Notes on the Geometry of Spacetime,

and some associated Vector, Tensor and matrix Notation and Conventions

0. Background:

Special relativity comes from the experimental facts that all observers in inertial reference frames measure the same values for the speed of light rays in vacuum, and from the notion that measurements by those observers for physical quantities are related by Lorentz transformations, in the 4 dimensional spacetime in which we live.

There are a number of conventions regarding descriptions of these phenomena that are important to understand from the beginning. I will use these notes to point out several of Finley's conventions, at appropriate locations throughout the text. It is also a good place to "sneak in" Finley's notations for vectors, matrices, etc.

To begin, I note that he uses "geometrized units," where the units of length, time, and mass are all the same, and are interconvertible. This is accomplished by setting the speed of light, c, and the gravitational constant, G, equal to +1, so that there is no need to explicitly indicate their presence in formulas. Not indicating their presence explicitly is a good approach to the relevant physics since it emphasizes the underlying geometry, and, even more importantly, the validity of both c and G as having exactly the same values for **all** types of phenomena. This makes the SI units of meters and seconds interconvertible, so that spatial and temporal coordinates have the same "dimensions," either meters or seconds, as is of course much needed in discussing special relaivity. Some texts almost always insert the letter c in the correct places, if you need to recall where they should be entered; however, a "better" approach is for you to use dimensional analysis to insert the needed c's, when numerical calculations require it.
In MKS units their values are $c = 2.99792458 \times 10^{8}$ m/sec, and $G = 6.6726 \times 10^{-11}$ m^3/(kg-sec^2).
Some useful conversion factors which result from this are, for example

i.) 1 solar mass = 1.47664 kilometer = 1.989×10^{33} g = 4.9255 microseconds;

ii.) the mass of a proton, $m_p = 0.93826$ GeV $= 6.764 \times 10^{-57}$ km $= 1.0888 \times 10^{13}$ Kelvins;

iii.) the charge on a proton, $e = 1.381 \times 10^{-39}$ km .

1. Spacetime is a 4-dimensional place, sometimes referred to as $\mathcal{M}^4$, with points referred to as "events." We often label these points with a quadruplet of coordinates, for example (x, y, z, t) or (r, θ, φ, t), or $(x, y, u \equiv z + t, v \equiv z - t)$. The first of these sets is clearly a generalization of the usual Cartesian coordinates, to include time; I usually refer to it as Minkowski coordinates. Obviously the relations between two different sorts of sets of 4 coordinates can be quite complicated, such as between the Minkowski ones and the polar coordinates including time, presented just above, which you doubtless remember where to find them written down.

We also often denote these 4 coordinates by a single symbol, but with an index symbol that takes on 4 values: $\{x^\mu \mid \mu = 1, 2, 3, 4\}$, where these are simply generic symbols for a choice of kind of coordinates, or we suppose that we have already agreed on the kind we are using, i.e., for instance, Minkowski coordinates. It is important here to note that Finley labels the coordinates so that indices $\{1, 2, 3\}$ are spatial, and time is labeled 4, so that it comes last in the sequence of coordinates. **It is to be noted that** roughly half of all the texts on these subjects label the temporal coordinate as 0, so that the temporal portion comes first in the sequence of coordinates. This becomes more important when one looks at matrices using these two different orderings of components, so that matrices describing the same physical action may appear rather different.

Also notice that Finley uses lowercase Greek letters to indicate indices which (at least almost always) take on values from 1 to 4. Then Finley will uses lowercase Roman letters to indicate indices that take values only from 1 to 3, which then are representing only the spatial portions of some otherwise 4-dimensional object, so that we could consider the following relationship

$$(x^1, x^2, x^3, x^4) = (x^\mu) = (x^m, t) = (x^1, x^2, x^3, t) = (x, y, z, t) \ . \tag{1.0}$$

He also sometimes uses lowercase Roman letters to simply indicate indices that take values from 1 to some yet-unspecified integer value, m, **and** he uses uppercase Roman letters to indicate indices on 2x2 matrices [or the associated 2-dimensional vectors], which then run only from 1 to 2.

Different *allowed observers*, usually referred to as *inertial observers*, will ascribe values to these coordinates in different ways. Considering just two such observers, let us label them, or, better said, their associated reference frames—which they use to create coordinates for the events they observe—by $\mathcal{S}$ and $\mathcal{S}'$, and indicate the coordinates of events that they observe by $\{x^\mu\}$ and $\{x'^\alpha\}$, respectively. Provided that the two observers are using the same sort of coordinates—usually Minkowski/Cartesian unless otherwise specified—there is a standard form which must relate the two sets of measured coordinates, known in general as *Poincaré transformations*:

$$x^\mu = \Lambda^\mu{}_\alpha \, x'^\alpha + a^\mu \ , \tag{1.1}$$

where we must also note that Finley has used the *Einstein summation convention* in this equality, which says that if, in any given single term of a mathematical expression there are two objects with indices—one of the objects with upper indices, i.e., superscripts, and the other with lower indices, i.e., subscripts, both using the same letter to denote their indices—such as the α in the equality just above—then one must always assume that a summation over all the allowed values of that index is implied—even though no explicit summation symbol has been written—thereby saving ink, and handwriting.

The $\Lambda^\mu{}_\alpha$ are the elements of a matrix which accomplishes this transformation between the measurements—in the same sort of coordinates, such as, often, Minkowski coordinates—made by these two observers, when they have agreed to have the same origin for the coordinates, i.e., the places where all 4 of the coordinates are 0. Because of our desire to agree with the experimental results already mentioned, concerning measurements of the speed of light, these transformations are restricted so that both observers do indeed measure the same speed of light rays. The elements of the matrix do not depend on the specific events, but only on the relation between the two observers. The set of all such matrices is referred to as the set of all *Lorentz transformations*, which we will discuss in much more detail later on. We will see then

