

# Useful Notes for the Lorentz Group

## O. Preliminary Comments about linear transformations of vector spaces

We study vectorial quantities, such as velocity, momentum, force, etc. We need to be able to combine different velocities, and also other systems of vectors of the same type; therefore, we visualize them as being elements of vector spaces. We can describe these vectors, for instance, via sets of numbers, which we think of as the components of these vectors relative to some choice of basis vectors. One can then suppose that it is desirable to make a change in this choice of basis vectors, which would then create a transformation of the components of the vectors. Such a change does not, of course, change the underlying physics of the situation.

It is then important to understand sets of such transformations, which we can visualize in terms of (sets of) matrices acting on the vector spaces. In these notes we will consider a very important such set of transformations, namely the entire Lorentz group, which describe changes of basis corresponding to different, allowed inertial observers. This set contains pure rotations, pure Lorentz boosts, i.e., changes of observers, moving with distinct velocities, and also products of such transformations. Physically, the insistence on products is just a way of saying we may want to make more than one of these transformations at a time, in order to create a desirable set of basis vectors.

We will begin by simply describing the rotations as they are easily perceived in 3-dimensional space, before we proceed onward to the entire (4-dimensional) spacetime, embedding the rotations into spacetime rather easily, and then looking at other transformations in that 4-dimensional situation. Each of these sets of transformations is described by a set of parameters that are allowed to vary continuously—and smoothly, remembering that smoothly means arbitrarily often differentiable—from one value to another of the parameters, in some appropriate parameter space. Such an entire set of transformations can be viewed in two quite different ways, which must be required to be consistent.

- a. Since there are many of these transformations it is important that products of them also be transformations. Since one could surely always not bother to perform a transformation

at all, an “identity” transformation must be included in the set, which does nothing to the vectors in question. Lastly, since there is an identity there must be a way to move backward toward doing nothing, after having done something; this is the same as saying that there must be an inverse transformation to each and every transformation we consider. A set of objects which satisfies the above notions, and which also has a product which is associative, is called a group. That our transformations are associative follows from the fact that they are transformations of elements of a vector space, of course preserving that structure, and therefore linear, which makes them associative.

- b. However, since the parameters are required to vary in a continuous manner, one can conceive of the complete parameter space as a manifold with the parameters as (local) coordinates on it. Therefore we have a manifold which is also a group. Such groups are referred to as *Lie groups*, named for Sophus Lie who first studied them extensively. Because some of the parameters may not vary continuously, an example being the determinant of a matrix, then it is often the case that the manifold in question comes in more than one connected part.

Although it could be that one is studying such transformations acting in a single vector space, it is usually the case that there is an underlying physical manifold, i.e., the place where “we live,” and there are several distinct vector spaces at each and every point on that manifold, which in these notes we will refer to as spacetime, which has a local set of Cartesian coordinates,  $\{x, y, z, t\} \equiv \{\vec{r}, t\}$  defined on it, among other possible choices for coordinates there. There are different sorts of vector spaces at each point, to accommodate different sorts of tensors that one wants to study and/or measure at each coordinate point in spacetime, such as one for 4-vectors, one for type  $[0,2]$  tensors, etc. For any particular sort of vector space the set of all of them together, over different spacetime points in some neighborhood is referred to as a *vector bundle*.

Since these transformations are linear transformations in the vector spaces defined at each point of the spacetime manifold, just as we traditionally describe vectors, and other tensors, in

terms of their components, relative to a given, well-understood, fixed choice of a (complete) set of basis vectors, we will also want to describe these transformations in terms of the components of more complicated tensors which act on those components. The traditional way to do that is via presentations of those components in terms of matrices which then act on the matrices containing the components of the vectors, all of course still relative to a given, fixed choice of basis set for the vector space(s) in question. Before we will be able to do that, we must distinguish between two rather different ways to present transformations of vectors in a vector space, and make a choice as to which mode we will use here: the so-called “passive” and “active” modes of thinking about them. To describe these two modes, we begin with the picture of a vector space along with a set of basis vectors, and some particular vector,  $\vec{A}$ .

- a. In the “*passive*” *approach* we hold the vector fixed, and rotate the set of axes. This makes sense from the point of view that the vectors are actually not changed simply because someone has chosen to pick some different set of basis vectors, i.e., ones that have been rotated relative to the original ones.

This is the most commonly-used approach in classical physics, especially mechanics.

- b. In the “*active*” *approach*, on the other hand, the vector is considered the physical quantity of interest, and therefore it is the object that should be rotated. The basis set is actually not physical, after all, and we shouldn’t waste our time rotating them.

This is the most common approach in quantum mechanics, where it is the “state vectors” on which attention is concentrated. It could also be used to describe measurements of a single vector made by two distinct allowed observers.

Either approach describes the rotations, in a very generic sense, in the same way. However, it should be clear that any particular rotation in one mode is the inverse of that rotation viewed in the other mode. To make that even more clear, suppose we rotate the vector and also rotate the basis vectors, then nothing will have really changed in the sense that the components are still the same as before. In any event, when making explicit analytic calculations, a decision must be made as to which approach is to be used, and, of course, only that choice must then be

followed. In these notes I intend to follow the passive approach, leaving the vectors unchanged, but allowing one to choose many different allowed sets of basis vectors, related to each other in various ways, and then considering the resulting changes of the components of vectors, when they are presented in the form, usually, of column vectors, i.e., matrices with only one column. This will arrange it so that the group elements, i.e., the transformation operators will be presented as square matrices.

## I. The Rotation Group, in 3-dimensional space

In a 3-dimensional (vector) space with standard (totally spatial) metric,  $\mathbb{E}^3$ , **rotations** are defined as (linear) transformations that preserve the lengths of vectors, while changing their directions. We will use the symbol  $R$  for the transformation itself, as it acts on the vectors  $\vec{A}$ , and also for the matrix that acts on the column vectors  $A$  that present the components of  $\vec{A}$ :

$$\begin{aligned} \forall \vec{A}, \vec{B} \in \mathbb{E}^3, \quad (R\vec{A}) \cdot (R\vec{B}) = \vec{A} \cdot \vec{B} \quad \implies \quad (RA)^T(RB) = A^T B = A^T \mathbf{I}_3 B \\ \iff \quad R^T R = \mathbf{I}_3, \end{aligned} \tag{1.1}$$

where  $\mathbf{I}_3$  is the  $3 \times 3$  identity matrix, and the second line follows because the identity was true for arbitrary vectors  $A$  and  $B$ , so that this line may be taken as the defining matrix relation for a rotation. As already noted above the set of all rotations forms both a group and a manifold, so that it is a Lie group. However, Eq. (1.1) allows us to imply that the determinant of  $R$  has value  $\pm 1$ , with no possible values in between, so that the set of all rotations must actually come in **two separate parts**, so that the manifold in question is *not connected*. A basic element from each part is the identity,  $\mathbf{I}_3$ , and its negative,  $-\mathbf{I}_3$ , this last one often called the *parity* transformation. The group of all rotations, in 3-dimensional space is often denoted by the symbol  $\mathbf{O}(3)$ , where the  $\mathbf{O}$  stands for the word “orthogonal,” a reminder that when this group of transformations acts on the basis set it maintains the property that they had when chosen, of being orthonormal, i.e., orthogonal to each other and “normal,” which implies that they had unit length.

The subset with determinant  $+1$  is, separately, a group, since the product of two  $+1$ 's is also  $+1$ ; therefore, it constitutes a *subgroup* of all rotations, called  $\mathbf{SO}(3)$ , where the  $S$  stands for the word “special.” It is also usual to refer to the elements of this subgroup as *proper rotations*; this language means that one excludes those with negative determinant. Of course the set with determinant  $-1$  is not a subgroup, i.e., a group all by itself, since the product of two of them lies in the other part.

A common choice of parameters to specify a rotation can be found by saying that we “rotate through some angle  $\theta$ , about some axis,  $\hat{n}$ ,” giving rise to the notation  $R(\theta, \hat{n})$ , where we use the (standard) “right-hand rule” to define the sign of the angle in question. That is, when the thumb of the right hand points in the direction of  $\hat{n}$ , then as the fingers curl closed they are moving in the direction of positive values of  $\theta$ . Notice then that the rotation in question is invariant when one both reverses the direction of  $\hat{n}$  and switches the sign of  $\theta$ . The axis in question is of course only a direction, and so has only two degrees of freedom, which we could specify, for instance, by giving its spherical coordinates, say  $(\vartheta, \varphi)$ , so that we see that three degrees of freedom are necessary to specify a rotation. (Since the direction is most easily specified by a unit vector, some people also use the notation  $\vec{\theta} \equiv \theta \hat{n}$  to specify a particular direction.

