

Notes and Conventions for Vector and Tensor Analysis:

Physics 495, *An Introduction to Special Relativity*

by Daniel Finley, Fall, 2001

I. The Spacetime of Special Relativity

1. **Spacetime** is a 4-dimensional *manifold*, 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 $(x, y, z+t, z-t)$. Different *allowed observers* will ascribe these coordinates in different ways. The allowed observers in spacetime are often referred to as *inertial observers*. (More will be said later about a manifold and such labelling of points on it by the use of coordinates.) Given two such points the differences of their coordinates may be indicated by Δx , Δy , etc. In the limit when these two points approach one another, we may treat this difference as infinitesimal, and denote it by dx , dy , etc.
- a. Spacetime also is provided with a notion of “distance,” or “length,” between pairs of events, referred to as *the interval*, Δs^2 . This interval is measured to have the same value by all *inertial observers*:

$$\Delta s^2 = (\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2 - (\Delta t)^2 = (\Delta x')^2 + (\Delta y')^2 + (\Delta z')^2 - (\Delta t')^2 . \quad (1.1)$$

One may think about the interval as a quadratic sum of squares (of differences) of coordinates. The function that generates that quadratic sum is referred to as the metric for spacetime, and plays the role of a scalar product. From the form shown above one sees that Finley uses the usual (“right”) sign convention for the metric, where a +1 in the metric corresponds to a spatial direction, and -1 to a temporal direction. It is also seen that Finley uses “geometrized units,” where the speed of light, c , has its value set equal to +1, which makes the SI units of meters and seconds interconvertible, and spatial and temporal coordinates have the same “dimensions,” either meters or seconds. This approach emphasizes the underlying geometry, and the validity of the meaning of c for **all** types of phenomena. Its more usual value is $c = 2.99792458 \times 10^8$ m/sec.

In infinitesimal form, where we use the symbol \mathbf{g} as a generic symbol for a metric, we have the following form for a (flat-space) metric in Minkowski (Cartesian) coordinates, $\{x, y, z, t\} \equiv \{x^\mu \mid \mu = 1, 2, 3, 4\} \equiv \{x^\mu\}_1^4$:

$$\mathbf{g} = dx^2 + dy^2 + dz^2 - dt^2 \equiv \eta_{\mu\nu} dx^\mu dx^\nu, \quad \mu, \nu = 1, 2, 3, 4. \quad (1.2)$$

Here we notice that Finley's conventions are to order the coordinates so that indices $\{1, 2, 3\}$ are spatial, with time being the coordinate with index 4, so that it comes last in the sequence of coordinates. We also notice that Finley uses the **Einstein summation convention**, which says that

any single term that contains the same index symbol twice, once as a subscript and once as a superscript is presumed to also contain (an unwritten) sum sign that indicates that a sum is to be performed over all allowed values of that index.

This means that a correctly written mathematical product of symbols should **not** contain the same index occurring **three times**. I also use the convention that 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. Lastly, we can see that Finley uses Greek, lower-case letters to take values from 1 to 4, and Roman, lower-case letters to take values only from 1 to 3, representing the spatial portions of some otherwise 4-dimensional object. He also sometimes uses Roman, lower-case letters to simply indicate indices that take values from 1 to some yet-unspecified integer value, m , **and** uses Roman, upper-case letters to indicate indices on 2x2 matrices [or the associated 2-dimensional vectors], which then run from 1 to 2.

- b. The set of inequivalent (allowed) inertial observers is labelled by the Poincaré transformation, which transforms the basis of the “standard” observer into that of some other one. Poincaré transformations include all rotations, Lorentz boosts, and 4-dimensional

translations, and any of their products, the set of such being 10-dimensional. We denote an arbitrary rotation by angle θ about an axis (of unit length), \hat{e} , by $R(\theta; \hat{e})$, and a Lorentz boost from an observer moving with velocity \vec{v} to a frame that measures that observer to be moving with that velocity by the symbol $B(\vec{v})$. The generic Poincaré transformation then moves from inertial reference frame S' , who uses coordinates \tilde{x}' , to inertial frame S , who uses coordinates \tilde{x} , via

$$\tilde{x} = R(\theta; \hat{e})B(\vec{v})\tilde{x}' + \tilde{a} \equiv P\{L(\theta; \hat{e}; \vec{v}); \tilde{a}\}\tilde{x}' , \quad (1.3)$$

where \tilde{a} is an arbitrary 4-vector (relating the choices of origins of the two observers), and one notices that we are using an “overtilde” to denote the (column-)matrix presentation of the coordinates of each observer as a matrix.

