinear.

Theorem II: An orthogonal triad of non-zero vectors cannot contain a null vector as a member.

To prove Theorem I, we assume that $\underline{A}$ and $\underline{B}$ are orthogonal null vectors, so that

$$\underline{A}\cdot\underline{A} = \underline{B}\cdot\underline{B} = \underline{A}\cdot\underline{B} = 0\,. \qquad (7.10)$$

From the last of these, we have

$$(A_1 B_1 + A_2 B_2)^2 = A_3^2 B_3^2\,, \qquad (7.11)$$

and hence by the other two of (7.10)

$$(A_1 B_1 + A_2 B_2)^2 = (A_1^2 + A_2^2)(B_1^2 + B_2^2)\,,$$

which may be written

$$(A_1 B_2 - A_2 B_1)^2 = 0 . \tag{7.12}$$

There are two variations on this argument, and we get

$$A_2 B_3 - A_3 B_2 = A_3 B_1 - A_1 B_3 = A_1 B_2 - A_2 B_1 = 0 . \tag{7.13}$$

But these imply the collinearity of the pair of orthogonal null vectors, and so Theorem I is proved.

To prove Theorem II, we assume that $\underline{A}$, $\underline{B}$, $\underline{C}$ is an orthogonal triad, $\underline{A}$ being null. Thus

$$\underline{A} \cdot \underline{A} = 0 , \quad \underline{A} \cdot \underline{B} = 0 , \quad \underline{B} \cdot \underline{C} = 0 , \quad \underline{C} \cdot \underline{A} = 0 . \tag{7.14}$$

From the second and fourth of these equations, we obtain

$$\underline{A} = \theta \, \underline{B} \times \underline{C} , \tag{7.15}$$

where θ is some number, not zero. Hence, by the first of (7.14),

$$(\underline{B} \times \underline{C}) \cdot (\underline{B} \times \underline{C}) = 0 . \tag{7.16}$$

But identically for any two vectors

$$(\underline{B} \times \underline{C}) \cdot (\underline{B} \times \underline{C}) = (\underline{B} \cdot \underline{B}) (\underline{C} \cdot \underline{C}) - (\underline{B} \cdot \underline{C})^2 , \tag{7.17}$$

and so this last expression vanishes. But $\underline{B} \cdot \underline{C} = 0$, and therefore either $\underline{B}$ or $\underline{C}$ is null. But this is impossible by Theorem I, and so Theorem II is proved.

<u>Orthogonal matrices and orthogonal transformations.</u>

A 3×3 matrix T with complex elements is said to be <u>orthogonal</u> if its transpose is its inverse:

$$\tilde{T} = T^{-1} \quad \text{or} \quad \tilde{T} T = I , \tag{7.18}$$

where I is the unit 3×3 matrix; we have also $T \tilde{T} = I$.

If such a matrix is applied to transform the vectors of E_3 , vectors A , B transform into A' , B' according to the rule

$$A' = T A , \qquad B' = T B . \tag{7.19}$$

Then the scalar product transforms according to

$$\tilde{A}' B' = \tilde{A} \tilde{T} T B = \tilde{A} B ; \tag{7.20}$$

in fact, the scalar product is conserved under the transformation. Hence it follows that an orthonormal triad transforms into an orthonormal triad, and a null vector transforms into a null vector. We call the transformation orthogonal (and this is also the reason for so naming the matrix) because it conserves orthogonality, but of course it does more than that.

The orthogonal transformation is proper or improper according as T is proper or improper.

The whole set of orthogonal transformations form a group, and so does the set of proper orthogonal transformations, but the set of improper orthogonal transformations does not form a group since the result of applying two improper transformations is a proper transformation.

It has been remarked above that when an orthonormal triad is submitted to an orthogonal transformation it changes into an orthonormal triad. But (and this is a possible source of confusion) an orthogonal matrix is itself an orthonormal triad, or may be so regarded: For any matrix can be written in the form

$$T = \begin{pmatrix} A_1 & A_2 & A_3 \\ B_1 & B_2 & B_3 \\ C_1 & C_2 & C_3 \end{pmatrix} \tag{7.21}$$

or, if we regard the elements in each row as the components of a vector,

$$T = \begin{pmatrix} \underline{A} \\ \underline{B} \\ \underline{C} \end{pmatrix} . \qquad (7.22)$$

Multiplication of this by its transpose gives

$$T \tilde{T} = \begin{pmatrix} \underline{A} \cdot \underline{A} & \underline{A} \cdot \underline{B} & \underline{A} \cdot \underline{C} \\ \underline{B} \cdot \underline{A} & \underline{B} \cdot \underline{B} & \underline{B} \cdot \underline{C} \\ \underline{C} \cdot \underline{A} & \underline{C} \cdot \underline{B} & \underline{C} \cdot \underline{C} \end{pmatrix} \qquad (7.23)$$

So far all is general. But now suppose that T is an orthogonal matrix. This means that the matrix (7.23) is the unit matrix, and so the three vectors $\underline{A}$, $\underline{B}$, $\underline{C}$ form an orthonormal triad. Conversely, if the three vectors form an orthonormal triad, then T, as in (7.21), is an orthogonal matrix. In fact, all orthogonal matrices may be exhibited as in (7.21) or (7.22) in terms of an orthonormal triad.

We follow convention in speaking of orthogonal matrices and orthogonal transformations; but it would be better in each case to use the word orthonormal.

Transformation of an eigenvalue problem.

Consider an equation as in (6.5),

$$K\,J = \lambda\,J . \qquad (7.24)$$

Let T be any orthogonal matrix. Then (7.24) implies

$$T\,K\,T^{-1}\,T\,J = \lambda\,T\,J , \qquad (7.25)$$

and if we define

$$J' = T\,J , \qquad K' = T\,K\,T^{-1} , \qquad (7.26)$$

we get

$$K'\,J' = \lambda\,J' . \qquad (7.27)$$

This, then, is how we transform an eigenvalue problem. The eigenvectors J transform according to (7.26) and the eigenvalues λ remain unchanged. We note that the matrix K transforms as in (7.26) and <u>not</u> according to $K' = TK$!

The purpose of such an orthogonal transformation is to simplify the problem, turning K into a simpler matrix K' . If the orthogonal matrix T is expressed as in (7.21) in terms of an orthonormal triad A, B, C, then, as is easily verified,

$$K' = T K T^{-1} = \begin{pmatrix} \tilde{A} K A & \tilde{A} K B & \tilde{A} K C \\ \tilde{B} K A & \tilde{B} K B & \tilde{B} K C \\ \tilde{C} K A & \tilde{C} K B & \tilde{C} K C \end{pmatrix} . \quad (7.28)$$

This formula will be essential later. Note that, if K is a symmetric matrix so that $\tilde{K} = K$ (as will be the case later), then K' is also symmetric. And it is easy to see that, if $\operatorname{tr} K = 0$, then $\operatorname{tr} K' = 0$.

<u>Canonical form of an orthonormal triad.</u>

In the totality of orthonormal triads, there are some of outstanding simplicity, namely those for which the three vectors are real. Among these is the orthonormal triad (1, 0, 0) , (0, 1, 0) , (0, 0, 1) , but that is not particularly interesting for our purposes because it is tied to the axes. The whole class of real orthonormal triads may be written $\underline{I}$, $\underline{J}$, $\underline{K}$ where these are real orthogonal unit vectors, so that

$$\underline{I}^2 = \underline{J}^2 = \underline{K}^2 = 1 , \qquad \underline{J} \cdot \underline{K} = \underline{K} \cdot \underline{I} = \underline{I} \cdot \underline{J} = 0 . \quad (7.29)$$

Here and henceforth we adopt the notation $\underline{A}^2 = \underline{A} \cdot \underline{A}$ as a convenience, but we shall not speak of the square roots of $\underline{A}^2$. We have

$$\underline{I} = \varepsilon\, \underline{J} \times \underline{K}\,, \qquad \underline{J} = \varepsilon\, \underline{K} \times \underline{I}\,, \qquad \underline{K} = \varepsilon\, \underline{I} \times \underline{J}\,, \tag{7.30}$$

where $\varepsilon = \pm 1$ according as the triad is proper or improper.

Any complex vector $\underline{A}$ may be resolved along the above real triad by a formula

$$\underline{A} = p\,\underline{I} + q\,\underline{J} + r\,\underline{K}\,, \tag{7.31}$$

and indeed a complex orthonormal triad may be so resolved, the expressions involving 9 complex coefficients are connected by 6 conditions of orthonormality, so that 3 are free. Our purpose now is to use these three complex degrees of freedom to simplify the 9 coefficients. Or we may say that, given an orthonormal triad $\underline{T}_\sigma$ (the subscript here being a distinguishing label and not a component index), we seek a real orthonormal triad $(\underline{I}\,, \underline{J}\,, \underline{K})$ so that the expressions of the type (7.31) are as simple as possible.

Query: Why not choose $(\underline{I}\,, \underline{J}\,, \underline{K})$ collinear with the vectors of the given complex triad? Answer: There exists in general no real vector collinear with a given complex vector.

