"Quaternionen und Semivektoren," Comm. Math. Helv. **7** (1935), 141-149.

Quaternions and semi-vectors

By W. SCHERRER, Bern

Translated by D. H. Delphenich

The theory of spinors has been founded in a very satisfactory way by Einstein and Mayer [1]) with the help of new quantities: semi-vectors. Semi-vectors mediate the transition from vectors to spinors in the following sense: One the one hand, they have a natural connection with the group of Lorentz transformations and, on the other hand, they include spinors as a special symmetry type.

The algebraic basis for the Einstein-Mayer theory is defined by a statement that we make in the form of:

Theorem 1: *Any real Lorentz transformation $\mathfrak{C}$ can be represented uniquely, up to sign, as the product of two special Lorentz matrices $\mathfrak{A}$ and $\mathfrak{B}$ that possess the following properties:*

a) *Any arbitrary $\mathfrak{A}$ commutes with any arbitrary $\mathfrak{B}$.*

b) *$\mathfrak{A}$ and $\mathfrak{B}$ are complex conjugates.*

From this theorem, it follows immediately that the theorem in question belongs to Hamilton's theory of quaternions and can therefore be effortlessly derived from it.

We carry out the examination, as far as possible, in the realm of complex orthogonal matrices by recalling the known fact that any orthogonal matrix:

$$\begin{Vmatrix} a_{00} & a_{01} & a_{02} & a_{03} \\ a_{10} & a_{11} & a_{12} & a_{13} \\ a_{20} & a_{21} & a_{22} & a_{23} \\ a_{30} & a_{31} & a_{32} & a_{33} \end{Vmatrix} \tag{1}$$

can be associated with two Lorentz matrices:

$$\begin{Vmatrix} a_{00} & \mp a_{01} & \mp a_{02} & \mp a_{03} \\ \pm a_{10} & a_{11} & a_{12} & a_{13} \\ \pm a_{20} & a_{21} & a_{22} & a_{23} \\ \pm a_{30} & a_{31} & a_{32} & a_{33} \end{Vmatrix} \tag{2}$$

[1]) A. Einstein and E. Mayer, "Semivektoren und Spinoren," Sitzungsberichte der preussischen Akademie de Wissenschaften, 1932.

and conversely.

Hamilton's quaternion algebra defines the appropriate instrument for the treatment of four-dimensional orthogonal matrices. For the reader's enlightenment, I will summarize the basic properties of quaternions, and indeed in such a way that nothing beyond the complex case shall be considered.

Let quaternions be denoted by small German symbols $\mathfrak{a}$, $\mathfrak{b}$, $\mathfrak{c}$, …, $\mathfrak{x}$, $\mathfrak{y}$, $\mathfrak{z}$. An individual quaternion is defined to be a linear combination:

$$\mathfrak{a} = a_0 + a_1\,\mathfrak{e}_1 + a_2\,\mathfrak{e}_2 + a_3\,\mathfrak{e}_3, \tag{3}$$

where the "components" a_0, a_1, a_2, a_3 can be arbitrary complex numbers, while the units that appear along with the usual unit 1 obey the following multiplication rules:

$$\begin{aligned}
&\mathfrak{e}_1^2 = \mathfrak{e}_2^2 = \mathfrak{e}_3^2 = -1, \\
&\mathfrak{e}_2\,\mathfrak{e}_3 = \mathfrak{e}_1, \qquad \mathfrak{e}_3\,\mathfrak{e}_1 = \mathfrak{e}_2, \qquad \mathfrak{e}_1\,\mathfrak{e}_2 = \mathfrak{e}_3, \\
&\mathfrak{e}_3\,\mathfrak{e}_2 = -\,\mathfrak{e}_1, \quad \mathfrak{e}_1\,\mathfrak{e}_3 = -\,\mathfrak{e}_2, \quad \mathfrak{e}_2\,\mathfrak{e}_1 = -\,\mathfrak{e}_3\,.
\end{aligned} \tag{4}$$