One might normally have thought of the angle of rotation as some positive value between  $0$  and  $2\pi$ . However, there is an approach that creates a better “visualization” of the manifold of all possible rotations, with positive determinant, in other words the manifold already named as  $\mathbf{SO}(3)$ . We surely can agree that it is an equivalent description of such an angle by subtracting off  $\pi$ , so that the range for  $\theta$  varies, instead, from  $-\pi$  to  $+\pi$ . Then, for those values that are positive, with  $\hat{n}$  taking on all possible directions, we may treat the value of  $\theta$  as the length of a vector pointing in the direction  $\hat{n}$  away from the origin, the set of all of these rotations will fill an entire *ball* of radius  $\pi$ . [Note that the word “ball” is intended to mean a “solid sphere,” i.e., a sphere along with its interior.] This would appear to be troublesome since we have now ignored all those angles which are negative. However if one changes both the sign of the angle and

the sign of the axis of rotation, this amounts to no change at all, i.e.,  $R(\theta, \hat{n}) = R(-\hat{n}; -\theta)$ . This is a “counting problem” with this parametrization where one actually counts all the rotations twice; therefore, if  $\theta$  is negative, this correspondence gives us a different rotation with a positive angle, in the opposite direction. By only including angles from 0 to  $\pi$  and allowing all possible directions for  $\hat{n}$  we are in fact only counting each rotation once, **except** for a separate duplication problem when one rotates half way around, i.e., when the value for the angle is  $\theta = \pi$ . It is clear that we have an additional double counting problem in this instance because  $R(\hat{n}; \pi) = R(\hat{n}; -\pi)$ , i.e., both rotations bring us to the same place, so that they are in fact completely equivalent. These two rotations of course occur on the boundary of our ball; in fact they are best seen as at the opposite ends of any arbitrary diameter of our ball, i.e., any straight line through the center. (I note that the pair of points at opposite ends of a diameter are usually referred to as *antipodal points*.) Therefore our manifold for the rotation group, which seemed fairly normal so far, is in fact a “little bit strange, since the apparently two points at opposite ends of any diameter, in fact constitute only one point, so that the manifold is in fact a curved space.

Because of this particular problem with the manifold it is sufficiently curved that it is not possible to view it in its entirety in a straightforward way in the 3-dimensional space in which we live, unless one simply says that every pair of antipodal points must be “identified” as the same point. Such a manifold is referred to as “*doubly connected*,” which has the following meaning. There are two distinct closed paths from any point, and back again, that cannot be deformed into each other. As examples, consider a circular path that goes from the origin (of the group, or of the ball, since those are the same point on the manifold) around on a circle of diameter less than  $\pi$ , say, perhaps,  $\pi/2$ . By slowly changing the diameter from this value to zero, in a continuous way, we deform that circular path into one which did not move at all. On the other hand, now consider a path that goes, up along the  $\hat{z}$ -axis, say, all the way to the boundary of the ball, at the North pole. However, the North pole is exactly the same point as the South pole, because of our identification of antipodal points. Therefore, now the path

proceeds from the South pole back along the  $\hat{z}$ -axis to the origin again. This is obviously a closed path; however, if we try to change the radius to become smaller we no longer have this identification of points so that such a change cannot be made. This is the second closed path, not deformable to the first one.

Having created a “visualization” of the manifold  $\mathbf{SO}(3)$ , our next step is to define explicit presentations, using  $3 \times 3$  matrices, for the action of these rotations,  $R(\theta, \hat{n})$ , in our vector space. We remember that it has been agreed that we will use the passive approach to this, where it is actually the basis set that is being rotated, and we ask what changes this creates in the components of the vector, which has remained fixed during the rotation. Pretending that we have inserted here a figure, suppose we have a 2-dimensional plane, with the usual  $\{\hat{x}, \hat{y}\}$ -axes, and a vector  $\vec{A}$  drawn in the plane, in the first quadrant, making some positive angle  $\alpha$  with the  $\hat{x}$ -axis which is less than  $\pi/2$ . If I now rotate—about the  $\hat{z}$ -axis—the axes shown in this figure by some positive angle, since the vector is in the first quadrant, the angle  $\alpha$  will decrease, as the  $\hat{x}$ -axis approaches the location of  $\vec{A}$ . At the same time, the complement of that angle, which is the angle with respect to the  $\hat{y}$ -axis, originally  $\pi/2 - \alpha$ , will increase. Choosing the vector  $\vec{A}$  to have unit length, for simplicity, this allows us to see that the correct matrix presentation for such a transformation would be the following:

$$\begin{aligned} \vec{A} \implies A &= \begin{pmatrix} \cos \alpha \\ \sin \alpha \\ 0 \end{pmatrix}, & R(\theta, \hat{z}) &\implies \begin{pmatrix} \cos \theta & +\sin \theta & 0 \\ -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix}, \\ \implies RA &= \begin{pmatrix} \cos \theta & +\sin \theta & 0 \\ -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos \alpha \\ \sin \alpha \\ 0 \end{pmatrix} & (1.2) \\ &= \begin{pmatrix} \cos \alpha \cos \theta + \sin \alpha \sin \theta \\ -\cos \alpha \sin \theta + \sin \alpha \cos \theta \\ 0 \end{pmatrix} = \begin{pmatrix} \cos(\alpha - \theta) \\ \sin(\alpha - \theta) \\ 0 \end{pmatrix}. \end{aligned}$$

The expression above was very straightforward, except possibly for where the minus sign goes in the matrix. It was by way of determining that sign that this discussion was given in such detail, as well as the discussion relative to passive and active modes of presentation for the transformations.

At this point it is not difficult to generalize, and obtain the matrix form of the following simple rotations, around each of the three coordinate axes:

$$\begin{aligned}
R(\hat{z}; \theta) &\implies \begin{pmatrix} \cos \theta & \sin \theta & 0 \\ -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix}, \\
R(\hat{y}; \varphi) &\implies \begin{pmatrix} \cos \varphi & 0 & -\sin \varphi \\ 0 & 1 & 0 \\ \sin \varphi & 0 & \cos \varphi \end{pmatrix}, \\
R(\hat{x}; \zeta) &\implies \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \zeta & \sin \zeta \\ 0 & -\sin \zeta & \cos \zeta \end{pmatrix}.
\end{aligned} \tag{1.3}$$

Before looking at more general cases, it is useful to write out the lowest order expansions, for very small angles, for each of these three cases:

$$\begin{aligned}
R(\hat{z}; \theta) &\implies I_3 - \theta \begin{pmatrix} 0 & -1 & 0 \\ +1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + O^2(\theta) \equiv I_3 - \theta \mathcal{J}_z + O^2(\theta), \\
R(\hat{y}; \varphi) &\implies I_3 - \varphi \begin{pmatrix} 0 & 0 & +1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix} + O^2(\varphi) \equiv I_3 - \varphi \mathcal{J}_y + O^2(\varphi), \\
R(\hat{x}; \zeta) &\implies I_3 - \zeta \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & +1 & 0 \end{pmatrix} + O^2(\zeta) \equiv I_3 - \zeta \mathcal{J}_x + O^2(\zeta),
\end{aligned} \tag{1.4}$$

where the “big- $O$ ” symbols indicate the existence of additional terms of second-order or higher in the various angles. The equations define the three *infinitesimal generators* of these three (basic) rotations, namely  $\{\mathcal{J}_x, \mathcal{J}_y, \mathcal{J}_z\}$ . They are skew-symmetric matrices, as will be shown to be necessary, in the next section, for generators of group elements which have determinant +1, and their signs have been chosen so that their commutation relations have a form which is both reasonable and traditional. We display below a generic form of writing out their matrix elements, where the notation  $\mathcal{J}_i$  with  $i$  taking on the values, 1, 2, or 3, denotes the subscript  $x$ ,  $y$ , or  $z$ , respectively, and we enclose the symbol for a matrix within double parentheses and then label the elements of that matrix outside those parentheses, so that the first index is the row index and the second index the column index, independent of whether they are upper or



lower:

$$\begin{aligned}
((\mathcal{J}_i))^j{}_k &\implies -\epsilon_{ijk} , \\
[\mathcal{J}_i, \mathcal{J}_j] &= \epsilon_{ij}{}^k \mathcal{J}_k , \\
\vec{\mathcal{J}} &\equiv \mathcal{J}_x \hat{x} + \mathcal{J}_y \hat{y} + \mathcal{J}_z \hat{z} ,
\end{aligned} \tag{1.5}$$

where the  $\epsilon_{ijk}$  are the usual, completely skew-symmetric Levi-Civita symbols, that generate determinants of square matrices, and we have used the Einstein summation convention for repeated, one upper and one lower, indices. I have also taken this opportunity to define a more-complicated object that is quite useful, namely a 3-dimensional vector,  $\vec{\mathcal{J}}$ , with components that are matrices rather than the more usual scalars. Treating this as a vector allows me to consider scalar products of it with more standard sorts of vectors. For instance, the quantity  $\hat{n} \cdot \vec{\mathcal{J}}$  is a single matrix which is the sum of the 3 matrices involved multiplied by the 3 components of  $\hat{n}$ . It is then this matrix that can be used to generate rotations in the direction  $\hat{n}$ . Lastly I should note that it is customary, in quantum mechanical descriptions of rotations to desire that your important, physical operators be Hermitean, while these are skew-symmetric, but of course real-valued as they certainly should be. Therefore, when studying rotations in quantum mechanics, one introduces associated Hermitean (angular momentum) matrices,  $\{J_x, J_y, J_z\}$ , where  $\mathcal{J}_i = -i J_i$ . However, we will not concern ourselves with these quantities further here.