Consider then a triad of vectors $\underline{T}_\sigma$, satisfying the orthonormality conditions

$$\underline{T}_\rho \cdot \underline{T}_\sigma = \delta_{\rho\sigma}\,. \tag{7.32}$$

From their unit character, none of these vectors can be purely imaginary. If they are all real, we have no problem. Assume then that $\underline{T}_1$ is complex, so that we can write it

$$\underline{T}_1 = \underline{P} + i\,\underline{Q}\,, \tag{7.33}$$

where $\underline{P}$ and $\underline{Q}$ are real vectors, with $\underline{Q}$ not the zero vector. Let P and Q be the magnitude of these real vectors. From $T_1^{\,2} = 1$ we obtain

$$P^2 - Q^2 = 1 , \qquad \underline{P} \cdot \underline{Q} = 0 . \tag{7.34}$$

Obviously $P \neq 0$, and we have assumed $Q \neq 0$. We can therefore divide by P or Q , and so we can define a pair of orthogonal real unit vectors

$$\underline{I} = \underline{P} / P , \qquad \underline{J} = \underline{Q} / Q , \tag{7.35}$$

and write

$$\underline{T}_1 = P \underline{I} + i Q \underline{J} . \tag{7.36}$$

Now define a complex vector

$$\underline{S} = - i Q \underline{I} + P \underline{J} . \tag{7.37}$$

We find

$$\underline{S}^2 = 1 , \qquad \underline{S} \cdot \underline{T}_1 = 0 , \tag{7.38}$$

so that $\underline{S}$ is a unit vector orthogonal to $\underline{T}_1$.

Having defined $\underline{I}$ and $\underline{J}$ as above, we complete a real orthonormal triad by taking

$$\underline{K} = \underline{I} \times \underline{J} . \tag{7.39}$$

It is clear that $\underline{T}_1$, $\underline{S}$, $\underline{K}$ form a complex orthonormal triad, and any vector can be resolved along them. Thus $\underline{T}_2$ and $\underline{T}_3$, being orthogonal to $\underline{T}_1$, may be written

$$\underline{T}_2 = a \underline{S} + b \underline{K} , \qquad \underline{T}_3 = c \underline{S} + d \underline{K} . \tag{7.40}$$

The orthonormality conditions (7.32) impose on these coefficients the conditions

$$a^2 + b^2 = 1 , \qquad c^2 + d^2 = 1 , \qquad ac + bd = 0 , \tag{7.41}$$

and so there exist complex numbers θ , θ' such that

$$\begin{aligned} a &= \cos\theta , & b &= \sin\theta , \\ c &= \sin\theta' , & d &= \cos\theta' , \\ \sin(\theta' + \theta) &= 0 . \end{aligned} \tag{7.42}$$

The triad $\underline{T}_\sigma$ may be proper or improper, and we include them both by writing

$$\underline{T}_1 = \varepsilon \underline{T}_2 \times \underline{T}_3, \qquad \varepsilon = \pm 1 . \tag{7.43}$$

Accordingly from (7.40)

$$\begin{aligned}\underline{T}_1 &= \varepsilon(ad - bc)\, \underline{S} \times \underline{K} \\ &= \varepsilon\,(ad - bc)\,(-i\,Q\,\underline{I} + P\,\underline{J}) \times (\underline{I} \times \underline{J}) \\ &= \varepsilon\,(ad - bc)\,(i\,Q\,\underline{J} + P\,\underline{I}),\end{aligned} \tag{7.44}$$

and so, by (7.36),

$$\cos(\theta' + \theta) = ad - bc = \varepsilon . \tag{7.45}$$

Thus

$$\text{for } \varepsilon = 1, \quad \theta' = -\theta; \qquad \text{for } \varepsilon = -1, \quad \theta' = \pi - \theta, \tag{7.46}$$

and so

$$\begin{aligned} a &= \cos\theta, & b &= \sin\theta, \\ c &= -\varepsilon\sin\theta, & d &= +\varepsilon\cos\theta . \end{aligned} \tag{7.47}$$

Accordingly the most general complex orthonormal triad may be expressed in the canonical form

$$\begin{aligned} \underline{T}_1 &= P\,\underline{I} + i\,Q\,\underline{J}, \quad \underline{T}_2 = \cos\theta\,\underline{S} + \sin\theta\,\underline{K}, \\ \underline{T}_3 &= -\varepsilon\sin\theta\,\underline{S} + \varepsilon\cos\theta\,\underline{K}, \end{aligned} \tag{7.48}$$

where $\underline{I}$, $\underline{J}$ are two unit orthogonal real vectors, P, Q any two real positive numbers satisfying $P^2 - Q^2 = 1$, $\underline{S} = -i\,Q\,\underline{I} + P\,\underline{J}$, $\underline{K} = \underline{I} \times \underline{J}$, the factor ε being $+1$ or -1 according as the triad $\underline{T}_\sigma$ is proper or improper.

Further formulae for an orthonormal triad in E_3.

If we express the T's in terms of real vectors, writing

$$\underline{T}_\rho = \underline{P}_\rho + i\,\underline{Q}_\rho \tag{7.49}$$

(we recall that the suffixes are not components-indices but labels distinguishing the several vectors, and there is at present no summation convention), in the set $\underline{P}_\rho$, $\underline{Q}_\rho$ there are 6 real vectors with 18 components. In (7.32) we have 12 real relations which read

$$\underline{P}_\rho \cdot \underline{P}_\sigma - \underline{Q}_\rho \cdot \underline{Q}_\sigma = \delta_{\rho\sigma} , \qquad \underline{P}_\rho \cdot \underline{Q}_\sigma + \underline{Q}_\rho \cdot \underline{P}_\sigma = 0 , \qquad (7.50)$$

so that there are 6 degrees of freedom in the orthonormal triad, the same number as in the Lorentz transformation.

The implications of (7.50) are interesting. To explore them, we shall use (7.48), writing

$$\cos\theta = \alpha + i\beta , \qquad \sin\theta = \gamma + i\delta , \qquad (7.51)$$

these four real numbers satisfying

$$\alpha^2 - \beta^2 + \gamma^2 - \delta^2 = 1 , \qquad \alpha\beta + \gamma\delta = 0 . \qquad (7.52)$$

We have then

$$\underline{P}_1 = P\,\underline{I} , \qquad \underline{Q}_1 = Q\,\underline{J} ,$$

$$\underline{P}_2 = \beta Q\,\underline{I} + \alpha P\,\underline{J} + \gamma\,\underline{K} , \qquad \underline{Q}_2 = -\alpha Q\,\underline{I} + \beta P\,\underline{J} + \delta\,\underline{K} ,$$

$$\varepsilon\,\underline{P}_3 = -\delta Q\,\underline{I} - \gamma P\,\underline{J} + \alpha\,\underline{K} , \qquad \varepsilon\,\underline{Q}_3 = \gamma Q\,\underline{I} - \delta P\,\underline{J} + \beta\,\underline{K} .$$

$$(7.53)$$

These formulae give the totality of orthonormal triads, with $\varepsilon = \pm 1$ (the orientation index), P and Q arbitrary real positive numbers subject to

$$P^2 - Q^2 = 1 , \qquad (7.54)$$

and α , β , γ , δ being four real numbers subject to (7.52). We may think of the 6 degrees of freedom distributed as follows: 3 in the real orthonormal pair $\underline{I}$, $\underline{J}$, and one each in P , γ and δ .

From (7.53) we have, with use of (7.54),

$$P_1^{\,2} = P^2 , \qquad P_2^{\,2} = (\alpha^2 + \beta^2) P^2 - \beta^2 + \gamma^2 ,$$

$$P_3^{\,2} = (\gamma^2 + \delta^2) P^2 + \alpha^2 - \delta^2 , \qquad (7.55)$$

$$\underline{P}_1 \cdot \underline{P}_2 = \beta P Q , \qquad \varepsilon\,\underline{P}_1 \cdot \underline{P}_3 = -\delta P Q , \qquad \varepsilon\,\underline{P}_2 \cdot \underline{P}_3 = -(\alpha\gamma + \beta\delta) Q^2$$

Then, defining the S's as follows, we have

$$S_1 = -1 - P_1^{\,2} + P_2^{\,2} + P_3^{\,2} = P^2 (\alpha^2 + \beta^2 + \gamma^2 + \delta^2 - 1) ,$$
$$S_2 = -1 + P_1^{\,2} - P_2^{\,2} + P_3^{\,2} = -Q^2 (\alpha^2 + \beta^2 - \gamma^2 - \delta^2 - 1) ,$$
$$S_3 = -1 + P_1^{\,2} + P_2^{\,2} - P_3^{\,2} = Q^2 (\alpha^2 + \beta^2 - \gamma^2 - \delta^2 + 1) . \qquad (7.56)$$

Thus

$$S_1 S_2 = -P^2 Q^2 [(\alpha^2 + \beta^2 - 1)^2 - (\gamma^2 + \delta^2)^2] ,$$
$$S_1 S_3 = P^2 Q^2 [(\alpha^2 + \beta^2)^2 - (\gamma^2 + \delta^2 - 1)^2] , \qquad (7.57)$$
$$S_2 S_3 = -Q^4 [(\alpha^2 + \beta^2 - \gamma^2 - \delta^2)^2 - 1] .$$