The rules of ordinary algebra are required for addition, with the extension that the zero will be represented by:

$$0 + 0 \cdot \mathfrak{e}_1 + 0 \cdot \mathfrak{e}_2 + 0 \cdot \mathfrak{e}_3\,.$$

Along with the rules (4), the distributive law of multiplication shall be valid. With these assumptions, one finds the following expression for the product of two quaternions $\mathfrak{a}$ and $\mathfrak{b}$:

$$\begin{aligned}
\mathfrak{ab} = \quad & a_0\,b_0 - a_1\,b_1 - a_2\,b_2 - a_3\,b_3 \\
& + (a_1\,b_0 + a_0\,b_1 - a_3\,b_2 + a_2\,b_3)\,\mathfrak{e}_1 \\
& + (a_2\,b_0 + a_3\,b_1 + a_0\,b_2 - a_1\,b_3)\,\mathfrak{e}_2 \\
& + (a_3\,b_0 - a_2\,b_1 + a_1\,b_2 + a_0\,b_3)\,\mathfrak{e}_3\,.
\end{aligned} \tag{5}$$

One now easily verifies the fundamental *associative law:*

$$(\mathfrak{ab})\,\mathfrak{c} = \mathfrak{a}\,(\mathfrak{bc}). \tag{6}$$

From (4), the commutative law does not enter in.

Let the quaternion that is "conjugate" to $\mathfrak{a}$ be denoted by $\mathfrak{a}^*$:

$$\mathfrak{a}^* = a_0 - a_1\,\mathfrak{e}_1 - a_2\,\mathfrak{e}_2 - a_3\,\mathfrak{e}_3\,. \tag{3*}$$

One easily confirms the validity of the important relations:

$$(\mathfrak{ab})^* = \mathfrak{b}^*\,\mathfrak{a}^*, \tag{7}$$

$$\mathfrak{a}\mathfrak{a}^* = a_0^2 + a_1^2 + a_2^2 + a_3^2 = \mathfrak{a}^*\mathfrak{a} = \text{scalar}. \tag{8}$$

A scalar and a quaternion whose last three components vanish are obviously equivalent concepts in the algebra of quaternions. For the sake of brevity, if we refer to the quantity $\mathfrak{a}\mathfrak{a}^*$ as the "length" of the quaternion then we have:

Theorem 2: *If an arbitrary quaternion* $\mathfrak{x}$ *is multiplied on the left or the right by a quaternion of length* 1 *then its length remains unchanged.*

The assertion thus follows from the fact that $\mathfrak{a}\mathfrak{a}^* = 1$ and $\mathfrak{y} = \mathfrak{a}\mathfrak{x}$ imply the equation:

$$\mathfrak{y}\mathfrak{y}^* = \mathfrak{x}\mathfrak{x}^*.$$

On the basis of (7), (8), (9), calculation yields:

$$\begin{aligned} \mathfrak{y}\mathfrak{y}^* &= (\mathfrak{a}\mathfrak{x})\,(\mathfrak{a}\mathfrak{x})^* \\ &= (\mathfrak{a}\mathfrak{x})\,(\mathfrak{x}^*\mathfrak{a}^*) \\ &= \mathfrak{a}\,(\mathfrak{x}\mathfrak{x}^*)\,\mathfrak{a}^* \\ &= \mathfrak{a}\mathfrak{a}^* \cdot \mathfrak{x}\mathfrak{x}^* = \mathfrak{x}\mathfrak{x}^*. \end{aligned}$$

The corresponding proof for right multiplication proceeds naturally in an analogous way.

If we now regard $\mathfrak{x}$ as variable then, from the theorem that was just proved, the multiplication:

$$\mathfrak{y} = \mathfrak{a}\mathfrak{x}$$

seems to be a special orthogonal transformation whose complete expression, as given by (5), is:

$$\begin{aligned} y_0 &= a_0\,x_0 - a_1\,x_1 - a_2\,x_2 - a_3\,x_3 \\ y_1 &= a_1\,x_0 + a_0\,x_1 - a_3\,x_2 + a_2\,x_3 \\ y_2 &= a_2\,x_0 + a_3\,x_1 + a_0\,x_2 - a_1\,x_3 \\ y_3 &= a_3\,x_0 - a_2\,x_1 + a_1\,x_2 + a_0\,x_3\,. \end{aligned} \tag{9}$$