that there are exactly 6 independent degrees of freedom for such transformations; one good, physical way to look at those degrees of freedom is that they involve the 3 different directions that $\mathcal{S}$ might measure for the **constant** velocity of $\mathcal{S}^{'}$. Then the other 3 degrees of freedom correspond to the 3 distinct directions about which one might rotate the 3 spatial axes of the one observer relative to the other. (Do remember that inertial observers, those for whom Newton's laws of motion are appropriate, must be moving at a constant velocity.)
As well, the quantities a^μ allow the two observers to also have different origins; it should be clear that there are exactly 4 degrees of freedom in the choice of an origin, i.e., in the set of all possible values for a^μ. We should think of the 4 quantities, a^μ, as components of a vector that joins the two choices of origin for the two sets of coordinates. As well, we should think of Eq. (1.1) as a matrix equation, where we are presenting the vectors, and also the coordinates, in the form of column vectors, i.e., matrices with only one column. With this approach we also treat the 16 quantities $\Lambda^\mu{}_\alpha$ as the elements of a matrix, denoted by $\mathbf{\Lambda}$ in matrix notation. Then the product of $\mathbf{\Lambda}$ with the one set of coordinates, $\{x'^\alpha\}$, i.e., those with a prime, is conceived of as a matrix product—a square matrix multiplying a column vector on its left results in yet another column vector. In the form shown in that equation, the matrix multiplication is shown by the repeated value of α, *which invokes the Einstein summation convention* to imply a sum over all the allowed values of α. Notice that the sum is one that is equivalent to a matrix multiplication with $\mathbf{\Lambda}$ on the left. Denoting the explicit column vectors themselves, i.e., the ones with elements $\{x^\mu\}$, or $\{x'^\alpha\}$, by the symbols **x** and **x'**, respectively, along with **a** for the column vector with elements a^μ, we can re-write Eq. (1.1) as a matrix equation as follows:

$$\mathbf{x} = \mathbf{\Lambda}\mathbf{x}' + \mathbf{a}\ . \tag{1.1$'$}$$

Continuing onward then, when we are iven two events, each with 4 coordinates as ascribed by some allowed observer, we want to indicate the differences of their coordinates by Δx, Δy, etc., or Δx^μ in general. In the limit when these two points are very near one another, we may treat this difference as infinitesimal, and denote it by dx, dy, etc., or dx^μ.

Our spacetime, in addition to just having coordinates, also is provided with a notion of "distance," or "length," between pairs of events, often referred to as *the interval.* (This is analogous to the better-known property of length in 3-dimensional spaces, where Pythagoras is usually credited with discovering how best to talk about purely spatial lengths.) In our spacetime, it is usual to denote this quantity by Δs^2, when the two events are well-separated, or in the infinitesimal case, we will refer to it as ds^2. In either case, this is done even though it is not in general the square of anything, any more than the quantity given by Pythagoras' theorem is not the square of any single thing, either. There are, however, two distinct cases when it will be the square of something, which we will identify below.

The physical importance of the interval is that it is measured by all *inertial observers* to have the same value.

Using Minkowski coordinates $\{x, y, z, t\} \equiv \{x^\mu \mid \mu = 1, 2, 3, 4\} \equiv \{x^\mu\}_1^4$ we may write

$$\Delta s^2 = \left\{ \begin{array}{c} (\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2 - (\Delta t)^2 \equiv \eta_{\mu\nu}(\Delta x^\mu)(\Delta x^\nu) \; , \\ \text{and} \\ (\Delta x')^2 + (\Delta y')^2 + (\Delta z')^2 - (\Delta t')^2 \equiv \eta_{\mu\nu}(\Delta x'^\mu)(\Delta x'^\nu) \; , \end{array} \right\} \begin{array}{c} \mu, \nu = 1, 2, 3, 4 \; , \\ \text{and} \\ \text{both summed.} \end{array} \tag{1.2}$$

It is also plausible to consider an infinitesimal version of this quantity when one takes the two events as exceedingly near to one another, i.e., each of the 4 differences involved are very small.

$$\eta_{\mu\nu}(dx^\mu)(dx^\nu) = ds^2 = \eta_{\alpha\beta}(dx'^\alpha)(dx'^\beta) \; . \tag{1.3}$$

Here $\boldsymbol{\eta}$ is a matrix with the following elements, which is of course a very simple extension of the usual 3-dimensional Cartesian metric matrix:

$$\boldsymbol{\eta} = ((\eta_{\mu\nu})) \longrightarrow \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \; . \tag{1.4}$$

From now on we will use double parentheses around the names for the entries of a matrix to indicate that they are to be thought of as the entire set of them, i.e., the entire matrix.

We also try to use the equality arrow symbol, $\Longrightarrow$, to indicate that the values of the entries of a matrix are those on the right-hand side of the arrow when a particular basis for matrix elements has been chosen, and agreed to by all, but not explicitly presented on the page. The language that goes with that arrowed equality is often read as "is presented by."

We also know that matrices have rows and columns; therefore, when we specify the elements of a matrix by a symbol that has two indices we mean that the left index specifies the row and the right index the column. Whether the indices are superscripts or subscripts means something entirely different, which we will discuss belowl

This particular matrix, $\boldsymbol{\eta}$, can best be thought of as the generator of a quadratic sum, and is referred to as **the metric for spacetime**, and plays the role of a scalar product there. The basis for this particular matrix presentation, in Eq. (1.4), is the standard Minkowski/Cartesian one.

Note that Finley uses the so-called "right" sign convention for the metric, where a positive entry in the metric corresponds to a spatial direction and a negative one corresponds to a temporal direction. This approach leads to a so-called *"signature"* of the metric as $+2$, which is simply the sum of the diagonal elements, when it is has been put into a diagonal form. Note that those who do not use this "right" sign convention will have a negative value for the signature, so that this is a convenient way to check on an author's sign convention. The metric matrices in other texts may look rather different if, for instance, that text counts the entries from 0 to 3 instead of 1 to 4, so that when this matrix is displayed it has a -1 in the first row and first column.

Once again it is worthwhile to notice, in Eq. (1.2), that Finley uses the

Einstein summation convention. To repeat, perhaps in more detail this time, this summation convention means that any single term—in a mathematical expression—that contains the same index symbol twice, once as a subscript and once as a superscript is presumed to also contain a sum sign that indicates that a sum is to performed over all allowed values of that index, even though this sum sign is not actually written.

It is worth remembering well that this requires that a correctly written mathematical product of symbols should **not** contain the same index occurring **three times**, since then one would not know which pair it is that is being summed. If one does **absolutely need** the same index three or more times, as, for instance, in an eigenvalue equation with explicitly-presented matrix indices, then after the first two occurrences the others are indicated with the capital-letter version of the lower-case one that indicates the summation.