By (7.52)

$$\gamma^2 - \delta^2 = 1 - \alpha^2 + \beta^2 , \qquad \gamma\delta = -\alpha\beta , \qquad (7.58)$$

and so

$$(\gamma^2 + \delta^2)^2 = (1 - \alpha^2 + \beta^2)^2 + 4\alpha^2\beta^2 = (\alpha^2 + \beta^2 - 1)^2 + 4\beta^2 , \qquad (7.59)$$

and likewise

$$(\alpha^2 + \beta^2)^2 = (\gamma^2 + \delta^2 - 1)^2 + 4\delta^2 . \qquad (7.60)$$

Further

$$(\alpha^2 + \beta^2 - \gamma^2 - \delta^2)^2 - 1 = (\alpha^2 + \beta^2 - \gamma^2 - \delta^2)^2 - (\alpha^2 - \beta^2 + \gamma^2 - \delta^2)^2$$
$$= -4(\alpha\gamma + \beta\delta)^2 . \qquad (7.61)$$

Hence by (7.57)

$$S_1 S_2 = 4P^2 Q^2 \beta^2 = 4(\underline{P}_1 \cdot \underline{P}_2)^2 ,$$
$$S_1 S_3 = 4P^2 Q^2 \delta^2 = 4(\underline{P}_1 \cdot \underline{P}_3)^2 , \qquad (7.62)$$
$$S_2 S_3 = 4Q^4 (\alpha\gamma + \beta\delta)^2 = 4(\underline{P}_2 \cdot \underline{P}_3)^2 ,$$

and so, for $\rho \neq \sigma$,

$$S_\rho\, S_\sigma = 4\, (\underline{P}_\rho \cdot \underline{P}_\sigma)^2 \, . \tag{7.63}$$

We note also that

$$\begin{aligned} \underline{P}_1 \cdot \underline{Q}_2 &= -\underline{P}_2 \cdot \underline{Q}_1 = -\alpha\, P\, Q\,, \\ \varepsilon\, \underline{P}_1 \cdot \underline{Q}_3 &= -\varepsilon\, \underline{P}_3 \cdot \underline{Q}_1 = \gamma\, P\, Q\,, \\ \varepsilon\, \underline{P}_2 \cdot \underline{Q}_3 &= -\varepsilon\, \underline{P}_3 \cdot \underline{Q}_2 = (\beta\gamma - \alpha\delta)\, P^2 \,; \end{aligned} \tag{7.64}$$

$$\begin{aligned} \underline{P}_1 \times \underline{Q}_1 &= P\, Q\, \underline{K}\,, \\ \underline{P}_2 \times \underline{Q}_2 &= (\alpha\delta - \beta\gamma)\, P\, \underline{I} - (\alpha\gamma + \beta\delta)\, Q\, \underline{J} + (\alpha^2 + \beta^2)\, P\, Q\, \underline{K}\,, \\ \underline{P}_3 \times \underline{Q}_3 &= (\alpha\delta - \beta\gamma)\, P\, \underline{I} + (\alpha\gamma + \beta\delta)\, Q\, \underline{J} + (\gamma^2 + \delta^2)\, P\, Q\, \underline{K}\,; \end{aligned} \tag{7.65}$$

$$\begin{aligned} \underline{Q}_2 \cdot (\underline{P}_1 \times \underline{Q}_1) &= \delta\, P\, Q\,, \\ \varepsilon\, \underline{Q}_3 \cdot (\underline{P}_1 \times \underline{Q}_1) &= \beta\, P\, Q\,, \\ \underline{Q}_1 \cdot (\underline{P}_2 \times \underline{Q}_2) &= -(\alpha\gamma + \beta\delta)\, Q^2\,, \\ \varepsilon\, \underline{Q}_3 \cdot (\underline{P}_2 \times \underline{Q}_2) &= -\beta\, P\, Q\,, \\ \underline{Q}_1 \cdot (\underline{P}_3 \times \underline{Q}_3) &= (\alpha\gamma + \beta\delta)\, Q^2\,, \\ \underline{Q}_2 \cdot (\underline{P}_3 \times \underline{Q}_3) &= -\delta\, P\, Q\,. \end{aligned} \tag{7.66}$$

Thus there are only three distinct values in these mixed triple products. It is easy to see this otherwise. For we have

$$\varepsilon\, (\underline{P}_1 + i\, \underline{Q}_1) = (\underline{P}_2 + i\, \underline{Q}_2) \times (\underline{P}_3 + i\, \underline{Q}_3)\,, \tag{7.67}$$

and two similar equations (with the same ε in all) obtained by cyclic permutation. Hence

$$\varepsilon\, \underline{Q}_1 = \underline{P}_2 \times \underline{Q}_3 + \underline{Q}_2 \times \underline{P}_3\,, \tag{7.68}$$

and so

$$\varepsilon\, \underline{Q}_2 \cdot \underline{Q}_1 = \underline{Q}_2 \cdot (\underline{P}_2 \times \underline{Q}_3) = -\underline{Q}_3 \cdot (\underline{P}_2 \times \underline{Q}_2)\,. \tag{7.69}$$

But by cyclic permutation of (7.68),

$$\varepsilon\, \underline{Q}_2 = \underline{P}_3 \times \underline{Q}_1 + \underline{Q}_3 \times \underline{P}_1\,, \tag{7.70}$$

and so

$$\varepsilon \, \underline{Q}_1 \cdot \underline{Q}_2 = \underline{Q}_1 \cdot (\underline{Q}_3 \times \underline{P}_1) = \underline{Q}_3 \cdot (\underline{P}_1 \times \underline{Q}_1) \, . \qquad (7.71)$$

Hence

$$\underline{Q}_3 \cdot (\underline{P}_1 \times \underline{Q}_1) = - \underline{Q}_3 \cdot (\underline{P}_2 \times \underline{Q}_2) \, , \qquad (7.72)$$

and two similar equations obtained by cyclic permutation.

8. The connection between Lorentz transformations in space-time and orthogonal transformations in complex E_3 .

Throughout this paper I use imaginary time $(x_4 = it)$ except in the first four sections. Some people find imaginary time confusing, but in the present instance its use leads to great notational simplifications arising out of the fact that the metric tensor is simply the unit matrix. It may be noted that truly complex quantities enter naturally when we use imaginary time, and nowhere do we have occasion to bring in complex conjugates. Nevertheless certain cautions should be given.

When we use imaginary time, we are accustomed to check our work by applying a parity rule with respect to the suffix 4 , already mentioned in Section 6 : a component is real or imaginary according as the suffix 4 occurs an even or odd number of times. Thus F_{23} is real, F_{14} imaginary. When we define the dual by

$$F^*_{rs} = \tfrac{1}{2} i \, \varepsilon_{rsmn} F_{mn} \, , \qquad F_{rs} = - \tfrac{1}{2} i \, \varepsilon_{rsmn} F^*_{mn} \, , \quad (8.1)$$

we insert the factor i in order that the parity rule may be observed (F^*_{23} real, F^*_{14} imaginary) ; ε_{rsmn} is the usual real 4-index permutation symbol. To avoid confusion in interpreting what follows, it is essential to realise that we are going to violate the parity rule by introducing quantities which are complex.

Let F_{rs} be a skew-symmetric tensor ($F_{\alpha\beta}$ real, $F_{\alpha 4}$ imaginary). Define

$$A_{rs} = F_{rs} + \tfrac{1}{2}\varepsilon_{rsmn}F_{mn} . \qquad (8.2)$$

Note that there is no factor i here as in (8.1) : A_{rs} is a complex quantity with components

$$A_{23} = A_{14} = F_{23} + F_{14} , \quad A_{31} = A_{24} = F_{31} + F_{24} , \quad A_{12} = A_{34} = F_{12} + F_{34} . \qquad (8.3)$$

We have then

$$A_{ab} = \tfrac{1}{2}\varepsilon_{abcd}A_{cd} . \qquad (8.4)$$

We might say that A_{rs} is self-dual, but that would be confusing, since there is a factor i in (8.1) but not in (8.4). This distinction is essential.

The permutation symbol ε_{abcd} is a tensor with respect to proper Lorentz transformations, otherwise it changes sign. Any homogeneous Lorentz transformation may be written:

$$x'_a = L_{ab}x_b , \qquad (8.5)$$

where the matrix L satisfies

$$L\tilde{L} = 1 = \tilde{L}L , \quad L_{ac}L_{bc} = \delta_{ab} = L_{ca}L_{cb} ; \qquad (8.6)$$

it is proper if det L = 1 . We shall here take L to be proper.

