Tensors and Matrices

Hartmut Neff

January 18, 2021

Abstract

1 Introduction

The aim of physics is to describe the behaviour of objects in nature. These objects are defined to be, as the name suggest, objective and hence independent of who is looking at them. But since observers might use different methods, say different units, to measure an object, these differences have to be compensated by the numerical values in an inverse way. Therefore, the description of an object has a dualistic character, due to the fact that we want to describe an object invariantly.

Say, we want to work out how many cakes we have. Let's look at the dual structure. Let's assume that one observer counts the cakes in whole cakes, and the other observer cuts all the cakes in half to count them in half cakes. For 5 cakes, the observers would find:

$$(5)(c) \tag{1}$$

and

$$(2*5)(2^{-1}*c), \tag{2}$$

demonstrating, that the units and the numerical values transform inversely to each other.

For vector spaces, we find the same dual structure in two ways. Firstly, vectors themselves can describe an object in nature, for instance a distance in space-time. Since this distance is independent of the chosen basis system, the coefficients and basis vectors have to transform inversely to each other. Secondly, the length of a vector, defined through a metric, is an invariant object as well. Therefore, the contraction of vectors, which produces a scalar, will also result in a dual structure. This leads to the definition of a dual vector space, different from the original one. Vectors and dual vectors transform inversly to each other.

In a vector space, to create invariant objects, one has to combine objects that transform inversely to each other. It is common practice to distinguish these objects by upper and lower indices, and to combine them by the summation convention over pairs of equal upper and lower indices. Say, we start by giving vectors lower indices. Then, coefficients of vectors have to have upper indices and so do dual vectors. On the other hand, coefficients of dual vectors have to have lower indices.

It is important to remember that the mathematical meaning is given by the algebraic structure and that the upper and lower indices only serve as a reminder of what algebraic structure we are dealing with.

The same can be said about matrices, as they are collections of coefficients. The algebraic structure that a matrix represents, depends on its combination of vectors and covectors. Accordingly, the transformation property of a matrix depends on its type of combination as well. For instance, a matrix could represent a linear transformation or a metric. Or, in other words, a square of numbers does not have a mathematical meaning, instead it has to defined what it means.

Mostly, in teaching this subject, the algebraic structure is solely represented by coefficients and their upper and lower indices. Here, instead, we want to represent the algebraic structure by the use of bra and ket vectors. A linear transformation will then be given by a combination of a ket and a bra vector, a metric by a combination of two ket vectors. This approach, for instance, might help to clarify situations where the convention of upper and lower indices to represent the algebraic structure is given up, at least implicitly, in favour of the row times column matrix multiplication convention, as when dealing with the transpose of a matrix.

When writing down expressions with explicit upper and lower indices, it is important to remember, that they represent explicit sums. Therefore, the order of the terms is irrelevant. This kind of expressions are not matrix multiplications, but just sums of numbers.

2 Vector Spaces

A vector is given by the product of coefficients and a ket symbol, as

$$|x\rangle = x^i |e_i\rangle, \tag{3}$$

where

$$|e_i\rangle \tag{4}$$

represent basis vectors. $|x\rangle$ is an element of the vector space and given by the explicitly invariant expression on the right hand side, with its sum over one upper and one lower index.

A covector is given by the product of coefficients and a bra symbol, as

$$\langle x| = x_i \langle e^i|, \tag{5}$$

where

$$\langle e^i| \tag{6}$$

represent basis covectors.

The coefficients x^i constitue a vector space as well. Accordingly, the coefficients of the covectors x_i form the corresponding covector space.

Covectors map vectors to real numbers. We define an orthonormal basis of the vector and covector space as

$$\langle e^i|\big(|e_j\rangle\big) = \langle e^i|e_j\rangle = \delta^i{}_j, \tag{7}$$

where the first expression emphasizes that covectors are maps and where we left out the brakets in the second expression for convenience. The δ

symbol just represents numbers, 1 for $i = j$ and 0 for $i \neq j$. Therefore, it doesn't matter where we put the indices. If δ were to represent a trivial linear transformation instead, the postition of the indices would matter.

Later we will translate vectors and covectors into each other with the help of a metric. But first, we will take a look at linear trnsformations.