Back to Physics, now: Since the interval is not positive definite, it can be used to divide spacetime into 6 distinct regions, a division which, upon a choice of an origin, used for one of the events, is also invariant under transformations between inertial observers. (Choosing a first distinct event, and thinking of it as the origin, one may always connect any other event to it by a straight line.) The 6 distinct regions are then described as follows:

1. The origin itself, which is of course just a single event (point), with—presumably—has all 4 of its coordinates as 0.
2. Regions where the interval is negative, relative to the origin, and the quantity Δt is positive. Such events are said to be *future, timelike* relative to the origin, where the word future is used because $\Delta t > 0$, i.e., these events are in the future of observers at the origin.
3. Regions where the interval is negative, relative to the origin, and the quantity Δt is negative. Such events are said to be *past, timelike* relative to the origin.
4. Regions where the interval relative to the origin is positive—referred to as events spacelike relative to the origin.
5. Regions where the interval is zero, relative to the origin, and the quantity Δt is positive. Such events are said to be *future, lightlike (or null)* relative to the origin.

   These events lie on the (3-dimensional) surface of a (3-dimensional) cone that can be thought of as the histories of light rays sent out in all possible directions from a source at the origin. At any given choice of the time the intersection with this cone is a (2-dimensional) sphere.

   The future, timelike region described above as region 2 is the interior of this cone.

6. Regions where the interval is zero, relative to the origin, and the quantity Δt is negative. Such events are said to be *past, lightlike (or null)* relative to the origin.

These events also lie on the (3-dimensional) surface of a (3-dimensional) cone that can be thought of as all the light rays from everywhere in space that are striking the origin, having been sent out at all possible events in the past The interior of this cone is the past, timelike region described above as region 3.

The join of these two cones is usually referred to as *the light cone*, including the origin, while the interior of the two cones is the complete timelike region, the two parts of which are separated by the origin.

When considering two events neither of which is the origin, we may use the same sorts of language, and say that the pair of events is timelike, spacelike, or null-separated. Below we now have a number of additional comments about the different sorts of pairs of events:

Regions 2,3 . When two events are timelike separated, there is always a special allowed observer who measures them to be at the same phsyical location, but only differing in their time coordinates. We therefore define *the proper time*, $\Delta\tau$, between those two events as the difference in time as measured by this special observer. She clearly measures a negative value for the interval between these two events, so that one can write

$$\eta_{\mu\nu}\Delta x^\mu \Delta x^\nu = 0 - (\Delta t^{"})^2 \equiv -\Delta\tau^2 \ , \tag{1.5}$$

where our special observer's measurements are denoted by a double prime, and she measures the spatial differences as zero. As the interval is independent of which observer measures it, any observer can determine the propr time between the events. the interval along that straight line is negative, and the square root of its negative is called the proper time, $\Delta\tau$, between those two events. Its square **maximizes** the negative of the interval, i.e., the square of the length along arbitrary curves between the two points. Picking one event as a base, the proper time from that base along this straight is a physically-reasonable choice for a parameter along this line.

Region 4. When two events are spacelike separated, the interval along that straight line is positive, and its square root is called the proper length, $\Delta\ell$, between those two events:

$$(\Delta\ell)^2 \equiv (\Delta s)^2 = (\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2 - (\Delta t)^2 \ , \tag{1.6}$$

which is of course also an invariant—for spacelike separated pairs of events. Its square **minimizes** the (square of the) length along arbitrary curves between the two events. Picking one event as a base, the length from that base along this straight line is a physically-reasonable choice for a parameter along this line. There is always a special choice of inertial observer who perceives two spacelike separated events to be simultaneous, i.e., to have occurred at the same time as measured in her reference frame.

Regions 5,6. When two events are **null separated**, they both lie on the straight line which is the trajectory of some light ray. This means that the interval between any points on that light ray has value zero, so that there is not a usual definition for length available. Nonetheless, it is often desired to have some sort of indication as to how far along the light ray has traveled, say from the origin. It is therefore reasonable to think of this straight line as a path, or mapping, from some parameter to the points in spacetime. As the parameter increases, the points along the light ray move forward in time; such a parameter is usually referred to as an *affine parameter,* and is not unique, but rather could be subject to any sort of a linear transformation, i.e., a change in the place where it has value zero, and a possible re-scaling, just like any measurements of time.

2. Worldlines and related quantities:

a. The trajectory of any possible observer is the set of all events at which that observer is present. Such trajectories are called *worldlines.* We may easily think of the set of these events as a *path, or curve,* on the spacetime, which can be well described by the use of some single parameter that varies continuously and is ever-increasing along the worldline. The "wristwatch-time" of the observer is of course a very reasonable choice for a parameter along that curve, so we will think of it as the proper time along his trajectory, and often denote it by the Greek symbol τ.

It is useful to think of this curve as a mapping of some range of real numbers, i.e., from some subset of the set of all real numbers, $\mathbb{R}$, into the spacetime. It is then straightforward to think of the curve as having, at each point, a tangent vector that indicates its direction at that point. Since any observer must always travel slower than does light, two nearby points on the worldline will always be timelike separated; therefore, we always suppose that we have chosen a proper scaling and a choice of origin for the observer's "wristwatch-time" so that we may identify it with the locally-measured proper time, τ, at each event through which he lives.

b. Using the proper time as the parameter along the worldline, the tangent vector to the curve, at every point on the curve, is well-defined, and will be referred to as the 4-velocity, at that point, since it is obviously a 4-dimensional vector. That tangent vector should have as components the rate of change, with respect to the proper time, τ, of the coordinates of the events, x^μ, along the worldline. For now we consider the 4-vector which has components, $dx^\mu/d\tau$ (relative to some appropriate-chosen basis for the vector space in which the vector lives), and denote the vector itself by the symbol $d\widetilde{x}/d\tau$. We may then take the ratio of this to the infinitesimal change of proper time in which it occurs; this is surely the desired tangent vector to the worldline, the 4-velocity, which we will call $\widetilde{u}$:

$$\widetilde{u} \equiv \frac{d\widetilde{x}}{d\tau} \ . \tag{2.1}$$

Note carefully that, as in Eq. (2.1) above, Finley always uses an "over-tilde" to note that the object being considered is a vector in 4-dimensions; this is analogous to the standard usage of an "over-arrow," as in $\vec{x}$, to note that the object being considered is a vector in a 3-dimensional space. With this notational idea in mind it is worthwhile to briefly run back and re-write Eq. (1.1) in a notation without so many indices:

$$\widetilde{x}' = \mathbf{\Lambda}\widetilde{x} + \widetilde{a} \ , \tag{1.1$'$}$$

where of course we are considering the components of the vector as being the elements of a matrix with only one column.