Note that the set of such labellings, i.e., Poincaré transformations, forms a group:

$$\begin{aligned} P\{L_1; \tilde{a}_1\}P\{L_2; \tilde{a}_2\} &\equiv P\{L_1L_2; \tilde{a}_1 + L_1\tilde{a}_2\} , \quad \text{definition of the product} \\ P\{L_1; \tilde{a}_1\}\{P\{L_2; \tilde{a}_2\}P\{L_3; \tilde{a}_3\}\} &= P\{L_1; \tilde{a}_1\}P\{L_2L_3; \tilde{a}_2 + L_2\tilde{a}_3\} \\ &= P\{L_1L_2L_3; \tilde{a}_1 + L_1\tilde{a}_2 + L_1L_2\tilde{a}_3\} , \quad \text{associativity of the product} \\ &= P\{L_1L_2; \tilde{a}_1 + L_1\tilde{a}_2\}P\{L_3; \tilde{a}_3\} = \{P\{L_1; \tilde{a}_1\}P\{L_2; \tilde{a}_2\}\}P\{L_3; \tilde{a}_3\} \\ P\{I_4; \tilde{0}\}P\{L_1; \tilde{a}_1\} &= P\{L_1; \tilde{a}_1\} = P\{L_1; \tilde{a}_1\}P\{I_4; \tilde{0}\} , \quad \text{an identity} \\ P\{L^{-1}; -L^{-1}\tilde{a}\}P\{L; \tilde{a}\} &= P\{I_4; \tilde{0}\} = P\{L; \tilde{a}\}P\{L^{-1}; -L^{-1}\tilde{a}\} , \quad \text{an inverse .} \end{aligned}$$

I also recall (?) that the definition of a *group* is that it is a set of objects, G , along with a definition of a product of two such objects that has the properties presented above, namely that the result of the product should also be a member of the group, that the product should be associative (relative to triple products), that the set should contain an identity element—where the product with it leaves any other element alone—and that each element should have an inverse element—such that their product gives the identity element—that is also contained within the set. More on this later, and also on the question of presentation of vectors as matrices.

- c. Two distinct events may always be connected by a straight line.

We may then say that if they are

- i.) 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. Its square **minimizes** the (square of the) length along arbitrary curves between the two events; or if they are
- ii.) timelike separated, 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 squared) length along arbitrary curves between the two points; or if they are
- iii.) null separated, they both lie on the trajectory of some light ray.

The fact that the spacetime admits the interval, which allows the above statements, allows each observer to make a division of all displacement vectors, relative to her or his origin, into the (past) and (future) lightcones, the (past) and (future) timelike parts, and the spacelike part; this division is (of course) independent of which observer makes the measurements.

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*, and must have everywhere-timelike tangent vectors. Treating any continuous curve in spacetime, such as one of these trajectories, as a mapping from some one variable to the spacetime, we may use the fact that worldlines have everywhere-timelike tangent vectors to use the proper time, τ , along that curve to parametrize. (Clearly such a parametrization is non-unique in the sense that one must choose an origin and a scale for τ ; a more mathematically-oriented statement is that any choice from the set $a\tau + b$, for constant a and b is as acceptable a choice for a parameter as any other.)

Along such a trajectory, we may therefore talk about displacement vectors, $\Delta\tilde{x}$, or the appropriate infinitesimal version, divided by the intervening proper time:

$$\text{4-velocity} \equiv \tilde{u} \equiv \frac{d}{d\tau}\tilde{x}. \quad (1.4)$$

This then allows the introduction of the 4-momentum vector, $\tilde{p} \equiv m\tilde{u}$, 4-forces, $\tilde{K} \equiv d\tilde{p}/d\tau$, etc., which are all important mechanical quantities. Their components all transform in the same way as do coordinate differences when one changes basis from one (inertial) observer to another.

e.) The metric allows a mapping between tangent vectors and 1-forms, $p_\alpha \equiv \eta_{\alpha\beta} p^\beta$, and

II. A Summary of Symbols Related to Manifolds and their Vector/Tensor Spaces

1. Manifolds, \mathcal{M} , of dimension n :
 - a. We use charts (sets of coordinate systems), which generalize the intuitive notion of coordinates, and the symbol x^μ is common; one should think of a chart as an **invertible**, arbitrarily-often-differentiable mapping from (some portion of) the manifold into a portion of ordinary \mathbb{R}^n , for some minimal value of n . Referring to a particular chart as x , we could use the notation $x : \mathcal{M} \rightarrow \mathbb{R}^n$. [At this point, relative to conventions, we should note that Finley uses the (common) useful mathematical notation that \mathbb{R} stands for the set of all real numbers and \mathbb{R}^n for the space of “n-tuples” of real numbers. As well he will also use \mathbb{C} for the set of all complex numbers, and \mathbb{Z} for the set of all integers, as well as \mathbb{Z}^+ for the set of all non-negative integers.]