An analogous statement is true for $\mathfrak{y} = \mathfrak{x}\mathfrak{b}$, and we obtain the result:

The two quaternion products:

$$\mathfrak{y} = \mathfrak{a}\mathfrak{x}, \qquad \mathfrak{y} = \mathfrak{x}\mathfrak{b} \tag{10}$$

are, when:

$$\mathfrak{a}\mathfrak{a}^* = 1, \qquad \mathfrak{b}\mathfrak{b}^* = 1, \tag{11}$$

equivalent to the orthogonal transformations:

$$\mathfrak{y} = \mathfrak{A}\mathfrak{x}, \qquad \mathfrak{y} = \mathfrak{B}\mathfrak{x}, \tag{12}$$

respectively, whose matrices are given by:

$$\mathfrak{A} = \begin{Vmatrix} a_0 & -a_1 & -a_2 & -a_3 \\ a_1 & a_0 & -a_3 & a_2 \\ a_2 & a_3 & a_0 & -a_1 \\ a_3 & -a_2 & a_1 & a_0 \end{Vmatrix}, \qquad \mathfrak{B} = \begin{Vmatrix} b_0 & -b_1 & -b_2 & -b_3 \\ b_1 & b_0 & b_3 & -b_2 \\ b_2 & -b_3 & b_0 & b_1 \\ b_3 & b_2 & -b_1 & b_0 \end{Vmatrix}. \tag{13}$$

As a result of (11), these matrices depend upon three complex parameters. In addition, as one establishes immediately, both of them can be continuously transported to the identity.

From this it follows that the product:

$$\mathfrak{y} = \mathfrak{a}\mathfrak{x}\mathfrak{b} \tag{14}$$

encompasses a 6-parameter family of orthogonal transformations that is continuously connected to the identity, and thus all proper orthogonal transformations. We cannot go into the details here of this and similar "completeness results" that follow with the help of infinitesimal transformations.

We now prove conversely that the representation of all proper orthogonal transformations by means of (14) is unique; i.e., that from:

$$\mathfrak{a}\mathfrak{x}\mathfrak{b} \equiv \mathfrak{c}\mathfrak{x}\mathfrak{d} \tag{15}$$

and

$$\mathfrak{a}\mathfrak{a}^* = \mathfrak{b}\mathfrak{b}^* = \mathfrak{c}\mathfrak{c}^* = \mathfrak{d}\mathfrak{d}^* = 1, \tag{16}$$

it follows that either:

$$\mathfrak{a} = \mathfrak{c} \quad \text{and} \quad \mathfrak{b} = \mathfrak{d}$$

or

$$\mathfrak{a} = -\,\mathfrak{c} \text{ and} \quad \mathfrak{b} = -\,\mathfrak{d}.$$

To this end, we multiply the identity (15) on the left by $\mathfrak{c}^*$ and on the right by $\mathfrak{b}^*$, and from (16), we obtain:

$$\mathfrak{c}^*\mathfrak{a}\mathfrak{x} \equiv \mathfrak{x}\mathfrak{d}\mathfrak{d}^*.$$

From this, it follows that:

$$\mathfrak{c}^*\mathfrak{a} = \mathfrak{d}\mathfrak{d}^* = \text{scalar}.$$

We can also set:

$$\mathfrak{c}^*\mathfrak{a} = \lambda \cdot \mathfrak{a}^*\mathfrak{a},$$

where λ is a suitable scalar, or:

$$(\mathfrak{c}^* - \lambda\,\mathfrak{a}^*)\,\mathfrak{a} = 0.$$

Right-multiplication by $\mathfrak{a}^*$ yields:

$$\mathfrak{c}^* - \lambda\,\mathfrak{a}^* = 0,$$

and from this, it follows that:

$$\mathfrak{c} - \lambda \cdot \mathfrak{a} = 0.$$

One then has:

$$\mathfrak{c}\mathfrak{c}^* = \lambda^2\,\mathfrak{a}\mathfrak{a}^* \quad \text{or} \qquad \lambda^2 = \pm 1,$$

and we ultimately obtain $\mathfrak{c} = \pm\,\mathfrak{a}$, and in an analogous way, $\mathfrak{d} = \pm\,\mathfrak{b}$, from which the uniqueness of the representation (14) – except for a sign – follows.