If we now rewrite Eqs. (1.5) in infinitesimal form and divide by the scalar $(d\tau)^2$, we acquire a statement that says that

$$\left(\frac{dx}{d\tau}\right)^2 + \left(\frac{dy}{d\tau}\right)^2 + \left(\frac{dz}{d\tau}\right)^2 - \left(\frac{dt}{d\tau}\right)^2 = -1 \ , \tag{2.2a}$$

where the last equality follows from the definition of $d\tau$.

However, since we have used this relation to define a tangent vector, it seems reasonable to **extend** the definition of the **metric tensor** $\boldsymbol{\eta}$. It began as a way to define a quadratic combination of coordinates, on $\mathcal{M}^4$ which was a Lorentz invariant, i.e., a quantity which had the same value when measured by any inertial observer. Now we want to extend its definition so that it also acts as a "scalar product" for any two tangent vectors: let the space of all tangent vectors at the point $P \in \mathcal{M}^4$ be denoted by $\mathcal{T}_P$, and choose two such tangent vectors, $\widetilde{m}$ and $\widetilde{n}$; then this scalar product is just the following sum, where we are following the same "rule" we used earlier for the creation of the interval:

$$\widetilde{m} \cdot \widetilde{n} \equiv m^x n^x + m^y n^y + m^z n^z - m^t n^t = \eta_{\mu\nu}\, m^\mu n^\nu \ . \tag{2.2b}$$

Of course the two vectors could be the same, in which case we have something we will call "the square" of that vector:

$$\widetilde{u}^2 \equiv \widetilde{u} \cdot \widetilde{u} = (u^x)^2 + (u^y)^2 + (u^z)^2 - (u^t)^2 \equiv \eta_{\mu\nu}\, u^\mu u^\nu \ . \tag{2.2c}$$

We can also use these notations for 4-vectors, and 4×4 matrices to re-write the above in a useful way:

$$\widetilde{x}^2 = \widetilde{x}^T \boldsymbol{\eta} \widetilde{x} = \widetilde{x'}^T \boldsymbol{\eta} \widetilde{x}' \tag{2.2c'}$$

In this method of writing we recall that the superscript ^{T} indicates the transpose of a matrix. In the case of a matrix that is a column vector, i.e., only one column, the transpose turns it into a row vector, which has only one row. In the index notation, one can understand these by remembering that matrix multiplication requires a sum of the column indices for the matrix on the left with the row indices of the matrix on the right, and that for a square matrix the left-most index is the one for the rows, with the right-hand index for the columns.

Just a little earlier we extended the actions of the metric to the set/space of all tangent vectors; it is therefore worthwhile to remember that every vector space has associated with a *dual vector space*, which is the set of all linear functions that map those vectors into scalars, which is also a vector space. Let us denote the set of all these vectors, at the point $P \in \mathcal{M}^4$, by the symbol Λ^1_P, and write such (dual) vectors in the form $\underset{\sim}{\alpha}$ where the "under-tilde" reminds us that this is a dual vector, often referred to as a differential form—which will also be a subject for further study soon. As it is a dual vector it follows that $\underset{\sim}{\alpha}(\widetilde{m})$ is a scalar. Since we are understanding vectors via their components, and presenting them as elements of matrices, it is reasonable to suppose that these dual spaces also have basis vectors, and we will write the components as α_μ, so that we may have all the following sorts of presentations for these scalars:

$$\underset{\sim}{\alpha}(\widetilde{m}) = \alpha_\mu m^\mu = \alpha^T m \; ,$$

where in the last quantity we understand α as the column vector/matrix with the components of the differential form, analogous to the symbol m which is the column vector withe the components m^μ, and the upper- and lower-placement of the indices agrees with the Einstein summation convention.

c. Dynamical physics then allows the introduction of some important mechanical quantities, which are related to this tangent vector, when it is tangent to the worldline of a particle of mass m:

   i.) the *4-momentum* vector,

$$\widetilde{p} \equiv m\,\widetilde{u} \equiv \begin{pmatrix} \vec{p} \\ E \end{pmatrix} \; , \tag{2.3}$$

   where this is a bit of a generalization of the Newtonian version of the 3-momentum, but this definition for $\vec{p}$ has the property that it is the quantity which is conserved when expected. Likewise, this E is a generalization of the Newtonian concept of the total energy of the particle, but is indeed the correct one.

ii.) the (net) 4-force,

$$\widetilde{K} \equiv d\widetilde{p}/d\tau = \frac{d}{d\tau}\begin{pmatrix} \vec{p} \\ E \end{pmatrix} \; . \tag{2.4}$$

This includes appropriate generalizations of the usual 3-vector force and also the power, i.e., the time-rate of change of the energy.

iii.) the *4-acceleration* is the proper time derivative of the 4-velocity, i.e.,

$$\widetilde{a} \equiv d\widetilde{u}/d\tau = \frac{d}{d\tau}\gamma_u \begin{pmatrix} \vec{u} \\ 1 \end{pmatrix} \; . \tag{2.5}$$

This is related to the usual 3-acceleration in a fairly complicated way, but in just such a way that the 4-acceleration is indeed a 4-vector.

However, there is an additional interesting feature possessed by the 4-acceleration, because the Lorentz square of the 4-velocity is a constant. It therefore follows that the 4-velocity and the 4-acceleration are always perpendicular:

$$0 = \frac{d}{d\tau}(-1) = \frac{d}{d\tau}(\widetilde{u}\cdot\widetilde{u}) = 2\widetilde{u}\cdot\widetilde{a} \; . \tag{2.6}$$