2.1 Linear Transformations

A linear transformation A is a map between vector spaces and is given by

$$|\hat{x}\rangle = A|x\rangle = \left(a^i{}_j|e_i\rangle\langle e^j|\right)|x\rangle. \tag{8}$$

Hence,

$$\hat{x}^i|e_i\rangle = |\hat{x}\rangle = a^i{}_j|e_i\rangle\langle e^j|x\rangle = a^i{}_j|e_i\rangle\langle e^j|x^k|e_k\rangle \tag{9}$$

$$= a^i{}_j|e_i\rangle x^k\langle e^j|e_k\rangle = a^i{}_j x^j|e_i\rangle. \tag{10}$$

Therefore

$$\hat{x}^i = a^i{}_j x^j. \tag{11}$$

The horizontal displacements of the indices of $a^i{}_j$ are necessary to show that the matrix represents a map between vetor spaces. A map between covector spaces takes the following form,

$$\langle\hat{x}| = A\langle x| = \left(a_i{}^j\langle e^i|\right)\langle x|e_j\rangle \tag{12}$$

with

$$\hat{x}_i = a_i{}^j x_j. \tag{13}$$

Let's now define a map B, which linearly transforms a system of basis vectors to a new system of basis vectors. We will write this transformation backwards for reasons that will become clear later.

$$|e_j\rangle = |\hat{e}_i\rangle\langle\hat{e}^i|e_j\rangle = b^i{}_j|\hat{e}_i\rangle, \tag{14}$$

with the trivial transformation

$$I = \delta^i{}_j|e_i\rangle\langle e^j| = |e_i\rangle\langle e^i|. \tag{15}$$

With B^{-1} defined to be the inverse of B, it follows

$$(b^{-1})^j{}_i|e_j\rangle = |\hat{e}_i\rangle. \tag{16}$$

A covector basis changes likes this:

$$\langle\hat{e}^i| = \langle\hat{e}^i|e_j\rangle\langle e^j| = b^i{}_j\langle e^j| \tag{17}$$

and

$$(b^{-1})^i{}_j\langle\hat{e}^i| = \langle e^j| \tag{18}$$

The linear map A transforms under the above change of basis as follows:

$$\hat{a}^i{}_j|\hat{e}_i\rangle\langle\hat{e}^j| = \hat{a}^i{}_j(b^{-1})^k{}_i|e_k\rangle b^j{}_l\langle e^l| \tag{19}$$

$$= (b^{-1})^k{}_i\hat{a}^i{}_j b^j{}_l|e_k\rangle\langle e^l| = a^k{}_l|e_k\rangle\langle e^l|. \tag{20}$$

Hence

$$a^k{}_l = (b^{-1})^k{}_i\ \hat{a}^i{}_j\ b^j{}_l, \tag{21}$$

or in matrix form

$$A = B^{-1}\hat{A}B. \tag{22}$$

2.2 The Metric

A metric defines the length, or more general, angles between vectors. Hence it has to represent an algebraic structure that contracts 2 vectors. A metric G can therefore defined to be:

$$G = g_{ij} \langle e^i | \langle e^j |. \tag{23}$$

The length of a vector x is then given by

$$g_{ij} \langle e^i | \langle e^j || x \rangle | x \rangle = g_{ij} \langle e^i | \langle e^j | x^k x^l | e_k \rangle | e_l \rangle = \tag{24}$$

$$g_{ij} x^k x^l \langle e^i | e_k \rangle \langle e^j | e_l \rangle = g_{ij} x^i x^j = x^i g_{ij} x^j. \tag{25}$$

Under a transformation of the basis vectors, the metric changes as follows:

$$G = \hat{g}_{ij} \langle \hat{e}^i | \langle \hat{e}^j | = b^i{}_k \hat{g}_{ij} b^j{}_l \langle e^k | \langle e^l | = g_{kl} \langle e^k | \langle e^l |. \tag{26}$$

Hence

$$g_{kl} = b^i{}_k \hat{g}_{ij} b^j{}_l, \tag{27}$$

and in matrix notation

$$G = B^T \hat{G} B. \tag{28}$$

The use of the matrix transpose can lead to confusion. The transpose of a matrix

$$M = m^i{}_j | e_i \rangle \langle e^j | \tag{29}$$

is given by

$$M^T = N = n^i{}_j | e_i \rangle \langle e^j | \tag{30}$$

with

$$n^j{}_i = m^i{}_j. \tag{31}$$

That means, that the transpose of a matrix is still a linear map between vector spaces. The matrix notation seems to suggest that instead

$$n_i{}^j = m^i{}_j, \tag{32}$$

as required by the row times column matrix mulitplication convention. If we give up that the position of the indices resemble the algebraic structure, we can swap the coefficients horizontally for the purpose of row times column multiplication. But, since this is bound to lead to confusion, this is only used implicitly for the matrix notation and never explicitly.