A manifold is required to have (at least) a set of coordinate mappings large enough so that every point on the manifold lies within the domain of some chart, **and** the following requirement on these charts. Consider some point on the manifold that lies within the domain of definition of two charts; then there is a mapping from $\mathbb{R}^n \rightarrow \mathbb{R}^n$ generated by taking the inverse of one of these charts, from \mathbb{R}^n to \mathcal{M} , followed by the second chart, from \mathcal{M} to \mathbb{R}^n . We require that, for every such point, this mapping should be a mapping

that possesses arbitrarily many continuous derivatives. We refer to such a mapping as being either “of class $C^{(\infty)}$,” or simply as being “smooth.”

- b. We are interested in many different sorts of (*smooth*) functions, mapping manifolds into the real numbers, \mathbb{R} , and denote the set of all of them by the symbol \mathcal{F} , or $\mathcal{F}(\mathcal{M})$ if it is necessary to specify which manifold; and also in
 - c. curves on manifolds, mapping \mathbb{R} into the manifold, i.e., $\Gamma : \mathbb{R} \rightarrow \mathcal{M}$, as also with their tangent vectors at each point along them, $d\Gamma/dt$, where $t \in \mathbb{R}$ is the parameter along them; as well as simply the more general notion of maps of manifolds into one another.
2. We will need generalizations of the intuitive, flat-space notion of a “vector,” at each and every point of the manifold. Since we need these quantities to be completely specifiable at each and every point of the manifold, they must be defined locally in all generality. Therefore we must specify at each and every point of a manifold one or more “vector spaces,” so that the resulting manifold now looks “something like a porcupine,” with all the vector spaces sticking out.
- a. We remember that a vector space is required to be a set of “things,” which have various properties. In particular, if a, b , and c are elements of this vector space, i.e., if they are vectors, then we must have a “sum” for them that retains the property of being a vector, and a “multiplication by scalars” which also retains this property; in addition, these two different mappings must behave properly, one with respect to the other. More precisely, let our vector space be V , with $a, b, c \in V$, and the space of “scalars” be called K , with $\alpha, \beta, \gamma \in K$. By “scalars” we mean quantities which have “all” the usual properties of real, or complex, numbers, i.e., they can be added, can be multiplied together, have an additive identity (zero) and a multiplicative identity (one), additive and multiplicative inverses, etc. [Sets with these properties are usually called *fields*, but we will (at least usually) only be using \mathbb{R} and \mathbb{C} for our sets of scalars.] Then let the desired mappings for the sum, s ,

and for the scalar multiplication, m , have the following requirements:

$$s : V \times V \rightarrow V , \quad m : K \times V \rightarrow V ,$$

$$s(a, b) = s(b, a) \in V , \quad \text{“the sum is commutative”}$$

$$s(a, s(b, c)) = s(s(a, b), c) , \quad \text{“the sum is associative”}$$

V contains special elements:

$$e \in V \text{ such that } s(e, a) = a = s(a, e) , \text{ for all } a \in V , \quad \text{“an identity,”}$$

$$a \in V \implies \check{a} \in V \text{ such that } s(a, \check{a}) = e = s(\check{a}, a) , \quad \text{“an inverse,”}$$

$$m(\alpha, s(a, b)) = s(m(\alpha, a), m(\alpha, b)) \quad \text{and} \quad m(\alpha + \beta, a) = s(m(\alpha, a), m(\beta, a)) ,$$

scalar multiplication is “distributive”

$$\text{and } m(\alpha\beta, a) = m(\alpha, m(\beta, a)) , \quad \text{scalar multiplication is “associative” ,}$$

$$m(0, a) = e , \quad m(1, a) = a .$$

It is usual and normal for the sum to be denoted simply by $+$, and for scalar multiplication to be denoted simply by writing the symbols for the two next to each other, usually with the scalar on the left.

Particular examples of vector spaces are provided by

- a. tangent vectors to a curve, which act as linear operators on functions by determining directional derivatives of those functions; I use the symbol \mathcal{T} for the bundle of tangent-vector spaces over some neighborhood of points of a manifold, such as spacetime,
- b. and, separately, by the space of linear, continuous maps of tangent vectors to \mathbb{R} , which we refer to as co-tangent vectors or 1-forms, or differential forms; use the symbol Λ for the bundle of spaces of co-vectors or 1-forms over a neighborhood of our spacetime.
- c. Tensor spaces at any given point $p \in \mathcal{M}$, are also vector spaces; in general they are linear, continuous maps of some number, s , of tangent vectors and some number, r , of 1-forms into \mathbb{R} . One may also say that they are contravariant of type r and covariant of type s , or simply that they are “of type $[r, s]$.”