We now go from the quaternionic representation (14) to the matrix representation, and must therefore pay special attention to the associativity of the composition in the product $\mathfrak{a}\mathfrak{x}\mathfrak{b}$. The equivalence of the representations (10) and (12) gives, in a notation that is simple to understand:

$$\mathfrak{a}\,(\mathfrak{x}\mathfrak{b}) \equiv \mathfrak{A}\mathfrak{B}\mathfrak{x},$$
$$(\mathfrak{a}\mathfrak{x})\;\mathfrak{b} \equiv \mathfrak{B}\mathfrak{A}\mathfrak{x}.$$

From the associative law (6), it then follows that:

$$\mathfrak{A}\mathfrak{B}\mathfrak{x} \equiv \mathfrak{B}\mathfrak{A}\mathfrak{x},$$

or

$$\mathfrak{A}\mathfrak{B} = \mathfrak{B}\mathfrak{A}. \tag{17}$$

On the other hand, we can set:

$$\mathfrak{a}\mathfrak{x}\mathfrak{b} \equiv \mathfrak{C}\mathfrak{x},$$

and thus deal with any proper orthogonal matrix $\mathfrak{C}$.

We express the results that we have arrived at as:

Theorem 3: *Any proper orthogonal matrix $\mathfrak{C}$ can be represented uniquely, up to sign, as the product of two special proper orthogonal matrices $\mathfrak{A}$ and $\mathfrak{B}$ that possess the following property:*

Any arbitrary $\mathfrak{A}$ commutes with any arbitrary $\mathfrak{B}$.

The types for $\mathfrak{A}$ and $\mathfrak{B}$ that come into consideration are given explicitly by equations (13) and must then fulfill, in addition to the conditions (11), only the equations:

$$a_0^2 + a_1^2 + a_2^2 + a_3^2 = b_0^2 + b_1^2 + b_2^2 + b_3^2 = 1. \tag{11$'$}$$

Obviously, Theorem 3 agrees with Theorem 1, up to reality conditions, to which we would now like to turn.

The reality of the matrices (13) is evidently sufficient for the reality of the proper orthogonal transformations, as a result of the completeness and uniqueness of the representation (14) that is likewise true here, and also necessary.

As far as the complex Lorentz transformations are concerned, a completely corresponding theorem follows immediately from Theorem 3, with the help of the transition from (1) to (2), and conversely. Thus, the following matrices enter in place of the matrices (13):

$$\mathfrak{A} = \begin{Vmatrix} a_0 & ia_1 & ia_2 & ia_3 \\ ia_1 & a_0 & -a_3 & a_2 \\ ia_2 & a_3 & a_0 & -a_1 \\ ia_3 & -a_2 & a_1 & a_0 \end{Vmatrix}, \qquad \mathfrak{B} = \begin{Vmatrix} b_0 & ib_1 & ib_2 & ib_3 \\ ib_1 & b_0 & b_3 & -b_2 \\ ib_2 & -b_3 & b_0 & b_1 \\ ib_3 & b_2 & -b_1 & b_0 \end{Vmatrix}. \tag{18}$$

In contrast to the matrices (13), these matrices can be assumed to be complex conjugate to each other, but the individual matrices cease to depend upon three complex parameters. In this way, one obtains a proper and real Lorentz transformation that depends upon 6 real parameters as a product. We therefore also have a representation here of all real proper Lorentz transformations. Then, in turn, it follows from the uniqueness of the product representation that the factors are necessarily complex conjugate. Thus, the Einstein-Mayer theorem that is at the forefront of their work is proved completely.

Quaternion algebra also seems to offer certain advantages for the construction of invariants. For this, it is generally sufficient for one to carry out all developments for orthogonal transformations, and then go over to Lorentz transformations at the conclusion with the help of the transformation $x_0 = it$.