Let's look again at the contraction of two vectors with the metric:

$$g_{ij} \langle e^i | \langle e^j | \; | x \rangle | y \rangle = g_{ij} \langle e^i | x \rangle \langle e^j | y \rangle = \tag{33}$$

$$= \left[\; x^i g_{ij} \langle e^j | \; \right] \left[\; | y \rangle \; \right]. \tag{34}$$

We know that the result of this computation is a scalar. The term in the first braket is therefore a dual vector, generated from the vector x. We can write this as

$$= \left[\; x_j \langle e^j | \; \right] \left[\; | y \rangle \; \right], \tag{35}$$

with

$$x_j = x^i g_{ij}. \tag{36}$$

x_j is the coefficient covector to x^j. The metric therefore translates between vectors and covectors. Therefore

$$\langle x|y\rangle = x_i y^i. \tag{37}$$

We could have chosen instead

$$g_{ij}\langle e^i|\langle e^j|\ |x\rangle|y\rangle = g_{ij}\langle e^i|x\rangle\langle e^j|y\rangle \tag{38}$$

$$= \left[\ g_{ij}y^j\langle e^i|\ \right]\ \left[\ |x\rangle\ \right] = \left[\ y_i\langle e^i|\ \right]\ \left[\ |x\rangle\ \right] \tag{39}$$

with

$$y_i = g_{ij}y^j \tag{40}$$

and therefore

$$\langle y|x\rangle = y_i x^i. \tag{41}$$

This shows that we can use the metric to translate between vectors and covectors and that it is practically achieved, on the coefficient level, by lowering indices.

$$\langle x|y\rangle = x^i g_{ij} y^j = x^i y_i = x_j y^j. \tag{42}$$

We can equally define a metric for the dual space, which allows to raise indices, by

$$g^{ij}|e_i\rangle|e_j\rangle. \tag{43}$$

With this we can contract two covectors to get

$$\langle x|y\rangle = x_i g^{ij} y_j = x^j y_j = x_i y^i. \tag{44}$$

We can use this to raise or lower indices of linear transformations as well

$$g_{ki}a^i{}_j = a_{kj}, \tag{45}$$

or

$$a^i{}_j g^{jk} = a^{ik}. \tag{46}$$

We emphasize here again, that the order of the terms is irrelevant, since the epressions constitute explicit sums over numbers. From this we can rederive the above statements about swapping indices horizontally.

$$g_{ki}a^i{}_j g^{jl} = a_k{}^l. \tag{47}$$

We found earlier that a linear map A transforms as

$$A = B^{-1}\hat{A}B \tag{48}$$

and a martic G as

$$G = B^T\hat{G}B. \tag{49}$$

Lowering the first index of the linear map A leads to

$$GA = B^T\hat{G}BB^{-1}\hat{A}B = B^T\hat{G}\hat{A}B \tag{50}$$

Let's assume that, for a give basis system, the linear map A is symmetric

$$A = A^T. \tag{51}$$

After changing the basis, we find

$$\hat{A}^T = (BAB^{-1})^T = (B^{-1})^T A^T B^T = (B^{-1})^T AB^T \neq \hat{A} \tag{52}$$

Therefore, a linear map cannot be symmetric, since if it was symmetric in a given basis, it wouldn't be symmetric in another basis. Let's do the same analysis for GA, with

$$GA = (GA)^T = A^T G^T. \tag{53}$$

After changing the basis, we find

$$(\hat{G}\hat{A})^T = (BGAB^T)^T = BGAB^T = \hat{G}\hat{A}. \tag{54}$$

If GA was symmetric in one basis system, it would be in any. Hence it makes sense only to talk about symmetric matrices for matrices with two lower (or two upper) indices. This will be important, for instance, when we talk about symmetries of the generators of the Lorentz algebra. This symmetries have to be analyzed for generators with two lower (or two upper) indices, and not for generators representing the linear Lorentz map (with one upper and one lower index).

Let's assume now, that the linear map A leaves the contraction of vectors unchanged, i.e.

$$\langle x|y\rangle = \langle \hat{x}|\hat{y}\rangle. \tag{55}$$

It then follows

$$\langle \hat{x}|\hat{y}\rangle = g_{ab}\langle e_a|\langle e_b|a^c{}_d x^d|e_c\rangle a^e{}_f x^f|e_e\rangle \tag{56}$$

$$= a^a{}_d g_{ab} a^b{}_f x^d x^f. \tag{57}$$

Therefore,

$$a^a{}_d g_{ab} a^b{}_f = g_{df}\,. \tag{58}$$

Let's multiply this equation with A^{-1} from the right

$$a_{hd} = a_{bd} a^b{}_f (a^{-1})^f{}_h = g_{df}\,(a^{-1})^f{}_h = (a^{-1})_{dh}. \tag{59}$$

That means, that if A leaves the the angles between vectors unchanged, GA is an orthogonal matrix. Again, the characteristic of a matrix, here the orthogonality, is given for the matrix version with two lower indices.