- d. Physically interesting examples of tensors are given by
 - i.) an ordinary tangent vector, which is of type $[1,0]$,
 - ii.) a differential form, or 1-form, which is of type $[0,1]$,
 - iii.) the metric tensor, of type $[0,2]$, and symmetric
 - iv.) the electromagnetic field tensor, also of type $[0,2]$ but skew-symmetric,
 - v.) the stress-energy tensor, of type $[1,1]$, and
 - vi.) the curvature of the manifold, of type $[1,3]$.
 - e. Of special interest are the tensor spaces made up of combinations of skew-symmetric tensor products of a number, p , of 1-forms, these objects being called *p-forms*.
 - i.) They involve the use of the “wedge” product and the exterior derivative, and have Poincaré’s Lemma as an important corollary. They are also important for integration over manifolds.
 - ii.) Area and volume forms, and the totally anti-symmetric tensor density, related to Levi-Civita’s symbol, $\epsilon^{\mu\nu\lambda\eta}$. (Hodge) duality is a very useful map from $\Lambda^p \rightarrow \Lambda^{n-p}$, which is created by the Levi-Civita symbol and the metric.
 - f. Choices of sets of basis vectors for either of these two vector spaces must be made! Each coordinate chart determines the so-called coordinate basis for our initial vector spaces: the set of tangent vectors $\{\partial_{x^a}\}$, which are tangent to the various curves along which only one coordinate varies at a time, form *the coordinate basis for \mathcal{T}^1* , while the differentials of those coordinates, $\{dx^a\}$, form the coordinate basis for Λ^1 . *Anholonomic* basis sets are choices of linear combinations of the elements of the coordinate basis which are **not** simply partial derivatives with respect to some other set of coordinates. The value of such basis sets is that they allow components of vectors to correspond to measurements made by actual physical observers; an example is the anholonomic basis set $\{dr, r d\theta\}$, arranged so that all elements have the same physical dimension, i.e., length.
3. Symbols to denote particular kinds of geometrical objects:
- a.) We begin by having two physically-different kinds of vector spaces to consider, namely the space of “**tangent vectors**,” \mathcal{T}^1 , and the space of “**differential forms**”, Λ^1 ,

which is *dual* to the first sort. We will generalize the more familiar use of “arrows” over symbols, to denote 3-dimensional vectors, by using an “over-tilde” to indicate a tangent vector, and an “under-tilde” to indicate a differential form:

- i. \tilde{x} is known to be an element of the tangent bundle; and
 - ii. $\underline{\omega}$ is known to be a differential form.
- b.) For **p-forms** of rank higher than 1, we will also use an “under-tilde” so that such a symbol does not automatically tell us “the value of p.” It is very unlikely that this will cause much confusion, however, since we will only truly discuss relatively few so that each one of interest will generally have its own particular symbol, never used for anything else.
- c.) For higher-rank tensors that are not p-forms, we will their special symbol in boldface letters, such as $\boldsymbol{\eta}$ for the metric.
2. Vectors in any vector space are usually measured by determining their (scalar) components with respect to some choice of basis.
- a.) A basis for that vector space is a maximal, linearly-independent set of vectors within that vector space. Our most common spaces of interest will be 4-dimensional; therefore, most of my examples will come from there. Let us use the standard symbols, $\{\tilde{\mathbf{e}}_1, \tilde{\mathbf{e}}_2, \tilde{\mathbf{e}}_3, \tilde{\mathbf{e}}_4\}$, to denote the elements of a basis for the space of tangent vectors—more simply, we can simply write $\{\tilde{\mathbf{e}}_\mu\}_{\mu=1}^4$ or just $\{\tilde{\mathbf{e}}_\mu\}_1^4$ to mean the same thing. Any vector may be written as a linear combination of the basis vectors, i.e., there always exists **unique**, *scalar* quantities, x^μ , such that

$$\tilde{x} = x^\mu \tilde{\mathbf{e}}_\mu . \tag{0.1}$$

The index on the symbol x^μ is a **superscript**; this will always be true when the vector \tilde{x} is a “tangent vector.” We will often refer to these indices as “*contravariant*.”

- b.) We also need a choice for basis vectors for differential forms, i.e., for the co-tangent bundle. Habitually, we will use $\{\underline{\omega}^\alpha\}_1^4$ as a set of symbols for this basis; we will also

always relate the two basis sets, for tangent vectors and for differential forms, so that they are *reciprocal bases*. Recalling that differential forms map tangent vectors into scalars, the definition of reciprocal bases is

$$\varpi^\alpha(\tilde{e}_\beta) = \delta_\beta^\alpha, \quad \text{—the Kronecker delta: } \delta_\beta^\alpha = \begin{cases} 1, & \alpha = \beta, \\ 0, & \text{otherwise.} \end{cases}$$

- c.) For an arbitrary 1-form, $\mu \in \Lambda^1$, its components are then the set of scalars $\{\mu_\alpha\}_1^4$ such that

$$\mu = \mu_\alpha \varpi^\alpha. \quad (0.2)$$

The indices on the components of a 1-form are always lower indices, i.e., subscripts; they are referred to as “*covariant*.” We also have the following set of useful relations, rather analogous to the behavior of “dot products”:

$$x^\alpha = \varpi^\alpha(\tilde{x}) \quad , \quad \rho_\beta = \varrho(\tilde{e}_\beta) \quad . \quad (0.3)$$