4. This is also a good occasion to introduce the idea of *the co-moving, inertial reference frame* for an accelerated particle. As measured by an arbitrary observer, at any given time, t, the particle has 3-velocity $\vec{u}(t)$, along with its 4-velocity, $\widetilde{u}(\tau)$, and 3-acceleration, $\vec{a}(t)$. As a function of that time, there is always a true inertial reference frame that has always been—and always will be—moving at velocity $\vec{u}(t)$, and—at that instant— this reference frame will measure our particle to be at rest, i.e., the measurement of the particle's 4-velocity made by the current co-moving inertial observer is just $\widetilde{u} = (\vec{0}, 1)^T$. Therefore, because of the orthogonality of these two 4-vectors, the co-moving observer's measurement of the particle's 4-acceleration is $\widetilde{a} = (\vec{a}_0, 0)^T$, where the subscript 0 reminds us that this is the 3-acceleration as measured by the instantaneous co-moving (inertial) frame. It is this particular 3-acceleration that is referred to as the particle's *instantaneous 3-acceleration*

in the rest frame. It is, nonetheless, quite simple to measure because the Lorentz square of the 4-acceleration is just

$$\widetilde{a} \cdot \widetilde{a} = (\vec{a}_0)^2 \ , \tag{2.7}$$

which is of course an invariant, and therefore measureable in any inertial frame whatsoever. It also makes it clear that the 4-acceleration is a spacelike vector, provided of course that it is non-zero. For interest, we now skip ahead a bit in our discussion, and use the Lorentz transformation that takes the instantaneous, co-moving frame back to the original frame—where the velocity was $\vec{u}$. This transformation is explained in detail in a later section on Lorentz transformations. Using those equations, we find that the 4-acceleration in that frame is given by

$$\widetilde{a} \Longrightarrow \begin{pmatrix} \vec{a}_{0\perp} + \gamma_u \vec{a}_{0\parallel} \\ \gamma_u \vec{u} \cdot \vec{a}_0 \end{pmatrix} \ , \tag{2.8}$$

where of course $\vec{a}_0$ above is the instantaneous acceleration in the rest frame, while the subscript $\parallel$ picks out that portion of $\vec{a}_0$ that is parallel to $\vec{u}$, and the subscript $\perp$ that portion which is perpendicular.

It would also be nice to know how this relates to the ordinary 3-acceleration. To answer this question we need to determine the relation between the 4-acceleration and the ordinary 3-acceleration, in general. We first note that

$$\widetilde{u} \equiv \frac{d}{d\tau}\widetilde{r} \equiv \frac{d}{d\tau}\begin{pmatrix} \vec{r} \\ t \end{pmatrix} = \gamma_u \begin{pmatrix} \vec{u} \\ 1 \end{pmatrix} \ . \tag{2.9}$$

From this we say that the 4-acceleration, $\widetilde{a}$ is given by

$$\widetilde{a} \equiv \frac{d}{d\tau}\widetilde{u} = \gamma_u \frac{d}{dt}\gamma_u \begin{pmatrix} \vec{u} \\ 1 \end{pmatrix} \ . \tag{2.10}$$

It is now helpful to know the derivative of γ_u:

$$\frac{d}{dt}\gamma_u = -\tfrac{1}{2}\frac{1}{(1 - \vec{u} \cdot \vec{u})^{3/2}}[-2\vec{u} \cdot \vec{a}] = \gamma_u^3 \vec{u} \cdot \vec{a} \ . \tag{2.11}$$

Now we have enough information to evaluate the derivatives in Eq. (2.10):

$$\widetilde{a} = \gamma_u^2 \begin{pmatrix} \vec{a} \\ 0 \end{pmatrix} + \gamma_u^4 \vec{u}\cdot\vec{a} \begin{pmatrix} \vec{u} \\ 1 \end{pmatrix} = \gamma_u^2 \begin{pmatrix} \vec{a}_\perp + \gamma_u^2 \vec{a}_\parallel \\ \gamma_u^2 \vec{u}\cdot\vec{a} \end{pmatrix} , \tag{2.12}$$

where we have divided $\vec{a}$ into its perpendicular and parallel parts, and noted that $\gamma_u^2 u^2 + 1$ is just γ_u^2. Next we compare Eq. (2.12) with Eq. (2.8), noting that only the parallel parts of each of $\vec{a}$ and $\vec{a}_0$ occur in the 4th component, giving us a relation between those two parallel parts. Likewise, the perpendicular parts appear more straightforwardly, so that the result is that the 3-acceleration, $\vec{a}$, in the frame where the velocity is $\vec{u}$—at that particular time—is related to the 3-acceleration in the instantaneous rest frame in the following simple way:

$$\vec{a}_\perp = \gamma_u^{-2} \vec{a}_{0\perp} , \qquad \vec{a}_\parallel = \gamma_u^{-3} \vec{a}_{0\parallel} . \tag{2.13}$$

The components of these quantities, i.e., all the various 4-vectors, all transform in the same way as do coordinate differences when one changes basis from one (inertial) observer to another, which is the meaning of the statement that they are 4-vectors.

3. Lorentz transformations for 4-vectors

Referring back to Eqs. (1.1), we agree that our two inertial observers will have agreed on the same event for an origin for their coordinates, which causes the a^μ indicated in that equation to vanish. We then want to consider all possible options for the Lorentz transformations given by the matrices $\Lambda^\mu{}_\alpha$. We rewrite that equation here, thinking of it in terms of coordinate differences:

$$\Delta x^\mu \equiv \Lambda^\mu{}_\alpha \Delta x'^\alpha \quad \text{or} \quad \widetilde{\Delta x} = \Lambda\, \widetilde{\Delta x'} , \tag{3.1}$$

where the second form given above is the same equation in *matrix notation*, and we distinguish those matrices which are just column vectors by inserting an "overtilde" on them. Then, continuing onward to Eqs. (1.8), which tells us that the interval between two events is the same for all inertial observers, we now write down that fact, using the matrix form:

$$\widetilde{\Delta x'}^T \boldsymbol{\eta} \widetilde{\Delta x'} = (\Delta s')^2 = (\Delta s)^2 = \widetilde{\Delta x}^T \boldsymbol{\eta} \widetilde{\Delta x} = (\Lambda\, \widetilde{\Delta x'})^T \boldsymbol{\eta} (\Lambda\, \widetilde{\Delta x'}) = \widetilde{\Delta x'}^T \Lambda^T \boldsymbol{\eta} \Lambda \widetilde{\Delta x'} . \tag{3.2}$$

Since this is true for arbitrary value of $\Delta x'$, it follows that we may simply require the two matrices inside—at opposite ends of the string of equalities—to be equal:

$$\boldsymbol{\eta} = \Lambda^T \boldsymbol{\eta} \Lambda \quad \text{or} \quad \eta_{\alpha\beta} = \Lambda^\mu{}_\alpha \eta_{\mu\nu} \Lambda^\nu{}_\beta = \eta_{\mu\nu} \Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta \ . \tag{3.3}$$

The last equality above is just the statement that the (indicial) entries have the rules about multiplication built into the locations and types of their indices, so that the order of writing down the kernel letters is irrelevant. This is unlike matrices, which have rules about the order in which multiplication takes place built into the order in which they are written. Note that in the index form of the matrix equations the fact that the matrix Λ on the left-hand side was transposed simply means that in the indicial form it is the values of the row index are the values of the column index of the transposed matrix, and are then being (correctly) summed into a row index on the following matrix. As well what are the column indices for Λ are of course the row indices for Λ^T, and therefore do correctly correspond to the row indices on the $\boldsymbol{\eta}$ on the other side of the equality.