3 Special Relativity

In special relativity, the metric has the well known, non trivial structure, defined by

$$\eta = \eta_{ij}\langle e^i|\langle e^j| \tag{60}$$

with

$$\eta_{ij} = 0, i \neq j \tag{61}$$

and

$$\eta_{00} = 1,\ \eta_{11} = -1,\ \eta_{22} = -1,\ \eta_{33} = -1, \tag{62}$$

where we used the symbol η instead of g to indicate that we are using an explicitly given metric now.

3.1 Lorentz Transformation

The Lorentz transformation L relates different observers to each other. It therefore translates between inertial systems. It is restricted to inertial systems with a common origin, since

$$L|0\rangle = |0\rangle. \tag{63}$$

The Lorentz transformation is defined to leave angles between and lengths of vectors invariant. With the above findings, L is defined such that ηL is orthogonal

$$(\eta L)^T = (\eta L)^{-1}. \tag{64}$$

3.2 Lie Algebra

The Lorentz transformation contains a subgroup consisting of rotations in space-time. Since roatations are continously linked to the identity element, the Lorentz transformation can be represented by a Lie Algebra

$$L = l^i{}_j |e_i\rangle\langle e^j| = e^{\omega^i{}_j} |e_i\rangle\langle e^j| \tag{65}$$

$$\approx (1+\omega)^i{}_j |e_i\rangle\langle e^j| = (\delta^i{}_j + \omega^i{}_j)|e_i\rangle\langle e^j| \tag{66}$$

With

$$l^a{}_d \eta_{ab} l^b{}_f = \eta_{df}\,, \tag{67}$$

it follows

$$\eta_{df} \approx (\delta^a{}_d + \omega^a{}_d)\eta_{ab}(\delta^b{}_f + \omega^b{}_f) \tag{68}$$

$$= (\delta^a{}_d + \omega^a{}_d)(\eta_{af} + \omega_{af}) \tag{69}$$

$$\approx \eta_{df} + \omega_{df} + \omega_{fd}. \tag{70}$$

Therefore

$$\omega_{df} = -\omega_{fd}. \tag{71}$$

As explained above, the symmetry considerations relate to matrices with two lower indices and not to the linear transformation directly. The generators relate to the linear Lorentz transformation as

$$\omega^d{}_f = \eta^{de}\omega_{ef}. \tag{72}$$

ω_{df} has, because of the symmetry, six degrees of freedom. This six degrees of freedom related to 3 rotations and 3 boosts.

The Lorentz group L and ω are related through

$$L^d{}_f = \exp(\omega^d{}_f). \tag{73}$$

A group is determined by the way it combines two elements via multiplications into one. This is most apparent for finite groups, which can be defined by a multiplication table alone. Hence, if we know how to multiply any two elements, we know the group. A group is therefore the abstract structure of the multiplication and anything that follows this type of multiplication is a group representation.

The question that naturally arise is: Given the basis matrices ω only, how could we work out the group multiplication of the Lorentz group. The answer is given by the the Baker–Campbell–Hausdorff formula

$$e^X e^Y = e^Z, \tag{74}$$

with

$$Z = X + Y + \frac{1}{2}[X,Y] + \frac{1}{12}[X,[X,Y]] - \frac{1}{12}[Y,[X,Y]] + \dots \tag{75}$$

and

$$[X,Y] = XY - YX. \tag{76}$$

This means, that to work out the group structure, we need to know the commutators of the Lie Algebra. Or in other words, two Lie algebras with the same commutator structure lead to isomorphic groups.

If we define the basis matrices to be

$$f_{ij} = \langle e_i|\langle e_j| - \langle e_j|\langle e_i| \tag{77}$$

with $j > i$. Because of the definition of f,

$$f^i{}_j \neq \eta^{il} f_{lj}\,. \tag{78}$$

Since we cannot use symmetry arguments for the linear map, we have to argue explicitly. Due to the metric, each

$$|e^i\rangle\langle e_j| \tag{79}$$

with $i \geq 2$ picks up a minus sign. Hence,

$$f^i{}_j = (f_m)^i{}_j = -|e^i\rangle\langle e_j| + |e^j\rangle\langle e_i|, \tag{80}$$

for $i \geq 2$ and $j > i$, and

$$f^1{}_j = (f_p)^1{}_j = |e^1\rangle\langle e_j| + |e^j\rangle\langle e_1|, \tag{81}$$

$j > 1$.