3. Matrix representations of the components of geometrical objects.

- a.) Matrices are not, *a priori* geometrical objects but, rather, arrays of scalar quantities along with (standard) rules concerning their display, and their manipulation to create new matrices. One must therefore have (rather arbitrary) conventions/rules that relate the matrix arrays of scalar quantities with the arrays of scalar quantities that form the set of components of some geometrical object.
- b.) It is convenient, and very conventional, to display the set of components of a vector by means of a **matrix with only one column**, usually referred to as a **column-vector**. However, since we have two sorts of vectors, we generalize this convention—not done by all authors—so that we represent our geometrical vectors so that
- i. contravariant components are represented via column-vectors, and
 - ii. covariant components are represented via row-vectors,

where **row-vectors** are actually **matrices with only one row**.

- c.) Since a matrix representation does not display the actual basis vectors themselves, but does assume an ordering of those basis vectors, these choices affect the matrix display in question. A column-vector describing, say, the (contravariant) components of the 4-vector displacement of a particle has the following appearance, while those for the corresponding row-vector describing the covariant component of the same 4-vector displacement is instead this way:

$$\tilde{x} \implies x^\mu = \begin{pmatrix} \Delta x \\ \Delta y \\ \Delta z \\ \Delta t \end{pmatrix}, \quad \mu = 1, 2, 3, 4, \quad (0.4)$$

$$\underline{x} \equiv H(\tilde{x}, \cdot) \implies x_\alpha = (\Delta x, \Delta y, \Delta z, -\Delta t)$$

- d.) Since different choices of basis would generate different sets of scalar quantities for the very same vector, it is of course true that \tilde{x} is not **equal** to the set of scalars, x^μ ; therefore, instead of writing an equality, I use the symbol \implies , which is read as “is represented by.” When this symbol is used, it should remind us that we must know which particular choice of basis has been made, and what the choice of ordering is, before we may understand whatever comes next.
- e.) We also use matrices to display the components of other geometrical quantities, especially those sorts of tensors that relate two vectors at once—called second-rank tensors. Here we give appropriate **conventions** for two, very common examples, the metric tensor and the electromagnetic field tensor.
- f.) The metric tensor is a bilinear mapping that takes two tangent vectors and gives a scalar—their “scalar product.” This makes it an element of $\Lambda^1 \otimes \Lambda^1$, a basis for which is the set of all $\varpi^\alpha \otimes \varpi^\beta$, where \otimes is a symbol that denotes the tensor product of two vectors. Therefore we can define the components in either of two equivalent ways:

$$\boldsymbol{\eta} = \eta_{\alpha\beta} \varpi^\alpha \otimes \varpi^\beta, \quad \text{or} \quad \eta_{\alpha\beta} = \boldsymbol{\eta}(\tilde{e}_\alpha, \tilde{e}_\beta), \quad (0.5)$$

and then we can use a matrix representation to display these components:

$$\boldsymbol{\eta} \implies \eta_{\mu\nu} = \begin{pmatrix} +1 & 0 & 0 & 0 \\ 0 & +1 & 0 & 0 \\ 0 & 0 & +1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \quad \mu, \nu = 1, 2, 3, 4, \quad (0.6)$$

Notice that $\boldsymbol{\eta}$ is that particular metric tensor that is commonly used in special relativity, for an “orthonormal” metric. In a more general context, I would use the symbol $\mathbf{g} = g_{\alpha\beta} \varpi^\alpha \otimes \varpi^\beta$ to refer to the metric tensor, that determines the interval.

- g.) On the other hand, since the electromagnetic tensor is actually a 2-form, i.e., it is a skew-symmetric, second-rank tensor—an element of Λ^2 —we use skew-symmetric products of the original basis elements to provide a basis for Λ^2 :

$$\underline{F} \equiv \frac{1}{2} F_{\alpha\beta} \varpi^\alpha \wedge \varpi^\beta \equiv F_{\alpha\beta} \left\{ \frac{1}{2} (\varpi^\alpha \otimes \varpi^\beta - \varpi^\beta \otimes \varpi^\alpha) \right\} \quad , \quad (0.7)$$

where the $\frac{1}{2}$ is justified by the definition of the *wedge product*, which ensures that the tensor in question is skew-symmetric. In 4 dimensions, any skew-symmetric, second-rank tensor is specified by exactly 6 degrees of freedom; therefore, it is customary to label those 6 degrees of freedom via a pair of 3-dimensional vectors. In the case of the electromagnetic tensor, we refer to those two 3-vectors as “the electric field,” \vec{E} , and “the magnetic field,” \vec{B} , justifying the following representations, with respect to the same bases as we have been using above:

$$\underline{F} \implies F_{\mu\nu} = \begin{pmatrix} 0 & B^z & -B^y & E_x \\ -B^z & 0 & B^x & E_y \\ B^y & -B^x & 0 & E_z \\ -E_x & -E_y & -E_z & 0 \end{pmatrix} \quad , \quad \mu, \nu = 1, 2, 3, 4 \quad , \quad (0.8)$$

- h.) When there is more than one index involved, we **also** need **conventions** concerning the ordering of the indices, relative to the standard convention for matrices, that the entries within matrices are labelled by their **row** and their **column**. For 2-index arrays of components, such as the ones just above, I always follow the convention:

- i. the first index corresponds to the row index of the matrix, and
- ii. the second index corresponds to the column index of the matrix.