Equation (3.3) is then the defining requirement for a Lorentz transformation, since it is the requirement that the interval be left invariant under transformations between inertial reference frames. On a fairly general level, one may see how the inverse is related to the original one via the following algebraic manipulations, that begin with the matrix form of Eqs. (3.3):

$$\begin{gathered} \boldsymbol{\eta} = \Lambda^T \boldsymbol{\eta} \Lambda \quad \Longrightarrow \quad \boldsymbol{\eta}\Lambda^{-1} = \Lambda^T \boldsymbol{\eta} \quad \Longrightarrow \quad \Lambda^{-1} = \boldsymbol{\eta}^{-1} \Lambda^T \boldsymbol{\eta} \\ \text{or, in indicial notation,} \quad (\Lambda^{-1})^\alpha{}_\mu = \eta^{\alpha\beta} \Lambda^\nu{}_\beta \eta_{\nu\mu} \ , \end{gathered} \tag{3.4}$$

where, since the matrix $\boldsymbol{\eta}^{-1}$ is identical to the matrix $\boldsymbol{\eta}$, the rather lengthy calculation of the inverse of an arbitrary 4×4 matrix is not required.

There are a couple of other useful pieces of information we may obtain from the defining equation, Eqs. (3.3). First, let us take the determinant of both sides of the equation, remembering that the determinant of a product of matrices is just the product of the determinants of those matrices, and that the determinant of a transposed matrix is the same as the one that was not transposed:

$$\det(\boldsymbol{\eta}) = \det(\Lambda)^2 \det(\boldsymbol{\eta}) \quad \Longrightarrow \quad \det(\Lambda)^2 = +1 \; . \tag{3.5}$$

Therefore, there are two independent "pieces" of Lorentz transformations, those with determinant +1 and those with determinant −1, with no other options since the elements of the matrix are surely real-valued.

The second relation comes from taking the 4, 4-element of the matrix equation that defines Lorentz transformations:

$$\begin{aligned} -1 = \eta_{44} = & \Lambda^\mu{}_4 \, \eta_{\mu\nu} \, \Lambda^\nu{}_4 = -(\Lambda^4{}_4)^2 + \sum_{m=1}^{3} (\Lambda^m{}_4)^2 \\ & \Longrightarrow \quad (\Lambda^4{}_4)^2 = 1 + \sum_{m=1}^{3} (\Lambda^m{}_4)^2 \geq 1 \; . \end{aligned} \tag{3.6}$$

Since, again, the square roots of this equation, to determine $\Lambda^4{}_4$, are real, it follows that either we have

$$\Lambda^4{}_4 \geq +1 \qquad \text{or} \qquad \Lambda^4{}_4 \leq -1 \; , \tag{3.7}$$

and the two sets of them do not overlap. Since $\Lambda^4{}_4$ determines the relation between times in the two different observers, we say that those transformations with positive values for $\Lambda^4{}_4$ are called *orthochronous*, since they preserve the sign of the time.

We have now divided the set of all Lorentz transformation matrices into 4 distinct, non-overlapping sets, depending on the sign of the determinant and on the sign of $\Lambda^4{}_4$. We note that the identity matrix is of course a Lorentz transformation, so that it is in the set called *the special, orthochronous Lorentz group*, i.e., those with determinant and $\Lambda^4{}_4$ both positive. It is common to refer to this set by some shorter set of symbols, so that we will sometimes refer to it as $\mathrm{SO}(3,1)^\uparrow$, where the S refers to the fact that it is special, i.e., has determinant +1, the upward-pointing arrow tells us that it is maintaining the direction of time, and the O(3,1) tells us that they are a set of matrices that "preserve" the 4-dimensional metric in question, which has 3 plus signs and 1 minus sign in its diagonal form.

Each of these four sets has a particular member which plays a role something like the identity matrix, $\mathbf{I_4}$, in the sense that it is the matrix closest to an identity that lies in that particular set of matrices. We write all 4 of those generating matrices just below:

$$\mathbf{I}_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} , \quad \mathcal{P} \equiv \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} ,$$
$$\mathcal{T} \equiv \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} , \quad \mathcal{I} \equiv \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} = \mathcal{P}\mathcal{T} = -\mathbf{I}_4 , \tag{3.8}$$

where $\mathcal{P}$, called *parity*, changes the sign of all the spatial dimensions, and has negative determinant, $\mathcal{T}$, called *time reversal*, changes the sign of time, and also has negative determinant, while $\mathcal{I}$, is the complete negative of the identity, and therefore changes the signs of every coordinate.

We will spend the majority of our time investigating those transformations that are in the same set as the identity, especially since all the elements of the other 3 sets may be determined simply by multiplying the ones in $\mathrm{SO}(3,1)^{\uparrow}$ by their respective generators. There are several distinct, useful ways to divide this set, which we will comment on. The first one we will discuss is that any Lorentz transformation in $\mathrm{SO}(3,1)^{\uparrow}$ may be written as the product of some (ordinary, 3-dimensional) rotation and a boost, that takes us between two reference frames without re-orienting any of the directions.

Rotations in 3 spatial dimensions:

Since our 4-dimensional metric is just the identity in its 3 spatial dimensions, the usual requirement for rotations is that they should be orthogonal, i.e., that some matrix R is a rotation provided $R^T \mathbf{I}_3 R = \mathbf{I}_3$, which is the same as saying that $R^T = R^{-1}$. It also follows that the determinant of a rotation must be ± 1. Those that also have determinant $+1$ are said to be members of SO(3), and are the usual matrices about which we already know quite a bit. Since we are interested in them as a subset of all possible Lorentz transformations, we will write them as 4×4 matrices, but with the 4th row and 4th column just $(0, 0, 0, 1)$ or its transpose, which assures us that these rotations do not mix space and time variables. Therefore, the standard approach is to write, for instance, a rotation about the $\hat{z}$-axis by an angle η as follows, which I assume you remember:

$$R(\hat{z}; \eta) \Longrightarrow \begin{pmatrix} \cos\eta & -\sin\eta & 0 & 0 \\ \sin\eta & \cos\eta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} . \tag{3.9}$$