Then A_{rs} is a tensor, transforming according to

$$A'_{rs} = L_{ra} L_{sb} A_{ab} \; . \tag{8.7}$$

Thus, with Greek suffixes in the range 1 , 2 , 3 , we have

$$A'_{\rho 4} = L_{\rho\alpha} L_{4\beta} A_{\alpha\beta} + (L_{\rho\gamma} L_{44} - L_{\rho 4} L_{4\gamma}) A_{\gamma 4} \; . \tag{8.8}$$

But by (8.4)

$$A_{\alpha\beta} = \epsilon_{\alpha\beta\gamma} A_{\gamma 4} \; , \tag{8.9}$$

where $\varepsilon_{\alpha\beta\gamma}$ is the usual 3-index permutation symbol, and so (8.8) gives

$$A'_{\rho 4} = T_{\rho\gamma} A_{\gamma 4} \; , \tag{8.10}$$

where

$$T_{\rho\gamma} = L_{\rho\alpha} L_{4\beta} \varepsilon_{\alpha\beta\gamma} + L_{\rho\gamma} L_{44} - L_{\rho 4} L_{4\gamma} \; . \tag{8.11}$$

Thus when F_{rs} is transformed by a proper Lorentz transformation L , $A_{\rho 4}$ is transformed by the linear transformation T , of which the matrix is quadratic in L as in (8.11). Putting

$$V_\rho = A_{\rho 4} \tag{8.12}$$

we have a complex 3-vector which undergoes a linear transformation. That this is an <u>orthogonal</u> transformation (conserving $V_\rho V_\rho$ or more generally the scalar product $U_\rho V_\rho$) is most easily seen by consideration of the invariants $F_{ab} F_{ab}$ and $F_{ab} F^*_{ab}$. But it is an interesting exercise (left to the reader) to prove from (8.11), using (8.6), that

$$T_{\alpha\gamma} T_{\beta\gamma} = \delta_{\alpha\beta} \tag{8.13}$$

or in matrix language $T\tilde{T} = 1$.

We see then that a proper Lorentz transformation L defines by (8.11) a unique orthogonal transformation T in complex E_3 . We now ask whether an orthogonal transformation T defines a unique Lorentz transformation. We seek then to solve (8.11) for L , satisfying at the same time $L\tilde{L} = 1$, as required for a Lorentz transformation. It is obvious that if L is a solution, so also is $-L$. We shall see that apart from this ambiguity in sign, L is uniquely determined.

The trick for solving (8.11) is to use the ordinary notation for 3-vectors. If we insert parentheses round the first subscripts in (8.11), so that it reads

$$T_{(\rho)\gamma} = L_{(\rho)\alpha} L_{(4)\beta} \varepsilon_{\alpha\beta\gamma} + L_{(\rho)\gamma} L_{(4)4} - L_{(\rho)4} L_{(4)\gamma} , \qquad (8.14)$$

we may regard the suffixes in parentheses as labels distinguishing vectors or scalars and the other suffixes, if Greek, as indicating components. Defining real scalars a_ρ and an imaginary scalar a_4 by

$$i\, a_\rho = L_{(\rho)4} , \qquad i\, a_4 = L_{(4)4} , \qquad (8.15)$$

(the L's are to obey the parity rule and we want the a's also to obey it), we write (8.14) in vector form:

$$\underline{T}_{(\rho)} = \underline{L}_{(\rho)} \times \underline{L}_{(4)} + i\, a_4\, \underline{L}_{(\rho)} - i\, a_\rho\, \underline{L}_{(4)} . \qquad (8.16)$$

The vector $\underline{L}_{(\rho)}$ is to be real and $\underline{L}_{(4)}$ imaginary. The condition $L\tilde{L} = 1$ is equivalent to

$$\underline{L}_{(\rho)} \cdot \underline{L}_{(\sigma)} = \delta_{\rho\sigma} + a_\rho a_\sigma , \qquad \underline{L}_{(\rho)} \cdot \underline{L}_{(4)} = a_\rho a_4 ,$$

$$\underline{L}_{(4)} \cdot \underline{L}_{(4)} = 1 + a_4^{\,2} . \qquad (8.17)$$

To simplify the notation, we shall now omit the parentheses, so that the subscripts are now labels distinguishing vectors. We suspend the summation convention. Put

$$\underline{T}_\rho = \underline{P}_\rho + i\, \underline{Q}_\rho , \qquad (8.18)$$

where $\underline{P}_\rho$ and $\underline{Q}_\rho$ are real vectors satisfying (since $\underline{T}_\rho^2 = 1$)

$$P_\rho^2 - Q_\rho^2 = 1 , \qquad \underline{P}_\rho \cdot \underline{Q}_\rho = 0 . \tag{8.19}$$

Separating the real and imaginary parts of (8.16), we get

$$i\, a_4 \underline{L}_\rho - i\, a_\rho \underline{L}_4 = \underline{P}_\rho , \qquad \underline{L}_\rho \times \underline{L}_4 = i\, \underline{Q}_\rho . \tag{8.20}$$

Here we have 6 equations, and in (8.17) 10 equations more, these 16 equations to be solved for the 16 real numbers contained in $\underline{L}_\rho$, $i\,\underline{L}_4$, a_ρ , $i\, a_4$.

With the aid of (8.17), (8.20) gives

$$i\, a_\rho \underline{L}_\rho + i\, a_4 \underline{L}_4 = - i\, \underline{P}_\rho \times \underline{Q}_\rho . \tag{8.21}$$

Hence, using the first of (8.20),

$$A_\rho \underline{L}_\rho = - i\, a_4 \underline{P}_\rho - a_\rho \underline{P}_\rho \times \underline{Q}_\rho , \qquad A_\rho \underline{L}_4 = - i\, a_\rho \underline{P}_\rho - a_4 \underline{P}_\rho \times \underline{Q}_\rho \tag{8.22}$$

where

$$A_\rho = a_\rho^2 + a_4^2 . \tag{8.23}$$

It remains then only to find the four numbers a_ρ , a_4 . From $\tilde{L}\, L = 1$, it follows that

$$a_1^2 + a_2^2 + a_3^2 + a_4^2 = - 1 . \tag{8.24}$$

We note that, since a_ρ is real, $a_4 \neq 0$.

Taking the self scalar product of the first of (8.22), we get, by (8.17) and (8.19) ,

$$A_\rho^2 (1 + a_\rho^2) = - a_4^2 P_\rho^2 + a_\rho^2 P_\rho^2 Q_\rho^2 = a_\rho^2 P_\rho^4 - A_\rho P_\rho^2 , \tag{8.25}$$

and so

$$(A_\rho + P_\rho^2)\, [a_\rho^2 (A_\rho - P_\rho^2) + A_\rho] = 0 . \tag{8.26}$$

The scalar product of the two equations in (8.22) gives

$$A_\rho^{\,2}\, a_\rho\, a_4 = -\, a_\rho\, a_4\, P_\rho^{\,2} + a_\rho\, a_4\, P_\rho^{\,2}\, Q_\rho^{\,2}\,, \qquad (8.27)$$

or

$$a_\rho\, a_4\, (A_\rho^{\,2} - P_\rho^{\,4}) = 0\,. \qquad (8.28)$$

Thus

$$a_\rho = 0 \quad \text{or} \quad A_\rho^{\,2} = P_\rho^{\,4}\,. \qquad (8.29)$$

Let us follow the consequences of $a_\rho = 0$. By (8.23), $A_\rho \neq 0$. Hence (8.26) gives

$$A_\rho = -\,P_\rho^{\,2}\,. \qquad (8.30)$$

Hence, whether $a_\rho = 0$ or not, we have $A_\rho = \varepsilon\, P_\rho^{\,2}$, $\varepsilon = \pm 1$. Now $P_\rho \neq 0$, and so, if we take $\varepsilon = 1$, (8.26) gives a false result. Hence (8.30) must hold in all cases.

We have now found A_ρ, and it remains to find a_ρ and a_4. By (8.24) we have

$$A_1^{\,2} + A_2^{\,2} + A_3^{\,2} - 2\, a_4^{\,2} = -1 \qquad (8.31)$$

and hence

$$a_4^{\,2} = \tfrac{1}{2}\,(1 - P_1^{\,2} - P_2^{\,2} - P_3^{\,2})\,, \qquad (8.32)$$

a negative value as required, since each P is greater than or equal to unity. Thus

$$a_4 = i\,\varepsilon'\, 2^{-\frac{1}{2}}\,(P_1^{\,2} + P_2^{\,2} + P_3^{\,2} - 1)^{\frac{1}{2}}\,, \quad \varepsilon' = \pm 1\,. \qquad (8.33)$$

The ambiguity of sign cannot be avoided, since (8.14) is quadratic in L.

We could now get a_ρ from (8.23), but that would introduce three more ambiguities in sign, and these do not in fact exist, as we shall see, using another method. For putting $\underline{R}_\rho = \underline{P}_\rho \times \underline{Q}_\rho$, and writing down

$$A_\sigma (A_\rho \underline{L}_4) = A_\rho (A_\sigma \underline{L}_4) ,$$

the second of (8.22) gives

$$A_\sigma (i\, a_\rho \underline{P}_\rho + a_4 \underline{R}_\rho) = A_\rho (i\, a_\sigma \underline{P}_\sigma + a_4 \underline{R}_\sigma) . \qquad (8.34)$$

We wipe out the right hand side by taking the scalar product with $\underline{Q}_\sigma$, and obtain, since by (8.30) $A_\sigma \neq 0$,

$$a_\rho = i\, a_4 (\underline{R}_\rho \cdot \underline{Q}_\sigma) / (\underline{P}_\rho \cdot \underline{Q}_\sigma) . \qquad (8.35)$$

This solves our problem, a_ρ having a common ambiguity of sign with a_4 , and this ambiguity of sign is carried into $\underline{L}_\rho$ and $\underline{L}_4$ by (8.22).