We will use this convention independent of whether the indices of the components are upper or lower, i.e., contravariant or covariant! Therefore, for example, we may

consider the following instances, where we create a new geometrical quantity by the use of the metric—as discussed in detail in class.

$$\begin{aligned}
F^\mu{}_\nu \equiv \eta^{\mu\lambda} F_{\lambda\nu} &\implies \begin{pmatrix} +1 & 0 & 0 & 0 \\ 0 & +1 & 0 & 0 \\ 0 & 0 & +1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} 0 & B^z & -B^y & E_x \\ -B^z & 0 & B^x & E_y \\ B^y & -B^x & 0 & E_z \\ -E_x & -E_y & -E_z & 0 \end{pmatrix} \\
&= \begin{pmatrix} 0 & B^z & -B^y & E_x \\ -B^z & 0 & B^x & E_y \\ B^y & -B^x & 0 & E_z \\ E_x & E_y & E_z & 0 \end{pmatrix}, \tag{0.9}
\end{aligned}$$

$$\begin{aligned}
F_\mu{}^\nu \equiv F_{\mu\lambda} \eta^{\lambda\nu} &\implies \begin{pmatrix} 0 & B^z & -B^y & E_x \\ -B^z & 0 & B^x & E_y \\ B^y & -B^x & 0 & E_z \\ -E_x & -E_y & -E_z & 0 \end{pmatrix} \begin{pmatrix} +1 & 0 & 0 & 0 \\ 0 & +1 & 0 & 0 \\ 0 & 0 & +1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \\
&= \begin{pmatrix} 0 & B^z & -B^y & -E_x \\ -B^z & 0 & B^x & -E_y \\ B^y & -B^x & 0 & -E_z \\ -E_x & -E_y & -E_z & 0 \end{pmatrix}. \tag{0.10}
\end{aligned}$$

4. Display of matrix arithmetic without explicit indices

- a.) We sometimes want to conserve ink by not explicitly writing out the indices on symbols representing matrices. In fact, this justifies the fact that there are very standard conventions concerning matrix multiplication. Having just looked at some examples of matrix multiplication in the preceding section, I expound these conventions some more, using as an example the changes wrought in various distinct sets of components as a result of a change in the choice of basis elements.
- b.) We begin by considering two distinct choices for basis elements of the cotangent space (of differential forms), namely $\{\varpi^\alpha\}$ and $\{\varrho^\mu\}$. Since either set forms a basis, we may immediately write down each member of the one choice of basis in terms of the members of the other; i.e., there exists uniquely a (square array) of scalar quantities $A^\alpha{}_\mu$ and, alternatively, the quantities $B^\mu{}_\alpha$ such that

$$\varpi^\beta = A^\beta{}_\nu \varrho^\nu, \quad \varrho^\mu = B^\mu{}_\alpha \varpi^\alpha. \quad (0.11)$$

Using the elements of these arrays, we may phrase the fact that if one goes “backward,” she should surely arrive at the place she started; i.e., we need

$$A^\beta{}_\nu B^\nu{}_\alpha = \delta^\beta{}_\alpha, \quad B^\mu{}_\alpha A^\alpha{}_\nu = \delta^\mu{}_\nu. \quad (0.12)$$

On the other hand, it is certainly possible, and probably advisable to consider these square arrays as square matrices; in that case, we have two options. The first is simply to treat the previous equation as indicating the details of matrix multiplication. A second, and much superior option, however, is to give explicit **names** to the matrices, whose elements are labelled by the symbols $A^\alpha{}_\mu$, etc. Therefore, we now suppose given to us two matrices, which we name simply A and B , with those elements. It is even useful to have a notation for this operation; therefore, we use the following notational convention:

$$A = ((A^\alpha{}_\nu)) \quad \text{means } A \text{ is the matrix with components } A^\alpha{}_\nu \quad (0.13)$$

The Eqs. (0.12) are simply statements that say that the two matrices A and B are inverse to one another, which, in pure matrix notation is written

$$AB = I = BA, \quad \text{or } B = A^{-1}, \quad (0.14)$$

where, of course the Kronecker delta symbol, δ^μ_α , used earlier simply denoted the components of the identity matrix.