More generally, we want to think of rotations as being generated by some matrices that allow us to begin with rotations of zero angle, i.e., the identity transformation, and then build up to larger angles. Therefore, we introduce a triplet of 4×4 matrices—each of course not mixing space and time—$\{\mathfrak{J}^x, \mathfrak{J}^y, \mathfrak{J}^z\}$ that may also be thought of as the components of a 3-vector—although a 3-vector that has components that are matrices, either 3×3 or, as we are looking on them, as 4×4 matrices. Therefore the symbols $(\mathfrak{J}^a)^\mu{}_\alpha$ indicate the elements in the μ-row and the α-column of the matrix which is the rotation generator (angular momentum vector) in the direction of the a-th basis vector. We may then let $\hat{n} \equiv \sin\theta(\cos\varphi\,\hat{x} + \sin\varphi\,\hat{y}) + \cos\theta\,\hat{z}$ be an arbitrary unit 3-vector, using the usual spherical coordinates to define it, and then define these matrices so that they behave as we want, as the (infinitesimal) generators for rotations:

$$R(\hat{n}; \eta) \equiv e^{\eta\hat{n}\cdot\vec{\mathfrak{J}}} \equiv \mathbf{I}_4 + \eta\hat{n}\cdot\vec{\mathfrak{J}} + \tfrac{1}{2}\eta^2 \left(\hat{n}\cdot\vec{\mathfrak{J}}\right)^2 + \ \ldots \ ,$$

$$\hat{x}\cdot\vec{\mathfrak{J}} \equiv \mathfrak{J}^x \Longrightarrow \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & +1 & 0 & 0 \\ 0 & 0 & 0 & +1 \end{pmatrix} , \quad \hat{y}\cdot\vec{\mathfrak{J}} \equiv \mathfrak{J}^y \Longrightarrow \begin{pmatrix} 0 & 0 & +1 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & +1 \end{pmatrix} ,$$

$$\hat{z}\cdot\vec{\mathfrak{J}} \equiv \mathfrak{J}^z \Longrightarrow \begin{pmatrix} 0 & -1 & 0 & 0 \\ +1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & +1 \end{pmatrix} ,$$

$$\hat{n}\cdot\vec{\mathfrak{J}} \Longrightarrow \begin{pmatrix} 0 & -\cos\theta & +\sin\theta\sin\varphi & 0 \\ +\cos\theta & 0 & -\sin\theta\cos\varphi & 0 \\ -\sin\theta\sin\varphi & +\sin\theta\cos\varphi & 0 & 0 \\ 0 & 0 & 0 & +1 \end{pmatrix} . \tag{3.10}$$

With $\vec{a}$ and $\vec{b}$ two arbitrary 3-vectors, we find the useful identities:

$$(\hat{a}\cdot\vec{\mathfrak{J}})\vec{b} = \hat{a}\times\vec{b}\ , \quad (\hat{a}\cdot\vec{\mathfrak{J}})^3 = -(\hat{a}\cdot\vec{\mathfrak{J}})\ , \quad [(\hat{a}\cdot\vec{\mathfrak{J}}),(\hat{b}\cdot\vec{\mathfrak{J}})] = (\hat{a}\times\hat{b})\cdot\vec{\mathfrak{J}}\ . \tag{3.11}$$

where it is useful to pay attention to the fact that when we write $\hat{a}\cdot\vec{\mathfrak{J}}$ we are treating the 3 distinct components of $\vec{\mathfrak{J}}$ as the components of a 3-vector, so that this means

$$\hat{a}\cdot\vec{\mathfrak{J}} = a_x\mathfrak{J}^x + a_y\mathfrak{J}^y + a_z\mathfrak{J}^z\ , \tag{3.12}$$

so that this is the sum of three matrices, which is of course then a particular single matrix. Then, when we write $(\hat{a}\cdot\vec{\mathfrak{J}})\vec{b}$ we are taking that single matrix created above and multiplying it, acting on the left on the matrix $\vec{b}$, being treated as (the components of) a vector, with the result being some different vector.

Wondering about the action of rotations on this vector with elements that are matrices, we can resolve that by letting R be some other, arbitrary rotation, and noting the following results:

$$R^{-1}\vec{\mathfrak{J}}R = R\vec{\mathfrak{J}} \ \text{ or, in indicial notation} (R^{-1})^{\mu}{}_{\alpha}(\mathfrak{J}^b)^{\alpha}{}_{\beta}R^{\beta}{}_{\nu} = (R^b{}_a\mathfrak{J}^a)^{\mu}{}_{\nu}\ , \tag{3.13}$$

where, in the indicial notation approach, the left-hand side of the equation has R^{-1} and also R acting on the angular momentum generators relative to their behavior as matrices, i.e., creating an equivalence transformation, while the right-hand side of the equation has the matrix R acting as a matrix acting on the components of the 3-vector, as would be expected for a transformation to different coordinates for a vector. The meaning of the equality is that these two different sorts of actions give exactly the same result! Also note that the very last equality in Eqs. (3.9), namely the commutator of angular-momentum matrices in two different directions, can also be written in the following form, which may (or may not) be more familiar to you:

$$[\mathfrak{J}^a,\mathfrak{J}^b] = \epsilon^{ab}{}_c\,\mathfrak{J}^c\ , \qquad a,b,c, = 1,2,3\ , \tag{3.14}$$

where $\epsilon^{ab}{}_c$ is a way of writing the Levi-Civita symbol, which is totally skew-symmetric on all 3 indices, and has $\epsilon^{123} = +1$, and the placement of the indices, upper or lower, is irrelevant for the Levi-Civita symbol, but is written in such a way as to ensure the correct action of the Einstein summation convention.

Boosts:

Again I assume that you know how to write a transformation of coordinates between two observers who have parallel spatial axes, and coordinated origins, but who are moving with some non-zero, constant relative velocity—in a single particular direction, which I will pick as the $\hat{z}$-direction. The 4×4 matrix which effects this is given by

$$L(\hat{z}; v) \Longrightarrow \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \gamma_v & \gamma_v v \\ 0 & 0 & \gamma_v v & \gamma_v \end{pmatrix} , \qquad (3.15a)$$

where this supposes that there are two observers, which we describe as $\mathcal{S}$ and $\mathcal{S}'$, and acknowledge that $\mathcal{S}$ measures $\mathcal{S}'$ to be traveling with speed v in the positive $\hat{z}$-direction. (Also see the similar set of statements in Grøn, p. 70.) Then if $\mathcal{S}'$ makes measurements of the components of some 4-vector, say $\widetilde{a}\,'$, then the action of $L(\hat{z}; v)$ on the components of that 4-vector gives the measurements of the values of $\widetilde{a}$ as made by $\mathcal{S}$, i.e., we have

$$a^\mu = L^\mu{}_\alpha a'^\alpha . \qquad (3.15b)$$

It should of course be immediately noticed that $\mathcal{S}'$ measures $\mathcal{S}$ to be moving with speed v in the negative $\hat{z}$-direction, so that the inverse of the matrix above is obtained simply by changing v to $-v$.