The objects we have called generators for the rotations are seen above to be the matrices that are the first term, after the identity matrix, in the power series expansion for the group element, when we expand it in a power series in a relevant parameter that would be small when the group element is near the identity element. The standard method for creation of the group element from that first, non-trivial term is by use of the exponential function, ; for example

$$R(\theta, \hat{z}) = e^{-\theta \mathcal{J}_z} = e^{-\theta \hat{z} \cdot \vec{\mathcal{J}}} , \tag{1.6}$$

where the exponential function is defined by its infinite power series.

As a useful aside, because of the relationship of determinants and cross products of 3-dimensional vectors, it is straightforward to see that the matrices  $\mathcal{J}_i$  generate the usual cross

product, when we treat them as components of a 3-dimensional vector, and consider their action on the components of a vector  $\vec{r} = r^a \hat{e}_a$

$$\left(\vec{v} \cdot \vec{\mathcal{J}}\right) \vec{r} = \left\{v^a (\mathcal{J}_a)^b{}_c r^c\right\} \hat{e}_b = \left\{-v^a \epsilon_{abc} r^c\right\} \hat{e}^b = \left\{\epsilon_{acb} v^a r^c\right\} \hat{e}^b = \vec{v} \times \vec{r}, \quad (1.7)$$

where the  $\{\hat{e}_b | b = 1, 2, 3\}$  constitute a Cartesian basis for the (3-dimensional) space being studied, i.e., they are different labels for the set  $\{\hat{x}, \hat{y}, \hat{z}\}$ .

I now want to go ahead and consider completely general rotation matrices that are elements of  $\mathbf{SO}(3)$ , i.e., the component of the manifold that contains the identity,  $\mathbf{I}_3$ , rather than simply those which rotate about the 3 basis vectors. We again consider them in terms of their generators by looking at small angles. Since the identity is a rotation by exactly zero angle, we can “measure” how much an arbitrary rotation,  $R$  deviates from the identity in terms of its logarithm,  $Q$ , by first writing

$$R = e^Q, \quad Q \equiv \log[I_3 + (R - I_3)] = \sum_{s=1}^{\infty} \frac{(-1)^{s+1}}{s} (R - I_3)^s, \quad (1.8)$$

which will converge for  $R$  “sufficiently near” the identity, which is the same as saying that the elements of  $R - I_3$  should be “sufficiently small.” We may then take the defining requirement, Eq. (1.1), that  $R$  be a rotation, and find its counterpart for the logarithm  $Q$ , as follows:

$$\begin{aligned} e^0 = I_3 = R^T R &= (e^Q)^T e^Q = e^{Q^T} e^Q = e^{Q^T + Q} \\ \implies 0 &= Q^T + Q, \end{aligned} \quad (1.9)$$

which is the same as saying that  $Q$  must be a skew-symmetric matrix  $\iff R$  is an orthogonal matrix, i.e., a rotation, near the identity. Recall that this property was promised above, when we chose the generators to be skew-symmetric matrices for the special cases that constituted rotations around the three principal axes. Having decided that, one should ask the range of convergence, in the group manifold, of the power series for the logarithm given above. As it turns out, the series defining the logarithm happens to converge for all **proper** rotations.

However, this is certainly not true for general Lie groups where the range of convergence is usually smaller than the entire, connected group, and simply requires fairly sophisticated calculations to determine the range of convergence.

In general, in  $n$  dimensional space, a skew-symmetric matrix has  $\frac{1}{2}n(n-1)$  degrees of freedom. In our 3-dimensional space, this tells us that there are exactly 3 independent degrees of freedom to skew-symmetric matrices, so that it is also true that

- a. again, we have shown that the group of all (3-dimensional) proper rotations has exactly 3 degrees of freedom.
- b. The 3 skew-symmetric matrices,  $\mathcal{J}_i$ , already presented in Eq. (1.3), constitute a choice of basis for the vector space of all  $Q$ 's; i.e. any skew-symmetric  $3 \times 3$  matrix is a linear combination of those 3 matrices, say  $a\mathcal{J}_x + b\mathcal{J}_y + c\mathcal{J}_z \equiv \vec{\theta} \cdot \vec{\mathcal{J}}$ . The 3 scalars  $\{a, b, c\}$  may then be thought of as the components of a vector  $\vec{\theta}$ , which is a choice of parametrization for all proper rotations, as we already briefly mentioned. However, it seemed simpler to us, above, and more convenient to separate out the direction and magnitude of  $\vec{\theta} \equiv \theta \hat{n}$ , so that  $Q = \theta \hat{n} \cdot \vec{\mathcal{J}}$ .
- c. The vector space which is the set of all possible skew-symmetric matrices  $Q$  is referred to as the *Lie algebra* for the rotation group. One may think of it as the derivatives, with respect to the set of parameters, of group elements which are then evaluated at the identity element. From that point of view the Lie algebra is the tangent vector space, i.e., the set of all directional derivatives, at the identity of the group manifold. In general, for any Lie group, this tangent vector space at the identity is its associated Lie algebra.

It is reasonably straightforward to see that  $\hat{n}$  is an eigenvector of the rotation it describes, since it is the axis about which the rotation is being described, and that is unchanged by such a rotation. On the other hand, it can also be seen more explicitly by the following calculation of the form of an arbitrary (proper) rotation, which we determine by the calculation of the exponential:

$$R(\theta, \hat{n}) = e^{-\theta \hat{n} \cdot \vec{\mathcal{J}}} = \cos \theta I_3 + (1 - \cos \theta) \hat{n} \hat{n}^T - \sin \theta \hat{n} \cdot \vec{\mathcal{J}}, \quad (1.10)$$

or

$$R(\theta, \hat{n};) \vec{V} = \cos \theta \vec{V} + (1 - \cos \theta) (\hat{n} \cdot \vec{V}) \hat{n} - \sin \theta \hat{n} \times \vec{V}.$$

The exponential has been evaluated by an iteration procedure rather than by an explicit calculation of the matrices. This is found by beginning with the vector cross product in Eq. (1.7), which is of course the action of the first term after the identity; however, that then allows us to iterate that step to calculate the second, and third powers:

$$\begin{aligned}
(\hat{n} \cdot \vec{\mathcal{J}})^2 \vec{V} &= \hat{n} \times (\hat{n} \times \vec{V}) = \hat{n}(\hat{n} \cdot \vec{V}) - \vec{V} \implies [\hat{n}\hat{n}^T - \mathbf{I}_3] \vec{V}; \\
\implies (\hat{n} \cdot \vec{\mathcal{J}})^3 \vec{V} &= \hat{n} \times (\hat{n}(\hat{n} \cdot \vec{V}) - \vec{V}) = -(\hat{n} \times \vec{V}) \\
&\implies (\hat{n} \cdot \vec{\mathcal{J}})^3 = -(\hat{n} \cdot \vec{\mathcal{J}}),
\end{aligned} \tag{1.11}$$

where this last equality allows us to use this recursion relation to compute all the remaining terms of the infinite series.

Having written down the general form of a proper rotation, i.e., any rotation that is in that part of the manifold that is connected to the identity, we should remember that the entire manifold actually comes in two parts, with those elements in the other portion described by having a determinant of value  $-1$ . As already noted, the *parity* transformation, that lies in this other portion. Therefore, all the rotations in this other portion, i.e., those with determinant  $-1$ , can be thought of either as a proper rotation multiplied by the parity matrix, which is a good “physical” description for them, or, simply as the negative of a proper rotation, so that not only does it rotate the axes as desired, but it reverses all their directions as well.