Let's work out the commutators of the basis matrics f_m and f_p. Let's take, for instance,

$$\left[(f_m)^2{}_3, (f_m)^2{}_4\right] = \tag{82}$$

$$= \left(-|e^2\rangle\langle e_3| + |e^3\rangle\langle e_2|\right)\left(-|e^2\rangle\langle e_4| + |e^4\rangle\langle e_2|\right) \tag{83}$$

$$-\big(\left(-|e^2\rangle\langle e_4| + |e^4\rangle\langle e_2|\right)\left(-|e^2\rangle\langle e_3| + |e^3\rangle\langle e_2|\right)\big) \tag{84}$$

$$= -|e^3\rangle\langle e_4| + |e^4\rangle\langle e_3| = (f_m)^3{}_4. \tag{85}$$

Similarly,

$$\left[(f_p)^1{}_3, (f_p)^1{}_4\right] = \tag{86}$$

$$= \left(|e^1\rangle\langle e_3| + |e^3\rangle\langle e_1|\right)\left(|e^1\rangle\langle e_4| + |e^4\rangle\langle e_1|\right) \tag{87}$$

$$-\big(\left(|e^1\rangle\langle e_4| + |e^4\rangle\langle e_1|\right)\left(|e^1\rangle\langle e_3| + |e^3\rangle\langle e_1|\right)\big) \tag{88}$$

$$= |e^3\rangle\langle e_4| - |e^4\rangle\langle e_3| = -(f_m)^3{}_4. \tag{89}$$

Let's take a look at the three dimensional rotation group, with a trivial metric

$$G = \delta_{ij}\langle e^i|\langle e^j| \tag{90}$$

The orthogonal map is defined through

$$O|x\rangle = |\hat{x}\rangle \tag{91}$$

with

$$\langle x|y\rangle = \langle\hat{x}|\hat{y}\rangle. \tag{92}$$

Hence

$$\langle\hat{x}|\hat{y}\rangle = \delta_{ab}\langle e_a|\langle e_b|o^c{}_d x^d|e_c\rangle o^e{}_f x^f|e_e\rangle \tag{93}$$

$$= o^a{}_d \delta_{ab} o^b{}_f x^d x^f. \tag{94}$$

It follows

$$o^a{}_d \delta_{ab} o^b{}_f = \delta_{df}\,. \tag{95}$$

The we find for an infinitesimal transformation

$$(\delta_{ad} + \omega^a{}_d)\delta_{ab}(\delta^b{}_f + \omega^b{}_f) \tag{96}$$

$$= (\delta_{bd} + \omega_{bd})(\delta^b{}_f + \omega^b{}_f) \tag{97}$$

$$= \delta_{fd} + \omega_{fd} + \omega_{df} \, = \delta_{df}\,. \tag{98}$$

Hence

$$\omega_{fd} = -\omega_{df}\,. \tag{99}$$

Due to the symmetry, the Lie algebra has three degrees of freedom. Let's define the basis as

$$f_{ij} \, = \langle e_i|\langle e_j| - \langle e_j|\langle e_i|. \tag{100}$$

Since the metric is trivial

$$f^i{}_j \, = |e^i\rangle\langle e_j| - |e^j\rangle\langle e_i|. \tag{101}$$

This is the same basis as for the f_m for the Lorentz Lie algebra. Hence the commutators will be identical too. Therefore the f_m matrices of the Lorentz Lie algebra correspond to rotations in space.

Let's rename the basis matrices and summarize the commutators,

$$J_1 = f^3{}_4\,, \quad J_2 = -f^2{}_4\,, \quad J_3 = f^2{}_3 \tag{102}$$

and

$$K_1 = f^1{}_2\,, \quad K_2 = f^1{}_3\,, \quad K_3 = f^1{}_4\,. \tag{103}$$

Then we find

$$[J_1, J_2] = J_3, \quad [J_2, J_3] = J_1, \quad [J_3, J_1] = J_2, \tag{104}$$

$$[K_1, K_2] = -J_3, \quad [K_2, K_3] = -J_1, \quad [K_3, K_1] = -J_2, \tag{105}$$

$$[J_1, K_2] = K_3, \quad [J_2, K_3] = K_1, \quad [J_3, K_1] = K_2, \tag{106}$$

$$[J_1, K_3] = -K_2, \quad [J_2, K_1] = -K_3, \quad [J_3, K_2] = -K_1, \tag{107}$$