To sum up, given any orthonormal complex triad $\underline{T}_\rho = \underline{P}_\rho + i\,\underline{Q}_\rho$ (or equivalently an orthogonal 3×3 matrix T) , there exists a Lorentz transformation L , unique except for the factor $\varepsilon' = \pm 1$, satisfying (8.11) (or equivalently (8.16)); it is given by (no summation convention!)

$$\begin{aligned} P_\rho^{\,2}\, \underline{L}_\rho &= i\, a_4 \underline{P}_\rho + a_\rho \underline{P}_\rho \times \underline{Q}_\rho , \\ \underline{P}_\rho^{\,2}\, \underline{L}_4 &= i\, a_\rho \underline{P}_\rho + a_4 \underline{P}_\rho \times \underline{Q}_\rho , \end{aligned} \qquad (8.36)$$

with the a's as in (8.33) and (8.35).

Note that in (8.35) the fraction becomes $0/0$ if $\sigma = \rho$. We therefore understand that $\sigma \neq \rho$. This leaves two choices of σ for given ρ , and it is not clear at once that these give the same value to a_ρ . That they do in fact do so may be seen from (7.52) and its cyclic permutations.

As remarked after (8.33), we could get a_ρ from that equation and (8.30), but for ambiguity in signs. It is interesting to write down such values. In the notation of (7.56), we have

$$a_\rho^2 = \tfrac{1}{2} S_\rho . \tag{8.37}$$

Also in terms of the parameters α, β, γ, δ of the canonical form for the orthonormal triad $\underline{T}_\rho$ as in (7.51), we have

$$\begin{aligned} a_1^2 &= P^2 (\beta^2 + \delta^2) , \\ a_2^2 &= Q^2 (\gamma^2 - \beta^2) , \\ a_3^2 &= Q^2 (\alpha^2 - \delta^2) , \\ a_4^2 &= - P^2 (\alpha^2 + \gamma^2) . \end{aligned} \tag{8.38}$$

It is an interesting exercise to show that $\gamma^2 - \beta^2$ and $\alpha^2 - \delta^2$ are positive. We put $\alpha + i \beta = r_1 \exp (i \phi_1)$, $\gamma + i \delta = r_2 \exp (i \phi_2)$ and use (7.52) and the triangle inequality $r_2^2 + 1 > r_1^2$ in the complex plane.

Proper and improper transformations.

Let us go back to the formula (8.11) :

$$T_{\rho\gamma} = L_{\rho\alpha} L_{4\beta} \varepsilon_{\alpha\beta\gamma} + L_{\beta\gamma} L_{44} - L_{\rho 4} L_{4\gamma} . \tag{8.39}$$

This formula connects a 3×3 matrix T with a 4×4 matrix L. If L satisfies $\tilde{L} L = I_4$ then T satisfies $\tilde{T} T = I_3$, where I_3 and I_4 are the unit 3×3 and 4×4 matrices. Also, as we have seen, if T is given, satisfying $\tilde{T} T = I_3$, then there exist two matrices L, one the negative of the other, satisfying (8.39) and $\tilde{L} L = I_4$.

In the line following (8.6), we took L to be proper (det $L = 1$)

in order that ε_{rsmn} should behave like a tensor under the transformation L. Let us now take a wider view and inquire into the proper or improper characters of T and L, related by (8.39) and satisfying

$$\tilde{T}\,T = I_3\,, \qquad \tilde{L}\,L = I_4\,. \tag{8.40}$$

The squares of det T and det L are then unity, and so we may write

$$\det L = \varepsilon_L\,, \qquad \det T = \varepsilon_T \tag{8.41}$$

where these ε's are ± 1.

If we change L and T continuously, these ε's cannot jump from one discrete value to the other. Thus they may be investigated for all cases by considering the four special Lorentz transformations, representative of the types stated:

a) proper, future-preserving: $L = \text{diag}\,(1, 1, 1, 1)$: $\det L = 1$.

b) proper, future-reversing: $L = \text{diag}\,(-1, -1, -1, -1)$: $\det L = 1$.

c) improper, future-preserving: $L = \text{diag}\,(-1, 1, 1, 1)$: $\det L = -1$.

d) improper, future-reversing: $L = \text{diag}\,(1, -1, -1, -1)$: $\det L = -1$.

In cases (a) and (b), (8.39) gives $T = \text{diag}\,(1, 1, 1)$, so that $\det T = 1$. In cases (c) and (d) we get $T = \text{diag}\,(-1, 1, 1)$, so that $\det T = -1$. Thus in all cases we have

$$\varepsilon_T = \varepsilon_L\,, \tag{8.42}$$

and so T and L are either both proper or both improper; in this statement it is not necessary to mention whether L preserves or reverses the future, i.e. transforms the two sheets of the null cone into themselves, or interchanges them.

9. Two theorems about eigenvectors.

Going back to the equation (6.5)

$$K\ J = \lambda\ J\ , \tag{9.1}$$

and the associated characteristic equation

$$\det\ (K - \lambda\ I) = 0\ , \tag{9.2}$$

we recall that K is a complex symmetric 3×3 matrix with zero trace. The equation (9.2) has at least one root, which might be zero, and not more than three roots altogether. If any root λ of (9.2) is substituted in (9.1), there exists at least one eigenvector J corresponding to this eigenvalue λ. The vector J might be null, and there might be several vectors J corresponding to one eigenvalue λ. The complication of the ensuing argument is due to this multiplicity of possibilities. We start by proving the following theorems:

Theorem III: Let J' be any eigenvector corresponding to an eigenvalue λ' and J'' any eigenvector corresponding to an eigenvalue λ'' where $\lambda' \neq \lambda''$. Then J' and J'' are orthogonal and linearly independent.

We are given

$$K\ J' = \lambda'\ J'\ , \qquad K\ J'' = \lambda''\ J''\ , \qquad \lambda' \neq \lambda''\ , \tag{9.3}$$

and we have to prove that

$$\tilde{J}'\ J'' = 0\ . \tag{9.4}$$

From (9.3)

$$\tilde{J}''\ K\ J' - \tilde{J}'\ K\ J'' = \lambda'\ \tilde{J}''\ J' - \lambda''\ \tilde{J}'\ J''\ . \tag{9.5}$$

The left hand side vanishes since K is symmetric. The two scalar products on the right are equal, and so the orthogonality of J' and J'' follows from the last of (9.3).

To prove the second part of the theorem, assume that $J'' = \theta J'$. Then (9.3) give $(\lambda' - \lambda'') J' = 0$, and this contradicts the last of (9.3), it being understood throughout that when we speak of an eigenvector the zero vector is excluded.

Theorem IV: If λ' and λ'' are distinct eigenvalues, and J' and J'' eigenvectors corresponding to them respectively, then J' and J'' cannot both be null.

This follows on combining Theorem I (Section 7) with Theorem III.

10. Class 1 : general case (three distinct eigenvalues).

Suppose that the characteristic equation (9.2) has three distinct roots λ', λ'', λ'''. Since $\mathrm{tr}\, K = 0$, we have

$$\lambda' + \lambda'' + \lambda''' = 0 . \tag{10.1}$$

Let J', J'', J''' be eigenvectors corresponding to these three eigenvalues respectively. By Theorem III we know that these three vectors are orthogonal and no two are collinear. They therefore form an orthogonal triad, and so by Theorem II none of them is null. We can then normalise them so that they form an orthonormal triad, say A, B, C, and then we have

$$K A = \lambda' A , \quad K B = \lambda'' B , \quad K C = \lambda''' C ,$$
$$\tilde{A} A = \tilde{B} B = \tilde{C} C = 1 , \quad \tilde{B} C = \tilde{C} A = \tilde{A} B = 0 . \tag{10.2}$$

On applying the orthogonal transformation T as in (7.28), K transforms into K' :

$$K' = T K T^{-1} = \begin{pmatrix} \tilde{A}KA & \tilde{A}KB & \tilde{A}KC \\ \tilde{B}KA & \tilde{B}KB & \tilde{B}KC \\ \tilde{C}KA & \tilde{C}KB & \tilde{C}KC \end{pmatrix} . \qquad (10.3)$$

By (10.2) this becomes

$$K' = \begin{pmatrix} \lambda' & 0 & 0 \\ 0 & \lambda'' & 0 \\ 0 & 0 & \lambda''' \end{pmatrix} . \qquad (10.4)$$

We recall that K is the complex 3×3 matrix $K = M + N$, where M and N are matrices formed as in (5.5) from the elements of the Weyl tensor. We have then the following result:

If the characteristic equation (9.2) has three distinct roots, there exists a Lorentz transformation which transforms the matrix K into the diagonal form (10.4), and the 6×6 Weyl matrix (5.4) into

$$W = \begin{pmatrix} \alpha' & 0 & 0 & i\beta' & 0 & 0 \\ 0 & \alpha'' & 0 & 0 & i\beta'' & 0 \\ 0 & 0 & \alpha''' & 0 & 0 & i\beta''' \\ i\beta' & 0 & 0 & \alpha' & 0 & 0 \\ 0 & i\beta'' & 0 & 0 & \alpha'' & 0 \\ 0 & 0 & i\beta''' & 0 & 0 & \alpha''' \end{pmatrix} . \qquad (10.5)$$

where the α's and β's are real and

$$\alpha' + \alpha'' + \alpha''' = 0 , \qquad \beta' + \beta'' + \beta''' = 0 . \qquad (10.6)$$

11. <u>Class 2 : double-root case</u> ($\lambda' \neq \lambda'' = \lambda'''$).