## II. Representations of the proper Rotation Group

We have so far only created a method to look at the action of rotations on 3-dimensional vectors; however, we certainly alluded to the existence of other sorts of vector spaces, the elements of which might be interesting tensors of other sorts, or even, as it turns out, the objects called spinors. We therefore now want to extend our understanding of the rotation group so that we may answer how a rotation in some “underlying” vector space affects the behavior of geometrical objects that are more complicated. This is done via sets of matrices referred to as *representations*.

A representation of a group is a (continuous) mapping that sends each element of the group into a continuous linear operator that acts on some vector space, **and** which preserves the group product. An irreducible representation is one which does not leave invariant any proper subspace of the vector space on which it acts. One may also have representations of the Lie algebra of a Lie group. The requirement that this be a representation is similar, namely that the mapping preserve the commutation relations of the algebra elements. As it is a vector space it is of course sufficient to simply preserve the commutation relations of the basis vectors. Given a representation of the Lie group, the view of the Lie algebra as its tangent vector space at the identity allows one to easily calculate the representation of the Lie algebra that is associated with that representation of the group, just by taking appropriate derivatives in the parameter space. On the other hand, given a representation of the Lie algebra the exponential series acting on that representation **may not** give a representation of the Lie group; instead, depending on the representation, the result of exponentiation may be a representation of a larger group, instead. That larger group is the so-called *universal covering group* of the original group. In general the universal covering group has an  $n$  to 1 mapping down to the original group, where  $n$  is the number of non-equivalent closed paths one can draw in the group manifold, which is often referred to by saying that the group is  $n$ -ly connected.

The rotation group, as an example of a compact group, has some nice properties for its irreducible representations, which we do not prove here but simply quote, from well-established sources, since we actually only need to make use of the results rather than the proofs.

- a. We first recall that the rotation group is doubly connected. Therefore it has a universal covering group which is larger than itself; that group is  $SU(2)$ , the group of all  $2 \times 2$ , unitary, complex-valued matrices that have determinant  $+1$ . See more discussion of this group and its  $2 \rightarrow 1$  mapping onto the rotation group below, after the description of the  $j = \frac{1}{2}$  representation.
- b. All the irreducible representations are in terms of finite-dimensional matrices.
- c. The matrices in question are all unitary, which is of course the same as orthogonal when the matrix elements are all real-valued. **There is exactly one irreducible representation**

**for each finite dimension,  $n$ .** It is customary to label these representations not by  $n$ , however, but by the value of  $j$ , where  $n \equiv 2j + 1$ ; the value of  $j$  is often called the “spin” of the representation. In the case that  $j$  is an integer then there is a choice of basis for the matrices such that they have only real-valued elements. When  $j$  is, instead, a half-integer, so that the matrices are even-dimensional the matrices are not true representations, but, instead representations of  $\mathbf{SU}(2)$ , so that there is a 2 to 1 mapping back to the rotation group itself.

- c. There are also infinite-dimensional representations, created from infinite sums or integrals of irreducible ones, which are best visualized as unitary operators on some appropriate space of functions, which is best thought of as infinite-dimensional vector space.

We will denote an arbitrary (unitary) representation of the rotation group, by  $U$ , so that the linear operators which are its values can be denoted by  $U(R)$ , or  $U[R(\theta, \hat{n})]$ , or, perhaps, as a sometimes-shorthand, just by  $U(\hat{n}; \theta)$ . As already noted, a representation of the group can be used to determine, uniquely, a representation of the Lie algebra of that group, since the Lie algebra are just the generators of the group near the identity. Denoting the generators by  $\bar{\mathcal{J}}_i$  we can write, for any choice of axis,  $\hat{n}$ ,

$$U[R(\theta, \hat{n})] = e^{-\theta \hat{n} \cdot \bar{\mathcal{J}}} . \quad (2.1)$$

$$\hat{n} \cdot \bar{\mathcal{J}} = - \lim_{\theta \rightarrow 0} \frac{1}{\theta} \left\{ U[R(\theta, \hat{n})] - I \right\} = - \frac{d}{d\theta} U[R(\hat{n}; \theta)] \Big|_{\theta=0} . \quad (2.2)$$

For a given vector space dimension  $n = 2j + 1$ , following very standard conventions, one may always choose a basis for that vector space such that the representation of  $\mathcal{J}_z$  is a diagonal matrix, although, since it is a skew-symmetric matrix, it should be noted that the elements on the diagonal will be imaginary. Since this is a very particular representation, labelled by the value of  $j$ , instead of this generic notation for the representation matrix, we will use instead a notation that reminds us of the value of  $j$ , namely  $\mathcal{J}_i^{(j)}$ . One might label that basis set as  $\{\hat{e}_a \mid a = 1, 2, 3, \dots, 2j + 1\}$ . However, it is much more conventional to, instead, use the

following symbols and labelling procedure for the basis set, which also reminds us of the value of  $j$  in question:

$$\left\{ |j, m\rangle \mid m = -j, -j + 1, -j + 2, \dots, j - 2, j - 1, j \right\}. \quad (2.3)$$

where there are of course  $2j + 1$  values for  $m$  which are all either half-integer or integer, depending on whether  $j$  is half-integer or integer. Using the commutators of the representation, one may then determine the elements of the three representation matrices, in this basis, to be determined by the following relations, where the first one of course is simply a re-statement of the fact that  $\mathcal{J}_z^{(j)}$  is diagonal in this basis:

$$\begin{aligned} \mathcal{J}_z^{(j)} |j, m\rangle &= -i m |j, m\rangle, \\ \mathcal{J}_+^{(j)} |j, m\rangle &= -i \sqrt{(j - m)(j + 1 + m)} |j, m + 1\rangle, \\ \mathcal{J}_-^{(j)} |j, m\rangle &= -i \sqrt{(j + m)(j + 1 - m)} |j, m - 1\rangle, \end{aligned} \quad (2.4)$$

$$\mathcal{J}_\pm \equiv \mathcal{J}_x \pm i \mathcal{J}_y$$

where the last line shows how to relate the original generators,  $\mathcal{J}_x$  and  $\mathcal{J}_y$  to the so-called “raising” and “lowering” operators that transform any given basis vector into one with a value of  $m$  that is either +1 higher, or lower. It can be seen that, in this basis, not only is the matrix that represents  $\mathcal{J}_z$  imaginary, but also the one representing  $\mathcal{J}_x$  is imaginary, while the one representing  $\mathcal{J}_y$  is real-valued. One may also use these to calculate the values of the so-called Casimir operator on the group,  $(\vec{\mathcal{J}}^{(j)})^2$ , where a Casimir operator is one which is always a representation-dependent multiple of the identity matrix in any given, irreducible representation:

$$(\vec{\mathcal{J}}^{(j)})^2 |j, m\rangle = -j(j + 1) |j, m\rangle. \quad (2.5)$$

The matrices for the irreducible representations of the group elements are conventionally denoted by the symbol  $D^{(j)}[R(\theta, \hat{n})]$ —usually written simply as  $D^{(j)}(\theta, \hat{n})$ —and are unitary. Their matrix elements can be given as

$$D^{(j)}(R)^{m' m} \equiv \langle j, m' | U(R) | j, m \rangle = \langle j, m' | e^{-\theta \hat{n} \cdot \vec{\mathcal{J}}} | j, m \rangle = \left( e^{-\theta \hat{n} \cdot \vec{\mathcal{J}}^{(j)}} \right)^{m'}_m. \quad (2.6)$$

1. **The case  $\mathbf{j} = 1$ .**

Using the formulae above, we easily find that

$$\mathcal{J}_z^{(1)} \implies -i \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \quad \mathcal{J}_x^{(1)} \implies \frac{-i}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad \mathcal{J}_y^{(1)} \implies \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}. \quad (2.7)$$

Since these are  $3 \times 3$  matrices, and there is only one irreducible representation for each dimension, one might wonder why these matrices are distinct from the  $3 \times 3$  matrices given in Eqs. (1.4). The answer is that they are equivalent, via a similarity transformation, to those others; more precisely, the matrices in Eqs. (1.4) are relative to a Cartesian basis set, while these are relative to a basis in which  $\mathcal{J}_z$  is diagonal. Finding the eigenvectors for the Cartesian presentation of  $\mathcal{J}_z$ , and making the standard choices about signs and normalizations, one finds the following relationships between the two basis sets:

$$|1, 0\rangle = \hat{z}, \quad |1, 1\rangle = \frac{-1}{\sqrt{2}} (\hat{x} + i\hat{y}), \quad |1, -1\rangle = \frac{+1}{\sqrt{2}} (\hat{x} - i\hat{y}). \quad (2.8)$$