If we change the basis yet again, this time to

$$A^+{}_i = J_i + iK_i, \quad A^-{}_i = J_i - iK_i. \tag{108}$$

The commutators read then

$$[A^+{}_1, A^+{}_2] = A^+{}_3, \quad [A^+{}_2, A^+{}_3] = A^+{}_1, \quad [A^+{}_3, A^+{}_1] = A^+{}_2, \tag{109}$$

$$[A^-{}_1, A^-{}_2] = A^-{}_3, \quad [A^-{}_2, A^-{}_3] = A^-{}_1, \quad [A^-{}_3, A^-{}_1] = A^-{}_2, \tag{110}$$

and

$$[A^+{}_1, A^-{}_2] = 0, \quad [A^+{}_2, A^-{}_3] = 0, \quad [A^+{}_3, A^-{}_1] = 0. \tag{111}$$

4 Quantum states

A quantum state is described by a wave function over space-time. Let's see how we can generate a function like this in the bra and ket notation. Let's start with a scalar function. A scalar can be computed with the dot product of two vectors. Let's therefore define

$$\langle x|y\rangle \tag{112}$$

at every spacetime point x. If we chose now y at every space-time point such that the dot product equals the value of the function we want to generate, then we have to write more precisely

$$\langle x|y(x)\rangle. \tag{113}$$

We'll rephrase this as follows

$$|\psi\rangle = |y(x)\rangle \tag{114}$$

and

$$\psi(x) = \langle x|y(x)\rangle = \langle x|\psi\rangle. \tag{115}$$

We can work out how this function changes under a Lorentz transformation. We can use its bracket definition to do so. A Lorentz transformation is defined to leave the dot product unchanged, hence

$$\langle \hat{x}|\hat{y}(\hat{x})\rangle = \langle x|y(x)\rangle, \tag{116}$$

or

$$\hat{\psi}(\hat{x}) = \psi(x) \tag{117}$$

or

$$\hat{\psi}(x) = \psi(L^{-1}x). \tag{118}$$

In this notation, the state $|\psi\rangle$ represents a vector at each space-time point, which will form a dot product to create a scalar. The lorentz transformation transforms this vectors for each space-time point and like this induces a transformation on the state space

$$\psi \to \hat{\psi} \tag{119}$$

or

$$\hat{\psi} = U(\psi). \tag{120}$$

The state vectors form a new vector space. A basis system can be thought of as the direct sum of a coordinate system at every space-time point. We will just enumerate this basis vectors, so that we can write again U is then statisfies

$$|\psi\rangle = \psi^i |e_i\rangle, \tag{121}$$

only that now, we have an infinite sum. U is then give by

$$\hat{\psi}^i = u^i{}_j \psi^j \tag{122}$$

Since the dot product at each space-time point leaves the dot product unchanged and since the Lorentz transformation induced the map U, U will leave the dot product, defined in its vector space, invariant as well. Since we are now dealing with quantum states, we have to allow for multiplications with complex numbers as well. Hence a transition between vectors and covectors will include a complex conjugation, to keep the result of the dot product real. Therefore,

$$\eta_{il} u^l{}_j = u_{ij} \tag{123}$$

is unitary (since, as described above, symmetries don't apply to linear maps, as they are not independent of the choice of basis).

For a vector field, we just have to take a vector at each space-time point

$$|y(x)\rangle = |\psi\rangle. \tag{124}$$

We can still call this $|\psi\rangle$, or the wave function $\psi(x)$, like it is mostly done for the Dirac field. Under a Lorentz transformation, this transforms as

$$|\hat{y}(\hat{x})\rangle = |\hat{y}(Lx)\rangle = L|y(Lx)\rangle \tag{125}$$

or

$$|\hat{y}(x)\rangle = L|y(L^{-1}x)\rangle \tag{126}$$

Let's look an intuitive explanation for this transformation. y is a vector that sits at the end of vector x. If we write

$$|z\rangle = |x\rangle + |y\rangle, \tag{127}$$

then both, $|z\rangle$ and $|x\rangle$ start at the origin of the coordinate system and therefore clearly transform as $L|z\rangle$ and $L|x\rangle$. Therefore,

$$|z\rangle - |x\rangle = |y\rangle, \tag{128}$$

$$L(|z\rangle - |x\rangle) = L|y\rangle, \tag{129}$$

$$L|z\rangle = L|x\rangle + L|y\rangle, \tag{130}$$

Hence, both $|x\rangle$ and $|y\rangle$ have to undergo a Lorentz transformation. This is the reason why we find two Lorentz transformations in the transformation law of vector fields.