We shall explain this process to the extent that it remains in the realm of orthogonal transformations. As is known, by restricting to orthogonal coordinates the difference between upper and lower indices goes away. We now represent any transformation $\mathfrak{x}' = \mathfrak{C}\mathfrak{x}$ quaternionically by $\mathfrak{x}' = \mathfrak{a}\mathfrak{x}\mathfrak{b}$ and define:

1. A *vector* $\mathfrak{x} = (x_0, x_1, x_2, x_3)$ is a quaternion that transforms according to the equation:

$$\mathfrak{x}' = \mathfrak{a}\mathfrak{x}\mathfrak{b}. \tag{19}$$

2. A *left vector* $\underline{\mathfrak{x}} = (x_{\bar{0}}, x_{\bar{1}}, x_{\bar{2}}, x_{\bar{3}})$ is a quaternion that transforms according to:

$$\underline{\mathfrak{x}}' = \mathfrak{a}\underline{\mathfrak{x}}\,. \tag{19a}$$

3. A *right vector* $\underline{\underline{\mathfrak{x}}} = (x_{\bar{\bar{0}}}, x_{\bar{\bar{1}}}, x_{\bar{\bar{2}}}, x_{\bar{\bar{3}}})$ is a quaternion that transforms according to:

$$\underline{\underline{\mathfrak{x}}}' = \underline{\underline{\mathfrak{x}}}\mathfrak{b}\,. \tag{19b}$$

Naturally, the $\mathfrak{a}$ and $\mathfrak{b}$ that enter into (19), (19*a*), (19*b*) must always arise from one and the same orthogonal matrix $\mathfrak{C}$.

One can now define mixed tensors of arbitrarily higher degree in the usual way by component-wise multiplication and addition of the vector types that were defined above. In particular, one then has:

Theorem 4: *If:*

$$F(\mathfrak{x}, \mathfrak{y}, \mathfrak{z}, \ldots) \equiv a_{i\bar{k}\bar{\bar{l}}\ldots}\, x_i\, y_{\bar{k}}\, z_{\bar{\bar{l}}} \cdots$$

is an invariant multilinear form then its coefficients $a_{i\bar{k}\bar{\bar{l}}\ldots}$ *define a tensor of the type that is indicated by the indices, and conversely.*

Quaternionic multiplication obviously defines a process for generating invariants. From (19), (19*a*), and (19*b*), in conjunction with (7), one immediately finds the invariant quaternions that are fundamental to the Einstein-Mayer theory:

$$\underline{\mathfrak{x}}^*\underline{\mathfrak{y}}\,, \qquad \underline{\mathfrak{x}}\,\underline{\mathfrak{y}}^* \tag{20}$$

and

$$\underline{\mathfrak{x}}^*\mathfrak{y}\,\underline{\mathfrak{z}}^*\,, \qquad \underline{\mathfrak{x}}\,\mathfrak{y}^*\underline{\mathfrak{z}}\,. \tag{21}$$

Any of these quaternions produces four invariant forms, and therefore, from Theorem 4, just as many numerically invariant tensors. From (20) and (21), one thus obtains the four tensors of type $a_{\bar{i}\bar{k}}$, $a_{\bar{\bar{i}}\bar{\bar{k}}}$, $a_{\bar{i}k\bar{\bar{l}}}$, $a_{\bar{\bar{i}}k\bar{l}}$. The two types that arise from (21) are, however, equivalent, as one sees from the equation $(\underline{\mathfrak{x}}\,\mathfrak{y}^*\underline{\mathfrak{z}})^* = \underline{\mathfrak{x}}^*\mathfrak{y}\,\underline{\mathfrak{z}}^*$.