2. **The case  $\mathbf{j} = \frac{1}{2}$ .**

In this case, the vector space is only 2-dimensional, and the representations of the generators are

$$\mathcal{J}_z^{(\frac{1}{2})} = -\frac{i}{2} \begin{pmatrix} +1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \mathcal{J}_x^{(\frac{1}{2})} = -\frac{i}{2} \begin{pmatrix} 0 & +1 \\ +1 & 0 \end{pmatrix}, \quad \mathcal{J}_y^{(\frac{1}{2})} = -\frac{i}{2} \begin{pmatrix} 0 & -i \\ +i & 0 \end{pmatrix}, \quad (2.9)$$

where these three matrices are just a factor of  $-i/2$  times the standard three *Pauli matrices*, usually called  $\sigma_i$ . By explicit calculation of the exponential of these, one finds that

$$D^{(\frac{1}{2})}[R(\theta, \hat{n})] = \cos(\theta/2) \mathbf{I}_2 - \hat{n} \cdot \vec{\mathcal{J}}^{(\frac{1}{2})} \sin(\theta/2). \quad (2.10)$$

It is therefore immediately clear that

$$D^{(\frac{1}{2})}(\hat{n}; 2\pi) = -\mathbf{I}_2, \quad \forall \hat{n}, \quad (2.11)$$

which makes it clear that this is not a *faithful representation* of  $\mathbf{SO}(3)$ .



The general notion of a *ray representation* is a mapping,  $\check{U}$ , where a phase is allowed to “creep into” the representation of the product, as follows:

$$\check{U}(R_1)\check{U}(R_2) = e^{i\eta(R_1,R_2)}\check{U}(R_1R_2) , \quad (2.12)$$

where  $\eta$  is a real-valued function of the two rotations. Since this “extra” factor has an absolute value of +1, in quantum mechanics it will never be measured. Therefore, there is a general theorem, due to Eugene Wigner, that quantum mechanics “allows,” or even “wants,” ray representations rather than the more usual sort of representations. As well, Wigner showed that these ray representations, of a group  $G$ , can always be normalized so that they are true representations of the *universal covering group* of  $G$ , which we will denote by  $\overline{G}$ . In general, as well, the universal covering group is the unique group that has a homomorphism from  $\overline{G}$  to  $G$  and is simply connected. The homomorphism from the universal covering group to the rotation group is given, for  $g \in \mathbf{SU}(2)$  to  $R(g) \in \mathbf{SO}(3)$ , is given by the following equation:

$$[R(g)]^a_b = \frac{1}{2} \text{tr} \{g^\dagger \sigma^a g \sigma_b\} . \quad (2.13)$$

This equation is obviously (at least) two to one, since if  $g \in \mathbf{SU}(2)$  then  $-g$  is also a member of  $\mathbf{SU}(2)$ , and, since the expression is quadratic in  $g$ , then both  $g$  and  $-g$  give the same values for the matrix elements of  $R(g)$ .

Having already created a reasonable “visualization” of the manifold for  $\mathbf{SO}(3)$ , it is also useful to create such a “visualization” of the manifold for  $\mathbf{SU}(2)$ . First, notice that an arbitrary element of  $\mathbf{SU}(2)$  can be written in the form

$$g \in \mathbf{SU}(2) \quad \Longrightarrow \quad g \longmapsto \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix} , \quad \text{where } |a|^2 + |b|^2 = +1 . \quad (2.14)$$

Given the constraint on the complex numbers  $a$  and  $b$ , we have three real degrees of freedom available, as expected, since there is a 2-1 homomorphism onto  $\mathbf{SO}(3)$ . Therefore, the set  $\{\text{Re } a, \text{Im } a, \text{Re } b, \text{Im } b\}$ , subject to the constraint  $(\text{Re } a)^2 + (\text{Im } a)^2 + (\text{Re } b)^2 + (\text{Im } b)^2 = +1$ ,

constitutes a set of three real parameters to describe the group. However, the constraint equation is clearly just the equation for the (surface of the) sphere  $S^3$ , i.e., the 3-dimensional sphere in 4-dimensional flat space. Looking back at Eq. (2.11) which says that the ray-representation of the rotation by  $2\pi$  is the same no matter what is the direction of the associated axis. Therefore, a reasonable way to think about this sphere is that the origin is the North pole, and the rotation by  $2\pi$  is the South pole. This also makes sense because of the fact that a rotation by  $4\pi$  is represented by the identity matrix, again. Notice as well that the surface of a sphere, in whatever number of dimensions, is simply connected, so that this group would be its own universal covering group.

### 3. The case $\mathbf{j} = \mathbf{0}$ .

In that case the matrices that represent the rotations are just 1-dimensional, i.e., just numbers, and the matrices that represent the generators are skew-symmetric numbers, i.e., numbers that equal their own negatives. There is only one such number, namely 0, and the exponential of zero is just +1. Therefore, as might should have been expected, this is the completely trivial representation where every group element is simply represented by +1. Yes, it's a representation, but not at all interesting.

## III. Description of the Lorentz Group

One could just say that the Lorentz group is the set of all “orthogonal” transformations of vectors defined over spacetime, where here the fact that “orthogonal” means preserving lengths is “translated” to mean that it preserves the metric value of the length of a vector in spacetime. This approach would give us the name  $\mathbf{O}(3,1)$  for the Lorentz group, where the two integers after the orthogonality statement give us the number of spacelike directions and the number of timelike directions. A more explicit approach to this generic statement would be to remind ourselves about the metric in spacetime and the explicit matrix presentation of it and the transformations, in a manner very similar to the procedure used in Eq. (1.1) above, except that now the metric is not just the Cartesian one, but the Minkowski one, where we

denote the matrix of its components by  $\eta_{\mu\nu}$ , and the matrix which presents it by  $H$ , which is a capital Greek “eta”, where we use the components of a 4-vector,  $\tilde{A}$  to present it all:

$$\begin{aligned}
(A^1)^2 + (A^2)^2 + (A^3)^2 - (A^4)^2 &\equiv \vec{A}^2 - (A^4)^2 \equiv \tilde{A} \cdot \tilde{A} \equiv \eta_{\mu\nu} A^\mu A^\nu \equiv \tilde{A}^T H \tilde{A}; \\
\Lambda \in SO(3, 1) &\iff \Lambda^T H \Lambda = H \quad \text{or} \quad \Lambda^\mu{}_\alpha \eta_{\mu\nu} \Lambda^\nu{}_\beta = \eta_{\alpha\beta}, \\
\implies \Lambda^T H &= H \Lambda^{-1} \implies \Lambda^{-1} = H^{-1} \Lambda^T H \quad \text{or} \quad (\Lambda^{-1})^\beta{}_\mu = \eta^{\beta\gamma} \Lambda^\nu{}_\gamma \eta_{\nu\mu} = \Lambda_\mu{}^\beta,
\end{aligned} \tag{3.1}$$

where we have used  $\tilde{A}$  to denote **both** the actual 4-vector and the matrix, or column vector, which has its components as elements. Here the fact that  $\eta$  is a type  $\begin{bmatrix} 0 \\ \end{bmatrix}$  tensor can be seen easily when we write it in terms of its components with an indication as to whether they transform in a contravariant way—upper indices—or a covariant way—lower indices. On the last line we have also solved the constraining equation for  $\Lambda$  in a way to give an explicit expression for its matrix inverse in terms of itself and the metric,  $H$ , and have also used the additional property ascribed to  $H$ , of raising and lowering indices.

As before with the rotation group, we now want to characterize the Lorentz group in terms of the properties of its associated Lie algebra. Therefore, we begin by considering those elements of the group that are “not too far” from the identity element, and to determine their logarithms, which we refer to as  $Q$  for the moment. We of course also recall that the vector space of all possible  $Q$ ’s will constitute the Lie algebra for the Lorentz group. Therefore, we again re-consider the constraint on the group elements in terms of the constraint that it puts on the corresponding Lie algebra:

$$\begin{aligned}
\Lambda = e^Q \quad \text{or, in terms of components, } \Lambda^\mu{}_\alpha &= e^{Q^\mu{}_\alpha} \implies e^{Q^T} H e^Q = H \\
\implies e^{Q^T} &= H e^{-Q} H^{-1} = e^{-HQH^{-1}} \\
\implies Q^T &= -HQH^{-1} \implies Q^T H = -HQ \implies (HQ)^T = -(HQ) \\
\text{or, in terms of components, } Q^\mu{}_\alpha \eta_{\mu\beta} &= -\eta_{\alpha\mu} Q^\mu{}_\beta \iff Q_{\mu\alpha} = -Q_{\alpha\mu}.
\end{aligned} \tag{3.2}$$

The very last statement says that the matrix  $HQ$  may be any matrix whatsoever that is skew-symmetric.