If we were to describe an elementary particle with this, then the external degrees of freedom would be its position and location. Everything else is a internal. Hence, if we used a vector field to describe a particle, the vector would describe an internal degree of freedom. The internal degrees of freedom should be independent of the observer and invariant under Lorentz transformations. Let's put a particle, described by a vector field, at the origin of a coordinate system, which is therefore equal to zero everywhere, but at the origin. Therefore we only have to consider the vector at the origin

$$|\psi(0)\rangle. \tag{131}$$

This transforms as

$$|\psi(\hat{0})\rangle = |\psi(0)\rangle, \tag{132}$$

since $L(0) = 0$. The remaining transformation describes internal degrees of freedom. We postulate that states that describe elementary particles to be eigenstates of L. Hence, we have to find a wave function such that

$$L|\psi(0)\rangle = s|\psi(0)\rangle, \tag{133}$$

where L has to only be applied at the 4 vector at the origin. We already showed before, that for L, a basis can be chosen, so that L acts as $su(2) \bigoplus su(2)$, where the direct sum $\bigoplus$ refers to the fact, that we have two mutually exclusive subspaces. We therefore found a group representation of the form $(\frac{1}{2}, 0) \bigoplus (0, \frac{1}{2})$. Therefore, a vector particle corresponds to a Dirac field, staisfying the Dirac equation.

Since the Lorentz group splits up in 2 invariant subspaces, we can place vectors in multiples of 2 at every space-time point, such that the 2d vector space is left invariant by the Lorentz transformation. Let's therefore write for a vector field

$$|x\rangle|s\rangle. \tag{134}$$

This can be seen as the vector $|s\rangle$ sitting at the end of each $|x\rangle$. $|s\rangle$ can have any diemension a representation of $su(2)$ can act on. Or it could be a tensor product of states as well, for instance

$$|x\rangle|s_1\rangle|s_2\rangle. \tag{135}$$

We have to rewrite the transformation of such a vector field, since it is now a transformation induced by the Lorentz transformation acting on $|s\rangle$,

$$|\hat{y}(\hat{x})\rangle = L|y(Lx)\rangle \to |\hat{x}\rangle|\hat{s}\rangle = L|x\rangle U|s\rangle. \tag{136}$$

L and U describe the same transformation, but couple to different vector spaces.

For this analysis of the inner symmetries we assumed the particle to be at rest at the origin of the coordinate system. This is not possible for massless particles. A state, of length 0, that describes a massless particle, moving in the z direction can be given

$$|p\rangle = k|e_1\rangle + k|e_4\rangle. \tag{137}$$

We define again inner symmetries to be described by all Lorentz transformations that leave this state invariant. For this to be true, the generators of the Lorentz group have to vanish when acting on the state. Let's look at all generator's action on the state:

$$J_1|p\rangle = -|e_3\rangle \quad J_2|p\rangle = |e_2\rangle \quad J_3|p\rangle = 0 \tag{138}$$

$$K_1|p\rangle = |e_2\rangle \quad K_2|p\rangle = |e_3\rangle \quad K_3|p\rangle = |e_1\rangle + |e_4\rangle. \tag{139}$$

Therefore, the generators we are looking for are given by

$$J_3, \quad J_1 + K_2, \quad J_2 - K_1. \tag{140}$$

Let's define

$$P_1 = J_1 + K_2, \quad P_2 = J_2 - K_1, \tag{141}$$

and look at the commutators

$$[J_3, P_1] = P_2, \quad [J_3, P_2] = -P_1, \quad [P_1, P_2] = 0. \tag{142}$$

5 Representation of $SO(3)$

Since the bra and ket already indicate whether we are dealing with a vector or a covector, we will make notation simpler by replacing

$$|e_i\rangle \to |i\rangle, \quad \langle e^i| \to \langle i|. \tag{143}$$

$SO(3)$ is the rotation group in 3 dimension, that we encountered already as a subgroup of the Lorentz group. Its generators read

$$J_1 = |2\rangle\langle 3| - |3\rangle\langle 2|, \quad J_2 = -|1\rangle\langle 3| + |3\rangle\langle 1|, \quad J_3 = |1\rangle\langle 2| - |2\rangle\langle 1|. \tag{144}$$

The commutation rules are given by

$$[J_1, J_2] = J_3, \quad [J_2, J_3] = J_1, \quad [J_3, J_1] = J_2, \tag{145}$$

where $|i\rangle$ and $\langle i|$ are the basis vectors and covectors, respectively. Since the commutators do not commute, they can't be diagonalized simultaniously. Let's therefore diagonalize J_3. Since