The calculation of the first expression in (20) yields, on the basis of (3^*) and (5), the invariant bilinear form:

$$\left.\begin{array}{r} x_{\bar{0}}y_{\bar{0}} + x_{\bar{1}}y_{\bar{1}} + x_{\bar{2}}y_{\bar{2}} + x_{\bar{3}}y_{\bar{3}} \\ -x_{\bar{1}}y_{\bar{0}} + x_{\bar{0}}y_{\bar{1}} + x_{\bar{3}}y_{\bar{2}} - x_{\bar{2}}y_{\bar{3}} \\ -x_{\bar{2}}y_{\bar{0}} - x_{\bar{3}}y_{\bar{1}} + x_{\bar{0}}y_{\bar{2}} + x_{\bar{1}}y_{\bar{3}} \\ -x_{\bar{3}}y_{\bar{0}} + x_{\bar{2}}y_{\bar{1}} - x_{\bar{1}}y_{\bar{2}} + x_{\bar{0}}y_{\bar{3}}. \end{array}\right\} \tag{22}$$

Comparison of this with the Einstein-Mayer formulas (41) and (41a), after performing a substitution $x_{\bar{0}} = ix_{\bar{4}}$, $y_{\bar{0}} = iy_{\bar{4}}$, shows that the left-vectors correspond to the semi-vectors of the second kind. The overbar on the indices would then have to be doubled.

As we remarked already, the coefficients of the forms (22) define special numerically-invariant tensors. One arrives at those tensors that correspond to the tensors $c_{\bar{\rho}\bar{\sigma}}$ of Einstein and Mayer [2]) from them by linear combination. Analogously, one obtains the general third-rank tensors from the invariants (21).

However, the application of Theorem 4 to (22) also admits the conclusion that the four rows:

$$(\;\;x_{\bar{0}}\,,\quad x_{\bar{1}}\,,\quad x_{\bar{2}}\,,\quad x_{\bar{3}}\,) \tag{23_0}$$

$$(-x_{\bar{1}}\,,\quad x_{\bar{0}}\,,\quad x_{\bar{3}}\,, -x_{\bar{2}}\,) \tag{23_1}$$

$$(-x_{\bar{2}}\,, -x_{\bar{3}}\,,\quad x_{\bar{0}}\,, -x_{\bar{1}}\,) \tag{23_2}$$

$$(-x_{\bar{3}}\,,\quad x_{\bar{2}}\,, -x_{\bar{1}}\,,\quad x_{\bar{0}}\,) \tag{23_3}$$

---

[2]) A. Einstein and W. Mayer, "Die Diracgleichungen für Semivektoren," Amsterdamer Berichte, 1933.

simultaneously define a system of left-vectors. The same thing is therefore also true for all linear combinations of the rows. If one then adds $-i$ times the second row to the first one then one obtains the special left-vector:

$$\begin{aligned} &(x_{\bar 0} + i x_{\bar 1} ,\ x_{\bar 1} - i x_{\bar 0} ,\ x_{\bar 2} - i x_{\bar 3} ,\ x_{\bar 3} + i x_{\bar 2}) \\ &= (x_{\bar 0} + i x_{\bar 1} , - i (x_{\bar 0} + i x_{\bar 1}), x_{\bar 2} - i x_{\bar 3} , i (x_{\bar 2} - i x_{\bar 3})) \\ &= (z_{\bar 0} , z_{\bar 1} , z_{\bar 2} , z_{\bar 3}) , \end{aligned} \tag{24}$$

whose components fulfill the relations:

$$z_{\bar 1} = - i z_{\bar 0} , \qquad z_{\bar 3} = i z_{\bar 2} .$$

Therefore, two components suffice for the description of this left-vector. Analogously, by adding i times the second row to the first one, one obtains the left-vector:

$$(x_{\bar 0} - i x_{\bar 1} , i (x_{\bar 0} - i x_{\bar 1}), x_{\bar 2} + i x_{\bar 3} , - i (x_{\bar 2} + i x_{\bar 3})) = (u_{\bar 0} , u_{\bar 1} , u_{\bar 2} , u_{\bar 3}) , \tag{25}$$

with the relations:

$$u_{\bar 1} = i u_{\bar 0} , \qquad u_{\bar 3} = - i u_{\bar 2} .$$

The independent components of the left-vectors (24) and (25) thus constitute two 2-component quantities – so-called *spinors* – that are determined by the left-vector (23_0) and, in turn, determine it.

(Received on 8 September 1934)