- c.) Knowing that the geometrical objects themselves are independent of any choice of basis, we can induce transformation of various other associated objects. We begin with the components of an arbitrary 1-form ϱ :

$$\begin{aligned} \varrho &= \rho_\alpha \varpi^\alpha & \text{or } \varrho &\implies P \equiv ((\rho_\alpha)), \text{ relative to the basis } \{\varpi^\alpha\} \\ \varrho &= \rho'_\mu \varpi^\mu & \text{or } \varrho &\implies P' \equiv ((\rho'_\mu)), \text{ relative to the basis } \{\varpi^\mu\} \\ \varpi^\mu &= (A^{-1})^\mu_\alpha \varpi^\alpha & \implies \rho'_\mu &= A^\alpha_\mu \rho_\alpha \text{ or } P' = PA, \end{aligned} \quad (0.15)$$

where we have agreed that the matrix representation of the components of a 1-form, such as ϱ , would be taken as a row-vector, i.e., a 1×4 matrix, so that the order indicated above for the matrix multiplication is indeed the correct order. (The symbol, P , being used above, represents a **capital** Greek ρ .) We may then consider the reciprocal basis for tangent vectors, relative to the two bases $\{\varpi^\alpha\}$ and $\{\varpi^\mu\}$, respectively, as follows:

$$\begin{aligned} \varpi^\alpha(\tilde{e}_\beta) &= \delta^\alpha_\beta & \text{and } \varpi^\mu(\tilde{f}_\nu) &= \delta^\mu_\nu, \\ \varpi^\mu &= (A^{-1})^\mu_\alpha \varpi^\alpha & \implies \tilde{f}_\mu &= A^\alpha_\mu \tilde{e}_\alpha. \end{aligned} \quad (0.16)$$

We are then able to look at the components of an arbitrary tangent vector, \tilde{v} :

$$\begin{aligned} \tilde{v} &= v^\alpha \tilde{e}_\alpha & \text{or } \tilde{v} &\implies V \equiv ((v^\alpha)), \text{ relative to the basis } \{\tilde{e}_\alpha\} \\ \tilde{v} &= v'^\mu \tilde{f}_\mu & \text{or } \tilde{v} &\implies V' \equiv ((v'^\mu)), \text{ relative to the basis } \{\tilde{f}_\mu\} \\ \varpi^\mu &= (A^{-1})^\mu_\alpha \varpi^\alpha & \implies v'^\mu &= (A^{-1})^\mu_\alpha v^\alpha \text{ or } V' = A^{-1} V, \end{aligned} \quad (0.17)$$

where we have agreed that the matrix representation of the components of a tangent vector, such as \tilde{v} , would be taken as a column-vector, i.e., a 4×1 matrix, so that the order indicated above for the matrix multiplication is again the correct order.

- d.) Notice that the components of a 1-form, conventionally taken as lower indices, are such that they transform according to the same matrix, $A = B^{-1}$, as do the tangent vector basis elements, $\{\tilde{e}_\alpha\}$; it is for this reason that they are referred to as “covariant,” in the sense that they transform in the same manner. On the other hand, we see that the components of a tangent vector transform in the inverse manner to the basis elements for tangent vectors, for which reason they are referred to as “contravariant.”
- e.) Following the reasoning above, one may now generalize and discover the transformation laws of more complicated objects, i.e., higher-rank tensors. This time, again, I give two examples. Firstly, we consider the metric tensor, $\mathbf{g} \in \Lambda^1 \otimes \Lambda^1$:

$$\begin{aligned}
\mathbf{g} &= g_{\alpha\beta} \varpi^\alpha \otimes \varpi^\beta & \text{or } \mathbf{g} &\implies G \equiv ((g_{\alpha\beta})) , \text{ relative to the basis } \{\varpi^\alpha\} \\
\mathbf{g} &= g_{\mu\nu} \varrho^\mu \otimes \varrho^\nu & \text{or } \mathbf{g} &\implies G' \equiv ((g'_{\mu\nu})) , \text{ relative to the basis } \{\varrho^\mu\} \\
\varrho^\mu &= (A^{-1})^\mu{}_\alpha \varpi^\alpha & \implies & g'_{\mu\nu} = A^\alpha{}_\mu A^\beta{}_\nu g_{\alpha\beta} \text{ or } G' = A^T G A ,
\end{aligned} \tag{0.18}$$

The symbol, T , in the last line of the equation, indicates the transpose of the matrix in question, and is necessary because of the ordering of the indices. Matrix multiplication also sums the column indices of the matrix on the left with the row indices of the matrix on the right. Of the two summations indicated in Eqs. (0.18)—via the Einstein summation convention—with the indices in the equation, the second sum has this ordering, and so is indicated simply by the matrix product GA ; however, the first has the opposite ordering, necessitating the use of the transpose on the matrix A , in order that the matrix multiplication rules represent correctly the desired summation, namely $A^T G$.