Such a set of matrices is a 6-dimensional vector space; therefore, we begin by choosing a basis of 6 skew-symmetric matrices. We could simply number them from 1 to 6, but a notation that turns out to be much more convenient is to label them via a skew-symmetric labeling scheme, making the skew-symmetry more immediate. Therefore, we pick out a set of 6 distinct, linearly-independent skew-symmetric matrices, and label them by the symbols  $\{\mathcal{J}^{\alpha\beta} = -\mathcal{J}^{\beta\alpha} \mid \alpha, \beta = 1, 2, 3, 4\}$ , where the labelling arranges it so that there are only 6 independent skew-symmetric matrices, which could, for instance, be labelled as  $\{\mathcal{J}^{12}, \mathcal{J}^{31}, \mathcal{J}^{14}, \mathcal{J}^{23}, \mathcal{J}^{24}, \mathcal{J}^{34}\}$ . Since we know that the 3-dimensional rotations considered earlier obviously are also Lorentz transformations, we want to choose a basis for these that automatically includes the  $3 \times 3$  matrices we used for generators there, simply put into a 4-dimensional context. Remembering that those matrices simply had +1 in one location, a -1 in the alternate location, and zeroes everywhere else. We then continue with that idea for the 3 more that we need to complete our necessary six, giving them in our Cartesian basis set:

$$(\mathcal{J}^{\alpha\beta})_{\mu\nu} \equiv \eta_{\mu\lambda} (\mathcal{J}^{\alpha\beta})^{\lambda}_{\nu} \equiv -(\delta_{\mu}^{\alpha} \delta_{\nu}^{\beta} - \delta_{\nu}^{\alpha} \delta_{\mu}^{\beta}) = -\delta_{[\mu}^{\alpha} \delta_{\nu]}^{\beta}. \quad (3.3)$$

For example, the explicit presentation of two of these are the following:

$$\mathcal{J}^{12} = (((\mathcal{J}^{12})_{\mu\nu})) = \begin{pmatrix} 0 & -1 & 0 & 0 \\ +1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \mathcal{J}^{34} = (((\mathcal{J}^{34})_{\mu\nu})) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & +1 & 0 \end{pmatrix}, \quad (3.4)$$

where we notice that the matrix for  $\mathcal{J}^{12}$  is just the matrix we used earlier for  $\mathcal{J}^z$  but with lots of extra zeroes to put it into a  $4 \times 4$  format. Having the basis set our general form for an arbitrary logarithm,  $Q$ , i.e., an arbitrary element of the Lorentz Lie algebra, is just some arbitrary linear combination of all of these matrices. However, as they are labeled by their skew-symmetric pair of indices it is more reasonable to also label the 6 required scalar coefficients in the same way. Therefore, we choose a set of 6 scalar quantities, which are to be thought of as the 6 independent parameters that describe an arbitrary Lorentz transformation,  $\{q_{\alpha\beta} = -q_{\beta\alpha} \mid \alpha, \beta = 1, 2, 3, 4\}$ , or the set  $\{q_{12}, q_{31}, q_{14}, q_{23}, q_{24}, q_{34}\}$ , such that

$$\Lambda = e^Q, \quad Q = \frac{1}{2} q_{\beta\alpha} \mathcal{J}^{\alpha\beta}, \quad (3.5)$$

where the index order on the two entries has been chosen in the way that is most appropriate for matrix multiplication, while the factor 1/2 is simply because the double sum counts things twice because both sets of quantities are skew-symmetric in their indices.

While the method above is quite useful, at least in part because it is explicitly 4-dimensional at every possible place, we have better intuition about a breakdown of the 4 dimensions into 3 spatial ones and 1 temporal one. Therefore it is important to also divide these 6 basis matrices in a different way, into 2 sets of 3. We define  $\{\mathcal{J}_i\}_{i=1}^3$  and  $\{\mathcal{K}^j\}_{j=1}^3$  so that the following statements are true:

$$\mathcal{J}^{ij} = \eta^{ijk} \mathcal{J}_k = \eta^{ijk4} \mathcal{J}_k, \quad \mathcal{J}^{4i} = \mathcal{K}^i. \quad (3.6)$$

The following are a couple of examples, in fact the same examples as shown above in Eqs. (3.4), but where we have raised the matrix index for the row index—the left one—so that we have the form of an actual matrix that will act on 4-vectors:

$$\mathcal{J}_z = (((\mathcal{J}_z)^\mu{}_\beta)) = \begin{pmatrix} 0 & -1 & 0 & 0 \\ +1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \mathcal{K}^3 = (((\mathcal{K}^3)^\mu{}_\beta)) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}. \quad (3.7)$$

In this form we may then choose the parameters themselves,  $q_{\alpha\beta}$ , so that their physical meaning is clear. Working from Eqs. (3.5-6) we have the form of the most general logarithm matrix,  $Q$ , in terms of a different division of the 6 parameters that describe an arbitrary Lorentz transformation:

$$Q = \frac{1}{2} q_{\beta\alpha} \mathcal{J}^{\alpha\beta} = -\theta \hat{n} \cdot \vec{\mathcal{J}} - \lambda \hat{v} \cdot \vec{\mathcal{K}}, \quad q_{ij} = \theta \eta_{ijk} \hat{n}^k, \quad q_{4i} = -\lambda \hat{v}_i. \quad (3.8)$$

This arrangement has made it clear which are the Lorentz transformations that are really just rotations in 3-dimensional space, leaving the  $\hat{t}$  basis vector alone, and which are the so-called “pure” Lorentz transformations, affecting boosts between two reference frames that have their basis vectors aligned but which measure the other one to be moving with non-zero, constant velocity.

Before going ahead to determine irreducible representations of these transformations it is necessary to first determine the commutators of these 6 generators, where it is useful to have the commutators available to us in both forms of their labelling. There are various “sophisticated” ways to calculate their commutators; however, as we have explicit  $4 \times 4$  matrices one can quite easily calculate the commutators directly. The results are the following:

$$[\mathcal{J}^{\alpha\beta}, \mathcal{J}^{\gamma\delta}] = \eta^{\alpha\gamma} \mathcal{J}^{\beta\delta} - \eta^{\beta\gamma} \mathcal{J}^{\alpha\delta} + \eta^{\beta\delta} \mathcal{J}^{\alpha\gamma} - \eta^{\alpha\delta} \mathcal{J}^{\beta\gamma} = \eta^{\alpha[\gamma} \mathcal{J}^{\beta\delta]} - \eta^{\beta[\gamma} \mathcal{J}^{\alpha\delta]} , \quad (3.9)$$

where a way of seeing why there are 4 distinct terms in the commutator is by noting that the result must be skew-symmetric in  $\alpha$  and  $\beta$ , and also skew-symmetric in  $\gamma$  and  $\delta$ , and lastly must be skew-symmetric in switching those two pairs. However, we also need the commutators when they are divided into pure rotations and pure boosts:

$$\begin{aligned} [\mathcal{J}_m, \mathcal{J}_n] &= + \epsilon_{mnc} \mathcal{J}^c , \\ [\mathcal{J}_s, \mathcal{K}_t] &= + \epsilon_{stb} \mathcal{K}^b , \\ [\mathcal{K}^b, \mathcal{K}^d] &= - \epsilon^{bdf} \mathcal{J}_f . \end{aligned} \quad (3.10)$$