Suppose now that the characteristic equation (9.2) has a double root, so that the three roots are $\lambda' \neq \lambda'' = \lambda'''$, satisfying as in (10.1)

$$\lambda' + 2\lambda'' = 0 . \tag{11.1}$$

Then there exists at least one eigenvector J' corresponding to λ' and at least one eigenvector J'' corresponding to λ'' , so that

$$K J' = \lambda' J' , \qquad K J'' = \lambda'' J'' . \tag{11.2}$$

By Theorems III and IV we know that J' and J'' are orthogonal and cannot both be null. There are then three cases to consider:

2a: Neither J' nor J'' null.

2b: J' not null, J'' null, and no non-null J'' exists.

2c: J' null, J'' not null.

We recall that J'' corresponds to the double root.

Class 2a: Normalise J' , J'' into A , B , say, and complete an orthonormal triad with $C = A \times B$. We have then

$$\begin{aligned} &K A = \lambda' A , \qquad K B = \lambda'' B , \\ &\tilde{A} K A = \lambda' , \qquad \tilde{A} K B = 0 , \qquad \tilde{A} K C = 0 , \\ &\qquad\qquad\qquad\quad \tilde{B} K B = \lambda'' , \qquad \tilde{C} K B = 0 . \end{aligned} \tag{11.3}$$

Transforming by T as in (7.28), we get

$$K' = \begin{pmatrix} \lambda' & 0 & 0 \\ 0 & \lambda'' & 0 \\ 0 & 0 & \tilde{C}KC \end{pmatrix} . \tag{11.4}$$

But the trace of K , being invariant under the transformation, must vanish; so, by (11.1), we see that $\tilde{C}KC = \lambda''$, and the reduced matrix is

$$K' = \begin{pmatrix} \lambda' & 0 & 0 \\ 0 & \lambda'' & 0 \\ 0 & 0 & \lambda'' \end{pmatrix} . \qquad (11.5)$$

Note that this is the same as (10.4) with $\lambda''' = \lambda''$.

Class 2b: Now J' is not null and J'' is null, and so we cannot use the latter as part of an orthonormal triad. Normalise J' to A , say; take any unit vector B orthogonal to A , and complete an orthonormal trial with $C = A \times B$. We know that J'' is orthogonal to J' ; therefore it lies in the plane of B and C , and so

$$J'' = \beta B + \gamma C , \qquad (11.6)$$

where $\beta^2 + \gamma^2 = 0$ since J'' is null. Without loss of generality we may then write

$$J'' = B + i C \qquad (11.7)$$

for the null eigenvector corresponding to λ'' . Now

$$K J'' = \lambda'' J'' , \qquad (11.8)$$

and so

$$K A = \lambda' A , \quad K B + i K C = \lambda'' B + i \lambda'' C ; \qquad (11.9)$$

therefore

$$\tilde{A} K A = \lambda' , \quad \tilde{A} K B = \tilde{B} K A = 0 , \quad \tilde{A} K C = \tilde{C} K A = 0,$$
$$\tilde{B} K B + i \tilde{B} K C = \lambda'' , \quad \tilde{C} K B + i \tilde{C} K C = i \lambda''. \qquad (11.10)$$

Putting $\tilde{B} K C = \tilde{C} K B = i a$, we have then

$$\tilde{B} K B = \lambda'' + a , \quad \tilde{C} K C = \lambda'' - a . \qquad (11.11)$$

Transforming with T as in (7.28), we get

$$K' = \begin{pmatrix} \lambda' & 0 & 0 \\ 0 & \lambda'' + a & ia \\ 0 & ia & \lambda'' - a \end{pmatrix} . \tag{11.12}$$

A further reduction is possible. Consider the matrix

$$M = \begin{pmatrix} 0 & 0 & 0 \\ 0 & a & ia \\ 0 & ia & -a \end{pmatrix} \tag{11.13}$$

and the matrix

$$T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\phi & \sin\phi \\ 0 & -\sin\phi & \cos\phi \end{pmatrix} \tag{11.14}$$

ϕ being any complex number. It is easy to see that T is an orthogonal matrix $(T\tilde{T} = 1)$. Applying T to transform M, an easy calculation gives

$$T M T^{-1} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & ae^{2i\phi} & iae^{2i\phi} \\ 0 & iae^{2i\phi} & -ae^{2i\phi} \end{pmatrix} . \tag{11.15}$$

Choose ϕ to satisfy

$$a\, e^{2i\phi} = 1 . \tag{11.16}$$

Then

$$T M T^{-1} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & i \\ 0 & i & -1 \end{pmatrix} . \tag{11.17}$$

Further, for any ϕ,

$$T \operatorname{diag}(\lambda', \lambda'', \lambda'')\, T^{-1} = \operatorname{diag}(\lambda', \lambda'', \lambda'') . \tag{11.18}$$

Now, by (11.11),

$$K' = \operatorname{diag}(\lambda', \lambda'', \lambda'') + M, \tag{11.19}$$

and so its transform under T , with ϕ satisfying (11.16), is

$$K'' = T K' T^{-1} = \begin{pmatrix} \lambda' & 0 & 0 \\ 0 & \lambda'' + 1 & i \\ 0 & i & \lambda'' - 1 \end{pmatrix} . \tag{11.20}$$

This is the reduced form of K for Class 2b.

There is one point in the above argument requiring justification. We cannot solve (11.16) for ϕ if $a = 0$. But we can show that $a \neq 0$. It will be recalled that part of the definition of Class 2b is that no non-null J'' exists. This condition is required to separate Class 2b from Class 2a , for if a non-null J'' existed, then we would be back in Class 2a. Now if $a = 0$, we get by (11.12) $K' = \operatorname{diag}(\lambda', \lambda'', \lambda'')$. The components of any J'' are then to satisfy

$$\lambda' J''_1 = \lambda'' J''_1 , \quad \lambda'' J''_2 = \lambda'' J''_2 , \quad \lambda'' J''_3 = \lambda'' J''_3 . \tag{11.21}$$

Thus, for example, (0, 1, 0) is an eigenvector, and it is not null, contrary to hypothesis. Therefore in Class 2b we have $a \neq 0$, and so (11.16) can be solved.

Class 2c has been included for logical completeness. We shall now see that it cannot exist. For, interchanging λ' and λ'' , J' and J'' , the argument can be carried on precisely as in Class 2b down to (11.12), which now reads

$$K' = \begin{pmatrix} \lambda'' & 0 & 0 \\ 0 & \lambda' + a & ia \\ 0 & ia & \lambda' - a \end{pmatrix} . \qquad (11.22)$$

The characteristic equation $\det (K' - \lambda I) = 0$ is

$$(\lambda'' - \lambda) [(\lambda' + a - \lambda)(\lambda' - a - \lambda) + a^2] = 0 , \qquad (11.23)$$

and this reduces to

$$(\lambda - \lambda'')(\lambda - \lambda')^2 = 0 . \qquad (11.24)$$

But this is absurd, since it is λ'', not λ', which is the double root. Hence Class 2c cannot exist.

12. Class 3 : triple-root case $(\lambda' = \lambda'' = \lambda''' = 0)$.

When the characteristic equation has a triple root, that root is zero, since the sum of the roots is zero. Then $\det K = 0$, and there is at least one eigenvector J' satisfying

$$K J' = 0 . \qquad (12.1)$$

Three cases have to be discussed:

3a : $K \neq 0$, and there exists a non-null eigenvector J'.

3b : There exists no non-null eigenvector; equivalently, any vector J' satisfying (12.1) is null.

3c : $K = 0$.

Class 3a: Normalise the existing J' to A, say. Complete an orthonormal triad A, B, C. Then we have

$$K A = 0 , \quad \tilde{A} K A = \tilde{B} K A = \tilde{C} K A = 0 . \qquad (12.2)$$

Transforming with T as in (7.28), we get a matrix of the form

$$K' = \begin{pmatrix} 0 & 0 & 0 \\ 0 & a & if \\ 0 & if & -a \end{pmatrix} \qquad (12.3)$$

using the facts that K' is symmetric and of zero trace.

The characteristic equation now reads

$$\lambda \, (\lambda^2 - a^2 + f^2) = 0 \, , \qquad (12.4)$$

and hence $f = \pm a$ since all roots are to be zero. The ambiguous sign here is trivial, and we may write

$$K' = \begin{pmatrix} 0 & 0 & 0 \\ 0 & a & ia \\ 0 & ia & -a \end{pmatrix} \, . \qquad (12.5)$$

Since $K \neq 0$ for Class 3a , we have $a \neq 0$ and we can proceed as at (11.14), solving (11.16) for ϕ . Thus we reduce the matrix K to

$$K'' = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & i \\ 0 & i & -1 \end{pmatrix} . \qquad (12.6)$$

This is the same form as (11.19) with $\lambda' = \lambda'' = 0$.

We can now examine the eigenvectors a posteriori. An eigenvector J has to satisfy $K'' J = 0$. In terms of components, this means

$$J_2 + i J_3 = 0 \, , \qquad (12.7)$$

and nothing more. In fact, the totality of eigenvectors are of the form

$$J = (\alpha \, , \beta \, , i\beta) \qquad (12.8)$$

where α and β are any complex numbers. For $\alpha = 0$, we get a null vector, unique to within a scalar factor; for $\alpha \neq 0$ we get an infinity of non-null eigenvectors, all orthogonal to the null eigenvector.