Now let us obtain the general form for an arbitrary boost between frames moving with some arbitrary velocity $\vec{v}$, one as measured by the other. We find it convenient to write down the behavior under this boost in terms of what happens to the spatial coordinates parallel to and perpendicular to the velocity between the two frames:

$$\begin{gathered} \vec{r} = \vec{r}_\parallel + \vec{r}_\perp , \\ \vec{r}_\parallel = (\hat{v} \cdot \vec{r})\hat{v} , \qquad \vec{r}_\perp = \vec{r} - \vec{r}_\parallel . \end{gathered} \qquad (3.16)$$

The appropriate Lorentz transformation equations for the location vector are then

$$\begin{aligned} &\vec{r}_{\|} = \gamma[r'_{\|} + t'\vec{v}] \,, \qquad \vec{r}_{\perp} = \vec{r}'_{\perp} \,, \qquad t = \gamma[t' + \vec{v}\cdot\vec{r}'/c^2] \,, \\ \Longrightarrow \quad &\vec{r} = \gamma[(\hat{v}\cdot r')\hat{v} + t'\vec{v}] + [r' - (r'\cdot\hat{v})\hat{v}] = r' + (\gamma-1)(r'\cdot\hat{v})\hat{v} + \gamma t'\vec{v} \,. \end{aligned} \tag{3.17}$$

Therefore a general boost matrix has the form, $L(\vec{v})$:

$$\begin{pmatrix} \vec{r} \\ ct \end{pmatrix} \equiv \widetilde{r} = L(\vec{v})\widetilde{r}' = L(\vec{v}) \begin{pmatrix} r' \\ ct' \end{pmatrix} \,,$$

$$\begin{aligned} L(\vec{v}) &= \begin{pmatrix} \mathbf{I}_3 + (\gamma-1)\hat{v}\hat{v}^T & \gamma\vec{v}/c \\ \gamma\vec{v}^T/c & \gamma \end{pmatrix} \\ &= \begin{pmatrix} 1+(\gamma-1)(\hat{v})_x(\hat{v})_x & (\gamma-1)(\hat{v})_x(\hat{v})_y & (\gamma-1)(\hat{v})_x(\hat{v})_z & \gamma\, v_x/c \\ (\gamma-1)(\hat{v})_x(\hat{v})_y & 1+(\gamma-1)(\hat{v})_y(\hat{v})_y & (\gamma-1)(\hat{v})_y(\hat{v})_z & \gamma\, v_y/c \\ (\gamma-1)(\hat{v})_x(\hat{v})_z & (\gamma-1)(\hat{v})_y(\hat{v})_z & 1+(\gamma-1)(\hat{v})_z(\hat{v})_z & \gamma\, v_z/c \\ \gamma\, v_x/c & \gamma\, v_y/c & \gamma\, v_z/c & \gamma \end{pmatrix} . \end{aligned} \tag{3.18}$$

Again the inverse is obtained just by switching the sign of $\vec{v}$, which means the entries in the 4th column and 4th row, except for the 4, 4-element.

It is useful to take the same "infinitesimal" approach to these matrices as we did with the rotations. We therefore define a set of 3 4×4 matrices, $\{\boldsymbol{\mathfrak{K}}^a \mid a = 1,2,3\}$:

$$\begin{aligned} \hat{x}\cdot\vec{\boldsymbol{\mathfrak{K}}} \equiv K^x \Longrightarrow \begin{pmatrix} 0&0&0&1\\0&0&0&0\\0&0&0&0\\1&0&0&0 \end{pmatrix} , \quad \hat{y}\cdot\vec{\boldsymbol{\mathfrak{K}}} \equiv K^y \Longrightarrow \begin{pmatrix} 0&0&0&0\\0&0&0&1\\0&0&0&0\\0&1&0&0 \end{pmatrix} , \\ \hat{z}\cdot\vec{\boldsymbol{\mathfrak{K}}} \equiv K^z \Longrightarrow \begin{pmatrix} 0&0&0&0\\0&0&0&0\\0&0&0&1\\0&0&1&0 \end{pmatrix} , \end{aligned} \tag{3.19}$$

which have the property that

$$\begin{aligned} &L(v\hat{v}) = e^{\lambda\hat{v}\cdot\vec{\boldsymbol{\mathfrak{K}}}} = \mathbf{I}_4 + \lambda\hat{v}\cdot\vec{\boldsymbol{\mathfrak{K}}} + \ \ldots \ , \\ &v = \tanh\lambda \,, \qquad \gamma_v = \cosh\lambda \,, \ v\gamma_v = \sinh\lambda \,. \end{aligned} \tag{3.20}$$

Analogous to the relations that the matrices $\vec{\boldsymbol{\mathfrak{J}}}$ satisfy, as given in Eqs. (3.9), we may write that

$$(\hat{a}\cdot\vec{\boldsymbol{\mathfrak{K}}})^3 = +(\hat{a}\cdot\vec{\boldsymbol{\mathfrak{K}}}) \,, \quad [\hat{a}\cdot\vec{\boldsymbol{\mathfrak{K}}}, \hat{b}\cdot\vec{\boldsymbol{\mathfrak{K}}}] = -(\hat{a}\times\hat{b})\cdot\vec{\boldsymbol{\mathfrak{J}}} \,, \quad [\hat{a}\cdot\vec{\boldsymbol{\mathfrak{J}}}, \hat{b}\cdot\vec{\boldsymbol{\mathfrak{K}}}] = (\hat{a}\times\hat{b})\cdot\vec{\boldsymbol{\mathfrak{K}}} \tag{3.21}$$

We recall that the commutators of two of the rotation generators is another rotation, while, quite differently, the commutator of two of the boost generators is a related rotation. These are important differences which we will discuss later on. The most important fact that results from this is that the product of two boosts, in different directions, is not just a boost in some third direction, as is indeed the case for rotations.