The exponentials of the Lie algebra elements that are sums only of the three  $\mathcal{J}_i$  are of course just a  $3 \times 3$  rotation matrix put into a  $4 \times 4$  form with a +1 in the 4, 4-position and 0’s in the 4,  $i$ - and  $i$ , 4-positions, showing that it performs a spatial rotation and neither changes the temporal values nor mixes them with the spatial ones. We already understand their structure. On the other hand the set of linear combinations of the other three generators, the  $\mathcal{K}^m$ , are the usual Lorentz boosts that move the basis vectors from one reference frame to another reference frame that is moving with respect to the first one. It is these boosts that one needs to describe for a moment. We determine the signs of the boost transformation so that the general form of a Lorentz boost is given by the following:

$$L(\lambda, \hat{v}) \equiv e^{-\lambda \hat{v} \cdot \vec{\mathcal{K}}} = \mathbf{I}_4 - \sinh \lambda \hat{v} \cdot \vec{\mathcal{K}} + (\cosh \lambda - 1)(\hat{v} \cdot \vec{\mathcal{K}})^2 , \quad (3.11)$$

which is the transformation that takes us from the coordinates of some 4-vector that was measured in the frame  $\mathcal{O}$  to the coordinates of that same 4-vector which were measured in

the frame  $\mathcal{O}'$ , where  $\mathcal{O}$  measures that  $\mathcal{O}'$  is moving with (constant) velocity  $\vec{v}$ . Knowing the simplest forms of the Lorentz transformation laws, in, say, 1 space and 1 time dimension, we use  $\lambda$  as the rapidity for our boost, and have the following properties, using the matrix form for  $\vec{\mathcal{K}}$  given above, where it is very useful for presenting this to use  $4 \times 4$  matrices that have been divided into their  $3 \times 3$  part, their  $3 \times 1$  and  $1 \times 3$  parts, and the final 4, 4-component of the matrix:

$$\begin{aligned} \Rightarrow \left. \begin{aligned} \tanh \lambda &\equiv v = |\vec{v}|, \\ \cosh \lambda &= \gamma_v = (1 - v^2)^{-\frac{1}{2}}, \\ \sinh \lambda &= \gamma v; \end{aligned} \right\}, \quad \begin{aligned} \hat{v} \cdot \vec{\mathcal{K}} &= \begin{pmatrix} \mathbf{0}_3 & \hat{v} \\ \hat{v}^T & 0 \end{pmatrix}, \\ (\hat{v} \cdot \vec{\mathcal{K}})^2 &= \begin{pmatrix} \hat{v} \hat{v}^T & \vec{0} \\ \vec{0}^T & 1 \end{pmatrix}, \\ (\hat{v} \cdot \vec{\mathcal{K}})^3 &= \hat{v} \cdot \vec{\mathcal{K}}, \end{aligned} \quad (3.12) \\ \Rightarrow L(\lambda, \hat{v}) &= \begin{pmatrix} \mathbf{I}_3 + (\gamma - 1)\hat{v}\hat{v}^T & -\gamma\vec{v} \\ -\gamma\vec{v}^T & +\gamma \end{pmatrix}. \end{aligned}$$

We have already given the form of a general rotation, in Eq. (1.10). There are of course other sorts of elements of the Lie algebra, where, for instance, one picks a non-zero coefficient for  $\mathcal{J}_z$  and  $\mathcal{K}_z$ ; forms such as that are sometimes called parabolic transformations. There is an interesting and useful way of parametrizing the Lorentz group by picking out a particular null vector and a conjugate one, such as, for example,  $\hat{z} \pm \hat{t}$ . Then use 2 parameters to re-scale the affine parameters of the two null vectors, while the other 4 rotate all the basis vectors around the particular one with which one began. This method is quite important in more advanced studies of the solutions of the Einstein field equations.

However, there is a theorem that says that all proper Lorentz transformations may be made from a proper of a boost and a rotation. Moreover, it is not true that the series expansion for the logarithm converges over the entire manifold, so that it is not true that every proper Lorentz transformation may be written as the exponential of a single element of the Lie algebra. One should perhaps then think just a bit about the manifold constituted by the proper Lorentz transformations. We already know that the manifold of all rotations can be visualized as a solid ball of radius  $\pi$ , with its antipodal points identified, so that it is doubly connected. However,

the boosts are described by any possible direction, but their parameter  $\lambda$  should be considered as varying between 0 and  $+\infty$ , so that this is basically just the usual, complete 3-dimensional space,  $\mathbb{E}^3$ . One should note that in principle one could have  $\lambda$  negative; however, one can simply take a positive  $\lambda$  and reverse the direction associated with the negative one, so that we do not do this, to avoid double counting. The entire manifold is then the direct product of these two 3-dimensional manifolds, which is clearly not compact, i.e., closed and bounded (inside some sufficiently large sphere, say).

We also know that the product of two boosts in different directions is not just a boost, since the commutator of two non-parallel  $\mathcal{K}_i$  is a  $\mathcal{J}_k$ , so that the product of two different boosts is in general equal to the product of a third boost, which takes one directly to the frame which has the correct magnitude of velocity for the two separate boosts, but then one also needs a rotation to align the axes properly.

Lastly one should say some things about those parts of the entire group that are not in the connected part of the manifold containing the identity. From the fact that the rotations have a disconnected part, we may see that it is obvious that the Lorentz group must also have a disconnected part, generated by the parity transformation, which in 4-dimensional spacetime has the following form:

$$\textit{parity transformation:} \quad \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & +1 \end{pmatrix}. \quad (3.13)$$

This corresponds again to the fact that the defining equation for a Lorentz transformation allows the determinant of a general Lorentz transformation to be either  $+1$  or  $-1$ , but no other choices, so that this divides the manifold into two portions. On the other hand, since we are always leaving the time coordinate the same here, it would seem that there ought to be more symmetry. We can find another such situation by returning to Eqs. (3.1) and considering the



4, 4-component of the equation that requires a Lorentz transformation to preserve the metric:

$$\begin{aligned}
-1 = \eta_{44} &= (\Lambda^T H \Lambda)_{44} = \Lambda^\mu{}_4 \eta_{\mu\nu} \Lambda^\nu{}_4 = \sum_{i=1}^3 (\Lambda^i{}_4)^2 - (\Lambda^4{}_4)^2 \\
\implies (\Lambda^4{}_4)^2 &= +1 + \sum_{i=1}^3 (\Lambda^i{}_4)^2 \\
\implies \Lambda^4{}_4 &\geq +1 \quad \text{or} \quad \Lambda^4{}_4 \leq -1.
\end{aligned} \tag{3.14}$$

this is again a separation of the entire group into two separate pieces: those which do something or other but do not change the direction of time, and those which do; therefore they are referred to as *orthochronous transformations* and *non-orthochronous transformations*. The generating elements in each one are of course the identity element, which is orthochronous, and time reversal:

$$\text{time reversal transformation: } \begin{pmatrix} +1 & 0 & 0 & 0 \\ 0 & +1 & 0 & 0 \\ 0 & 0 & +1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{3.15}$$

There are however also Lorentz transformations which are improper and non-orthochronous, i.e., which have  $\det \Lambda = -1$  and  $\Lambda^4{}_4 \leq -1$ . The obvious example entry is the negative of the identity element, which is however seen more physically as the product of parity and time reversal. Therefore this manifold comes in 4 separate, disconnected pieces, only one of which contains the identity, and of course only that one is a subgroup.

#### IV. Representations of the full proper Lorentz Group

The Lorentz group has two triplets of generators, which do not commute with each other; however, it turns out that we may take complex linear combinations of them which do. We begin by defining

$$\mathcal{F}^j \equiv \frac{1}{2}(\mathcal{J}^j + i\mathcal{K}^j), \quad \mathcal{G}^k \equiv \frac{1}{2}(\mathcal{J}^k - i\mathcal{K}^k). \tag{4.1}$$

We now show below that the commutators of each of these triplets has exactly the same structure as the commutators of the  $\mathcal{J}^m$  themselves; i.e., they generators a group that has exactly the

same group-theoretical structure as the rotation group, albeit of course with complex-valued generators:

$$\begin{aligned}
[\mathcal{F}^j, \mathcal{F}^k] &= \frac{1}{4}[\mathcal{J}^j + i\mathcal{K}^j, \mathcal{J}^k + i\mathcal{K}^k] = \frac{1}{4}\epsilon_{jkm}\{\mathcal{J}_m + i\mathcal{K}_m - (-i\mathcal{K}_m) + \mathcal{J}_m\} = \epsilon^{jkm}\mathcal{F}_m, \\
[\mathcal{G}^j, \mathcal{G}^k] &= \frac{1}{4}[\mathcal{J}^j - i\mathcal{K}^j, \mathcal{J}^k - i\mathcal{K}^k] = \frac{1}{4}\epsilon_{jkm}\{\mathcal{J}_m - i\mathcal{K}_m - (+i\mathcal{K}_m) + \mathcal{J}_m\} = \epsilon^{jkm}\mathcal{G}_m, \\
[\mathcal{F}^j, \mathcal{G}^k] &= \frac{1}{4}[\mathcal{J}^j + i\mathcal{K}^j, \mathcal{J}^k - i\mathcal{K}^k] = \frac{1}{4}\epsilon_{jkm}\{\mathcal{J}_m - i\mathcal{K}_m - (-i\mathcal{K}_m) - \mathcal{J}_m\} = 0,
\end{aligned} \tag{4.2}$$

where the last line shows, additionally, that the two triplets commute between themselves, so that the two ‘‘copies’’ of the rotation group that are involved do not interact with each other, so that we could treat each one of them separately. Therefore we may now consider, independently, representations of each of these ‘‘rotation groups,’’ labeling them by  $f$  and  $g$ , where each of  $2f$  and  $2g$  are allowed to take on all non-zero integer values. Since we may resolve the original equations backwards, for  $\mathcal{J}^j$  and  $\mathcal{K}^j$ , as follows:

$$\mathcal{J}^m = \mathcal{F}^m + \mathcal{G}^m, \quad \mathcal{K}^m = i(\mathcal{G}^m - \mathcal{F}^m), \tag{4.3}$$

the desired representations of the Lorentz group are simply direct sums of the two representations of the rotation group, so that it is reasonable to label an irreducible representation of the Lorentz group by the two half-integers,  $f$  and  $g$ :  $D^{(f,g)}[\Lambda(\lambda; \hat{v})]$ , for example, or we could refer to just the representation itself by the simpler notation,  $D(f, g)$ .