Class 3b: In this case we have a null eigenvector J. Take A a unit vector orthogonal to J and complete an orthonormal triad with B and C. Then we have

$$K J = 0 , \quad J = B + i C , \quad K B + i K C = 0 , \qquad (12.9)$$

and so

$$\tilde{A}KB + i\tilde{A}KC = 0 , \quad \tilde{B}KB + i\tilde{B}KC = 0 , \quad \tilde{C}KB + i\tilde{C}KC = 0 . \qquad (12.10)$$

Transforming with T as in (7.28), and putting

$$\tilde{A} K B = h , \quad \tilde{B} K B = b , \qquad (12.11)$$

we get

$$K' = \begin{pmatrix} 0 & h & ih \\ h & b & ib \\ ih & ib & -b \end{pmatrix} , \qquad (12.12)$$

where we have put $\tilde{A}KA = 0$ because the trace must vanish.

The components of any eigenvector J must satisfy

$$h\,(J_2 + i\,J_3) = 0 , \quad h\,J_1 + b\,(J_2 + i\,J_3) = 0 . \qquad (12.13)$$

If $h = 0$, there exists a non-null eigenvector $J = (1, 0, 0)$, contrary to the specification of Class 3b. Hence $h \neq 0$.

Consider now the matrix

$$T = \begin{pmatrix} -\frac{1}{2} b\, h^{-2} & \frac{1}{2} h\,(1 + h^{-2} - \frac{1}{4} b^2 h^{-4}) & \frac{i}{2} h\,(1 - h^{-2} - \frac{1}{4} b^2 h^{-4}) \\ 1 & \frac{1}{2} b\, h^{-1} & \frac{i}{2} b\, h^{-1} \\ -\frac{i}{2} b\, h^{-2} & \frac{i}{2} h\,(-1 + h^{-2} - \frac{1}{4} b^2 h^{-4}) & \frac{1}{2} h\,(1 + h^{-2} + \frac{1}{4} b^2 h^{-4}) \end{pmatrix} . \qquad (12.14)$$

It is easy to verify that this is an orthogonal matrix for all values of b and h with $h \neq 0$. Further, by direct calculation we find

$$K'' = T K' T^{-1} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & i \\ 0 & i & 0 \end{pmatrix} \qquad (12.15)$$

As for eigenvectors, the components must satisfy

$$J_2 = 0 , \qquad J_1 + i J_3 = 0 , \qquad (12.16)$$

which means that there is essentially only one eigenvector, the null vector $(1, 0, i)$, and all vectors collinear with it.

Class 3c: Now the matrix K vanishes, and every vector is an eigenvector. This rather trivial case completes the classification, which is summarised in the next section.

13. Summary of classification of complex 3 x 3 matrices $K = M + N$, symmetric and of zero trace.

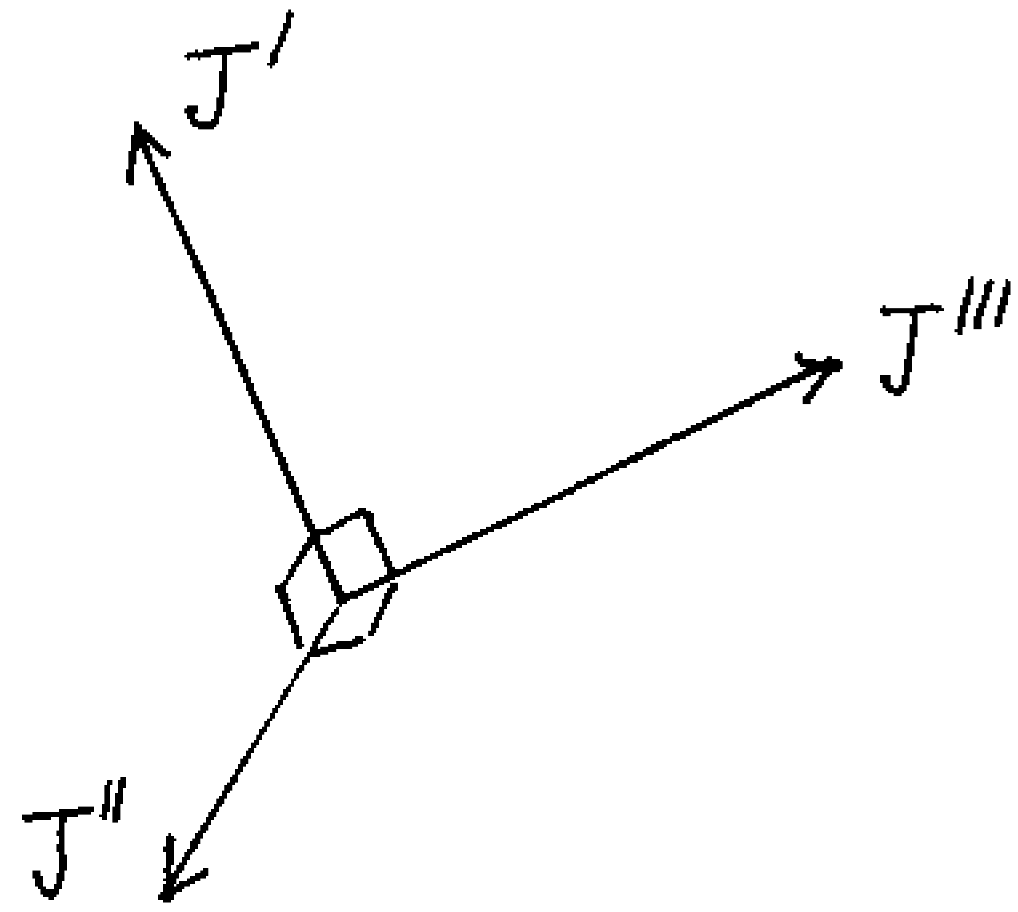
Class 1:

eigenvalues: λ', λ'', λ''' distinct $(\lambda' + \lambda'' + \lambda''' = 0)$;

eigenvectors: 3 non-null, mutually orthogonal, with lines uniquely determined;

reduced matrix: $K = \mathrm{diag}(\lambda', \lambda'', \lambda''')$;

diagram:



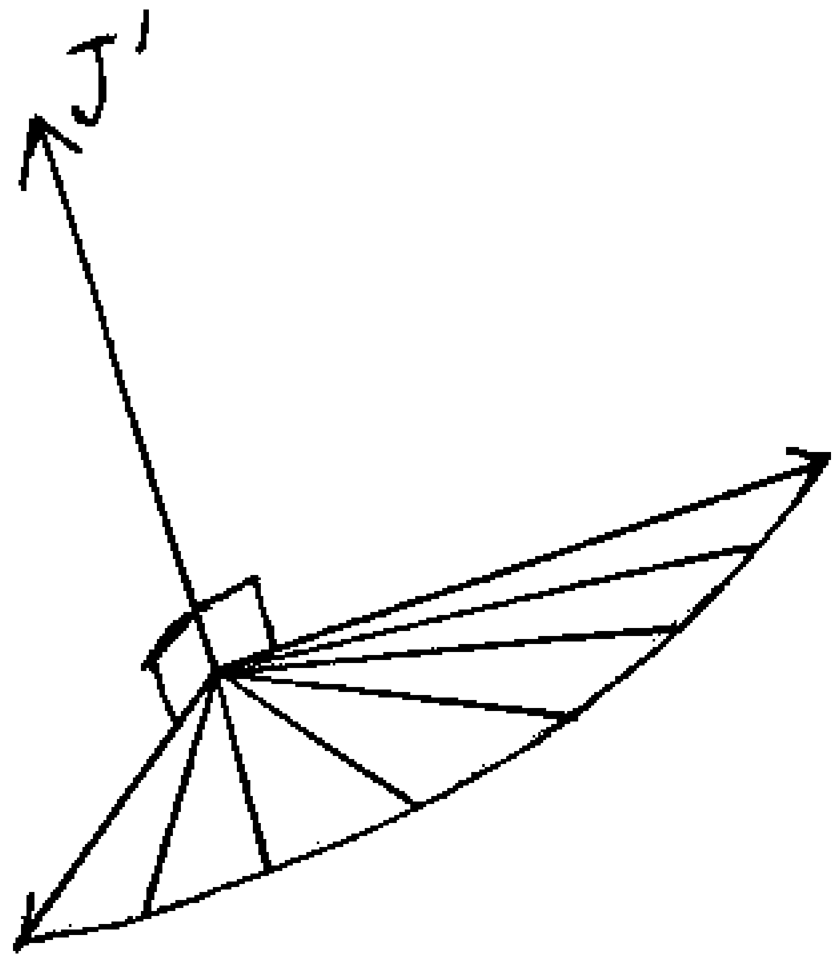
Class 2a:

eigenvalues: $\lambda' \neq \lambda'' = \lambda'''$ $(\lambda' + 2\lambda'' = 0)$;

eigenvectors: J' non-null with line uniquely determined, and all vectors in plane orthogonal to J' ;

reduced matrix: $K = \text{diag}(\lambda', \lambda'', \lambda''')$;

diagram:



<u>Class 2b</u>:

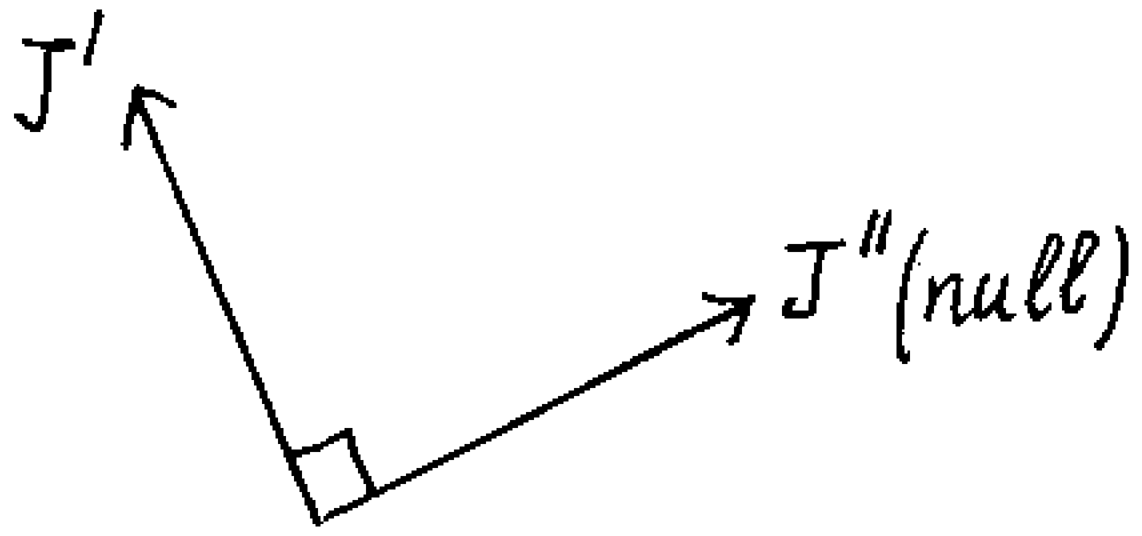
eigenvalues: $\lambda' \neq \lambda'' = \lambda''' \quad (\lambda' + 2\lambda'' = 0)$;

eigenvectors: J' non-null with line uniquely determined; J'' null with line uniquely determined, orthogonal to J' .

reduced matrix:

$$K = \begin{pmatrix} \lambda' & 0 & 0 \\ 0 & \lambda'' + 1 & i \\ 0 & i & \lambda'' - 1 \end{pmatrix}$$

diagram:



<u>Class 3a</u>:

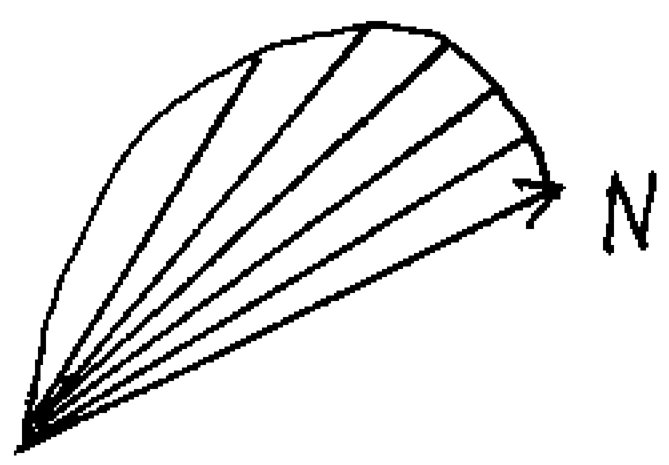
eigenvalues: $\lambda' = \lambda'' = \lambda''' = 0$;

eigenvectors: one null vector N with line uniquely determined, and all vectors orthogonal to N (forming a plane through N) ;

reduced matrix:

$$K = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & i \\ 0 & i & -1 \end{pmatrix}$$

diagram:



Class 3b:

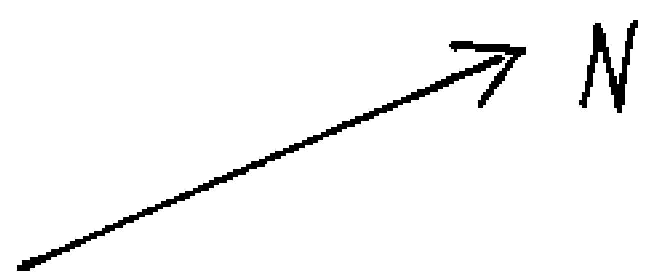
eigenvalues: $\lambda' = \lambda'' = \lambda''' = 0$;

eigenvectors: null vector with line uniquely determined.

reduced matrix:

$$K = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & i \\ 0 & i & 0 \end{pmatrix}$$

diagram:



Class 3c: $(K = 0)$

eigenvalues: $\lambda' = \lambda'' = \lambda''' = 0$;

eigenvectors: all vectors.

The above diagrams are useful as an aid for remembering the several classes. But since we are concerned with complex E_3 , while the diagrams cannot pretend to indicate more than <u>real</u> E_3, they must be used with considerable caution as a basis for reasoning.

The present classification differs in arrangement from that given by Petrov. The connection is as follows:

Petrov's classification	Present classification
I reduced $K = \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix}$ $(a + b + c = 0)$	1 and 2a
II reduced $K = \begin{pmatrix} a & 0 & 0 \\ 0 & b+1 & i \\ 0 & i & c-1 \end{pmatrix}$ $(a + b + c = 0)$	2b and 3a
III reduced $K = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & i \\ 0 & i & 0 \end{pmatrix}$	3b

14. Procedure in classification.

In order to carry out the classification according to the present method, we are to proceed as follows.

1) Suppose we start with real coordinates x^a and a given metric field $g_{ab}(x)$. We calculate the Riemann tensor R_{abcd} and the Weyl tensor W_{abcd} ($W_{abcd} = R_{abcd}$ in vacuo, by virtue of the field equations).

2) Fix attention on some event E at which the classification is to be made. At E we have a matrix g_{ab}. Find a matrix $X^{\bar{c}}_{a}$ so that

$$X^{\bar{c}}_{a} X^{\bar{c}}_{b} = g_{ab} , \qquad (14.1)$$

or, equivalently, so that identically

$$g_{ab}\, dx^a\, dx^b = d\bar{x}^c\, d\bar{x}^c , \qquad d\bar{x}^c = X^{\bar{c}}_{a}\, dx^a . \qquad (14.2)$$

This is best done by changing $g_{ab}\, dx^a\, dx^b$ step by step into a sum of squares of linear differential forms, one a pure imaginary necessarily on account of the assumed signature +2 of g_{ab}. We arrange to have $d\bar{x}^4$ imaginary, thus introducing imaginary time. The matrix $X^{\bar{c}}_{a}$ will then obey the parity rule with respect to the index 4 ; this matrix is of course not unique - we have the freedom of a Lorentz transformation.

3) Define the inverse matrix $X^{a}_{\bar{b}}$ by

$$X^{a}_{\bar{b}} X^{\bar{b}}_{c} = \delta^{a}_{c} , \qquad (14.3)$$

and transform the Weyl tensor at E by

$$\bar{W}_{abcd} = W_{pqrs} X^{p}_{\bar{a}} X^{q}_{\bar{b}} X^{r}_{\bar{c}} X^{s}_{\bar{d}} . \qquad (14.4)$$

4) Write $\bar{W}_{AB}$ out in the Petrov notation in the form of a 6 x 6 matrix

$$\begin{pmatrix} M & N \\ N & M \end{pmatrix} \quad .$$

Check that M and N are symmetric, with M real and N imaginary, both being 3×3 matrices with zero trace.

5) Form the complex 3×3 matrix $K = M + N$, and write out the cubic equation $\det (K - \lambda I) = 0$.

6) Examine the roots of this equation. If they are distinct, we are in Class 1 , and the classification is completed.

7) If there is a double root, we are in Class 2 , and it remains to distinguish between 2a and 2b . To do this, find the (unique) eigenvector J' corresponding to the unrepeated root λ' . Investigate the eigenvectors for the repeated root λ'' . If there are several, we are in Class 2a ; if only one, we are in Class 2b (this unique eigenvector must be null).

8) If all the roots are zero, we are in Class 3. If there are several eigenvectors, we are in Class 3a ; if there is only one, we are in Class 3b (and this unique eigenvector must be null). Class 3c can be recognised early in the work, because it corresponds to the vanishing of the Weyl tensor in any coordinate system.

REFERENCES

Debever, R. 1964. Le rayonnement gravitationnel. Université Libre de Bruxelles. (Stencilled report).

Géhéniau, J. 1957. Une classification des espaces einsteiniens. Comptes Rendus Acad. Sci. Paris 244, 723 .

Petrov, A. Z. 1962. Invariant classification of gravitational fields. Recent Developments in General Relativity (Pergamon Press, Oxford; PWN, Warszawa). p. 371.

Pirani, F. A. E. 1962. Gravitational radiation. Gravitation, edited by L. Witten (Wiley, New York). p. 199.

Synge, J. L. 1956. Relativity: the Special Theory (North-Holland, Amsterdam).

Synge, J. L. 1960. Relativity: the General Theory (North-Holland, Amsterdam).