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} = \begin{pmatrix} -i & 1 & 0 \\ 0 & 0 & \sqrt{2}i \\ i & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & i & 0 \\ -i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{i}{2} & 0 & -\frac{i}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & -\frac{i}{\sqrt{2}} & 0 \end{pmatrix} \tag{146}$$

$$\begin{pmatrix} 0 & -\frac{i}{\sqrt{2}} & 0 \\ -\frac{i}{\sqrt{2}} & 0 & \frac{i}{\sqrt{2}} \\ 0 & -\frac{i}{\sqrt{2}} & 0 \end{pmatrix} = \begin{pmatrix} -i & 1 & 0 \\ 0 & 0 & \sqrt{2}i \\ i & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{i}{2} & 0 & -\frac{i}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & -\frac{i}{\sqrt{2}} & 0 \end{pmatrix} \tag{147}$$

$$\begin{pmatrix} 0 & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & 0 \end{pmatrix} = \begin{pmatrix} -i & 1 & 0 \\ 0 & 0 & \sqrt{2}i \\ i & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & i \\ 0 & -i & 0 \end{pmatrix} \begin{pmatrix} \frac{i}{2} & 0 & -\frac{i}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & -\frac{i}{\sqrt{2}} & 0 \end{pmatrix} \tag{148}$$

Since we want real eigenvalues, we redefine the generators as

$$L_i = iJ_i. \tag{149}$$

The new basis vectors are simply the columns of

$$\begin{pmatrix} \frac{i}{2} & 0 & -\frac{i}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & -\frac{i}{\sqrt{2}} & 0 \end{pmatrix}. \tag{150}$$

The matrix entry $(3,2)$, given by $-\frac{i}{\sqrt{2}}$, is chosen, so that the ladder operators look nice and symmetric. Any other factor would be possible. We find,

$$L_3(i|1\rangle + |2\rangle) = i|1\rangle + |2\rangle, \tag{151}$$

$$L_3(-i|1\rangle + |2\rangle) = i|1\rangle - |2\rangle = -(i|1\rangle + |2\rangle), \tag{152}$$

$$L_3(-i|3\rangle) = 0, \tag{153}$$

and

$$|\hat{1}\rangle = i|1\rangle + |2\rangle, \quad |\hat{2}\rangle = -i|3\rangle. \quad |\hat{3}\rangle = -i|1\rangle + |2\rangle. \tag{154}$$

We can define ladder operators as

$$L_+ = (L_1 - iL_2) = \sqrt{2}\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \tag{155}$$

$$L_- = (L_1 + iL_2) = \sqrt{2}\begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \tag{156}$$

and therefore,

$$L_+|\hat{3}\rangle = |\hat{2}\rangle \quad L_+|\hat{2}\rangle = |\hat{1}\rangle \quad L_+|\hat{1}\rangle = 0, \tag{157}$$

$$L_-|\hat{1}\rangle = |\hat{2}\rangle \quad L_-|\hat{2}\rangle = |\hat{3}\rangle \quad L_-|\hat{3}\rangle = 0. \tag{158}$$

Let's define J^2 as

$$J^2 = {J_1}^2 + {J_2}^2 + {J_3}^2. \tag{159}$$

Therefore,

$$J^2 = \begin{pmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 1 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix} + \begin{pmatrix} \frac{1}{2} & 0 & -\frac{1}{2} \\ 0 & 1 & 0 \\ -\frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix} + \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{160}$$

$$= \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} = \begin{pmatrix} j(j+1) & 0 & 0 \\ 0 & j(j+1) & 0 \\ 0 & 0 & j(j+1), \end{pmatrix} \tag{161}$$

with $j = 1$.

The fact that we had to introduce complex numbers to get to real eigenvalues, indicates that there exists a more fundamental algebra, which is directly defined within complex numbers.

6 Representations of $su(2)$

An $su(2)$ matrix leaves a complex dot product invariant. For transformations between complex vector and covector spaces, the coeffiecient undergo conjugation. Hence, with

$$U|x\rangle = |\hat{x}\rangle \tag{162}$$

and

$$\langle x|y\rangle = \langle\hat{x}|\hat{y}\rangle, \tag{163}$$

it follows,

$$\langle\hat{x}|\hat{y}\rangle = \delta_{ab}\langle e_a|\langle e_b|{u^c}_d x^d|e_c\rangle {u^e}_f x^f|e_e\rangle \tag{164}$$

$$= o^{a}{}_{d}\delta_{ab}o^{b}{}_{f}x^{d}x^{f}. \tag{165}$$

It follows

$$o^{a}{}_{d}\delta_{ab}o^{b}{}_{f} = \delta_{df}\,. \tag{166}$$

References

[1] Author, *Title*, Journal/Editor, (year)