The simplest two of these representations would simply be  $2 \times 2$  complex matrices, given by either  $D(0, \frac{1}{2})$  or  $D(\frac{1}{2}, 0)$ . Given that these take either  $\mathcal{F}^m$  or  $\mathcal{G}^n$  to be like the  $j = 1/2$  representation of the rotation group and the other one to vanish, we may easily recall that this means that our representation for the rotation group is given by  $\vec{\mathcal{J}} \implies -\frac{i}{2}\vec{\sigma}$ , we may complete the case fairly easily to obtain the forms below:

$$\begin{aligned}
D(0, \frac{1}{2}) &\implies \begin{cases} \vec{\mathcal{J}} \implies -\frac{i}{2}\vec{\sigma} \\ \vec{\mathcal{K}} \implies +\frac{1}{2}\vec{\sigma} \end{cases} \implies \begin{cases} D[R(\theta; \hat{e})] = (e^{+i(\theta/2)\hat{e}\cdot\vec{\sigma}}) \\ D[\Lambda(\lambda; \hat{v})] = (e^{-(\lambda/2)\hat{v}\cdot\vec{\sigma}}) \end{cases}, \\
D(\frac{1}{2}, 0) &\implies \begin{cases} \vec{\mathcal{J}} \implies -\frac{i}{2}\vec{\sigma} \\ \vec{\mathcal{K}} \implies -\frac{1}{2}\vec{\sigma} \end{cases} \implies \begin{cases} D[R(\theta; \hat{e})] = (e^{+i(\theta/2)\hat{e}\cdot\vec{\sigma}}) \\ D[\Lambda(\lambda; \hat{v})] = (e^{+(\lambda/2)\hat{v}\cdot\vec{\sigma}}) \end{cases},
\end{aligned} \tag{4.4}$$

To study more complicated forms than these, such as, for instance,  $D(\frac{1}{2}, \frac{1}{2}) = D(0, \frac{1}{2}) \otimes D(\frac{1}{2}, 0)$ , we will need to understand how to create the generators when considering a product representation. We understand that the matrices representing the group elements of a product representation correspond to direct products of the individual representations of that group element. However, one must be somewhat more careful with respect to the generators. The following is the right approach for the generators, for any Lie group,  $G$ , with Lie algebra,  $\mathcal{G}$ . Let us suppose that we have been an element  $g \in G$  such that there exists  $Q \in \mathcal{G}$  such that  $g = e^Q$ . Further let us have two distinct representations of  $G$ , which we label by  $D_1(g)$  and  $D_2(g)$ . Therefore, we also can find, and label, the representations of the generator  $Q$ , simply by  $Q^{(1)}$  and  $Q^{(2)}$ , so that

$$g = e^Q \implies \begin{cases} D_1(g) = e^{Q^{(1)}} , \\ D_2(g) = e^{Q^{(2)}} . \end{cases} \quad (4.5)$$

Then we find that the generators of the product representation are related to those of the individual representations in the following way:

$$\{D_1 \otimes D_2\}(g) \equiv [D_1(g)] \otimes [D_2(g)] \equiv e^{Q^{(1 \oplus 2)}} , \quad Q^{(1 \oplus 2)} \equiv Q^{(1)} \otimes \mathbf{I}_2 + \mathbf{I}_1 \otimes Q^{(2)} , \quad (4.6)$$

where  $\mathbf{I}_i$  indicates the matrix which is the identity matrix in representation  $D_i$ . To verify this we consider the following, setting, first  $T \equiv Q^{(1)}$ :

$$\begin{aligned} e^{T \otimes \mathbf{I}_2} &= \mathbf{I}_{1 \otimes 2} + T \otimes \mathbf{I}_2 + \frac{1}{2!}(T \otimes \mathbf{I}_2)(T \otimes \mathbf{I}_2) + \frac{1}{3!}(T \otimes \mathbf{I}_2)(T \otimes \mathbf{I}_2)(T \otimes \mathbf{I}_2) + \dots \\ &= \left\{ \mathbf{I}_1 + T + \frac{1}{2!}T^2 + \frac{1}{3!}T^3 + \dots \right\} \otimes \mathbf{I}_2 = D_1(g) \otimes \mathbf{I}_2 . \end{aligned} \quad (4.7)$$

We have an entirely analogous construction for the exponential of  $\mathbf{I}_1 \otimes Q^{(2)}$ , so that we may finally write

$$\begin{aligned} e^{Q^{(1)} \otimes \mathbf{I}_2 + \mathbf{I}_1 \otimes Q^{(2)}} &= e^{Q^{(1)} \otimes \mathbf{I}_2} e^{\mathbf{I}_1 \otimes Q^{(2)}} = \{D_1(g) \otimes \mathbf{I}_2\} \{\mathbf{I}_1 \otimes D_2(g)\} \\ &= D_1(g) \otimes D_2(g) \equiv \{D_1 \otimes D_2\}(g) , \end{aligned} \quad (4.8)$$

where we have used the ‘‘obvious’’ fact that  $Q^{(1)} \otimes \mathbf{I}_2$  commutes with  $\mathbf{I}_1 \otimes Q^{(2)}$ .

With that notion in hand, we may immediately write down the generators for the representation  $D(\frac{1}{2}, \frac{1}{2})$ :

$$\vec{\mathcal{J}} \implies -\frac{i}{2}\{\vec{\sigma} \otimes \mathbf{I}_2 + \mathbf{I}_2 \otimes \vec{\sigma}\}, \quad \vec{\mathcal{K}} \implies \frac{1}{2}\{\vec{\sigma} \otimes \mathbf{I}_2 - \mathbf{I}_2 \otimes \vec{\sigma}\} \quad (4.9)$$

$$\begin{aligned} \mathcal{J}^x &\implies -\frac{i}{2} \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}, \quad \mathcal{J}^y \implies +\frac{1}{2} \begin{pmatrix} 0 & -1 & -1 & 0 \\ 1 & 0 & 0 & -1 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \end{pmatrix}, \quad \mathcal{J}^z \implies i \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \\ \mathcal{K}^x &\implies \frac{1}{2} \begin{pmatrix} 0 & 1 & -1 & 0 \\ 1 & 0 & 0 & -1 \\ -1 & 0 & 0 & 1 \\ 0 & -1 & 1 & 0 \end{pmatrix}, \quad \mathcal{K}^y \implies +\frac{i}{2} \begin{pmatrix} 0 & -1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ -1 & 0 & 0 & -1 \\ 0 & -1 & 1 & 0 \end{pmatrix}, \quad \mathcal{K}^z \implies \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \end{aligned}$$

You should now immediately point out that we already know what the generators  $\vec{\mathcal{J}}$  and  $\vec{\mathcal{K}}$  look like in 4 dimensions, as they act on tangent vectors in spacetime, say, and the ones above are not those at all. This is certainly true; however, as usual, the problem here is that these are given with respect to a different set of basis vectors. With some algebra it is not too hard to show the following:

$$\begin{aligned} W \vec{\mathcal{M}}^{(\frac{1}{2}, \frac{1}{2})} W^{-1} &= \vec{\mathcal{M}}^{(3+1)}, \\ W &= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 & 0 & -1 \\ i & 0 & 0 & i \\ 0 & -1 & -1 & 0 \\ 0 & +1 & -1 & 0 \end{pmatrix}, \end{aligned} \quad (4.10)$$

where  $\mathcal{M}$  is meant to indicate either  $\mathcal{J}$  or  $\mathcal{K}$ , while the representation we just created is labelled by the superscript  $(\frac{1}{2}, \frac{1}{2})$  and the more customary representation is labelled by  $(3+1)$ , i.e., that one is our original  $\{x, y, z, t\}$  representation for these generators. This tells us, as expected, that the standard matrices for  $D(\frac{1}{2}, \frac{1}{2})$  are simply given with respect to a different basis, built on  $(x \pm iy)/\sqrt{2}$  and  $(\pm z - t)/\sqrt{2}$ , a choice of new basis which is not too surprising.