- f.) As a second example, consider the type [1,1] tensor created in Eqs. (0.9), with components $F^\alpha{}_\beta$, which gives us

$$F'^\mu{}_\nu = (A^{-1})^\mu{}_\alpha A^\beta{}_\nu F^\alpha{}_\beta \quad \text{or } F' = A^{-1} F A \quad . \tag{0.19}$$

- g.) The study of matrix transformations has a long history. Transformations of the type appropriate for F , as given in Eqs. (0.19), are called *similarity transformations*,

$F' = A^{-1} F A$; they preserve all the eigenvalues of the matrix, and, therefore, also its determinant. For *normal matrices*, they can be used to bring the matrix into diagonal form. (Normal matrices are those which commute with their transpose; both symmetric and skew-symmetric matrices are examples.)

- h.) On the other hand, transformations of the type appropriate for the metric matrix G , as given in Eqs. (0.18) above, are called *congruency transformations*: $G' = A^T G A$. They do not preserve the determinant; instead they preserve the *signature of the matrix*, which is a particular set of +1's, -1's, and 0's for that matrix. Sylvester's theorem says every symmetric matrix has a congruency transformation which will not only diagonalize it, but in fact bring it to have only +1, -1, or 0 in all the places on the diagonal. This set, independent of order, is called the signature. Therefore, in a spacetime, the equivalence principle assures us that there is always a change of basis for \mathcal{T} , effected by an invertible matrix, M , such that the components of the metric can be converted from $g_{\alpha\beta}$ to just those of an arbitrary inertial frame of special relativity, $\eta_{\mu\nu}$:

$$G = M^T H M, \quad \text{or} \quad g_{\alpha\beta} = M^\mu{}_\alpha \eta_{\mu\nu} M^\nu{}_\beta, \quad (0.20)$$

where the symbol H is a **capital** Greek η .

- i.) Comments on Determinants of Matrices.

In addition to rules for matrix multiplication, matrices, for example $A \equiv ((A^a{}_b))$, also come equipped with a definition of various scalars created from them. The only important ones we are likely to use are *the trace*, $\text{tr } A$, and *the determinant*, $\det A$.

- i. The definition of trace must be given so that it will be invariant under tensor transformation rules, i.e., it should be a scalar; therefore, we can write

$$\text{tr } A \equiv A^\alpha{}_\alpha = g^{\alpha\beta} A_{\beta\alpha}, \quad (0.21)$$

where $g^{\alpha\beta}$ are the components of the inverse metric tensor;

- ii. the definition of the determinant, of an $n \times n$ matrix, involves taking products of n elements, one from each row and from each column, in certain orders, with certain signs.

The Levi-Civita symbol, $\epsilon^{b_1 b_2 \dots b_n} \equiv \epsilon_{b_1 b_2 \dots b_n}$, has n indices, is skew-symmetric under the interchange of those indices, and is such that $\epsilon^{1234} = +1$.

It has been defined precisely so that it creates determinants via summations, as expressed by the **Fundamental Theorem of Determinants**:

$$\epsilon^{b_1 b_2 \dots b_n} A^{a_1}_{b_1} A^{a_2}_{b_2} \dots A^{a_n}_{b_n} = \epsilon^{a_1 a_2 \dots a_n} \det(A). \quad (0.22)$$

In 4 dimensions we have

$$\begin{aligned} \epsilon^{1234} = +1 = \epsilon^{2143} = \epsilon^{4321} = \epsilon^{1423} = \dots \\ \epsilon^{2134} = -1 = \epsilon^{1243} = \epsilon^{4312} = \epsilon^{4123} = \dots \end{aligned} \quad (0.23)$$

$$\epsilon^{\alpha\beta\gamma\delta} = 0 \text{ whenever 2 indices are equal.}$$

One should notice that $\epsilon^{\alpha\beta\gamma\delta}$ is **not a tensor**.

5. The **volume 4-form** and its relationship to Hodge duality.

a.) Over 4-dimensional manifolds, there are 5 distinct spaces of p -forms, Λ^p :

- i. Λ^0 is just the space of continuous (C^∞) functions, also denoted by \mathcal{F} . We say that it has dimension 1, since no true “directions” are involved.
- ii. Λ^1 is the space of 1-forms, already considered; it has as many dimensions as the manifold, so for 4-dimensional spacetime, it has dimension 4.
- iii. Λ^2 is the space of 2-forms, i.e., skew-symmetric tensors, or linear combinations of wedge products of 1-forms; therefore in general it has dimension $\frac{1}{2}n(n-1)$, which becomes 6 for 4-dimensional spacetime. A basis can be created by taking all wedge products of the basis set for 1-forms: $\{\varpi^\alpha \wedge \varpi^\beta \mid \alpha, \beta = 1, \dots, 4; \alpha < \beta\}$.

- iv. Λ^3 is the space of 3-forms, i.e., linear combinations of wedge products of 1-forms, three at a time; in general it has dimension $\binom{n}{3} = \frac{1}{6} n(n-1)(n-2)$, which becomes 4 for 4-dimensional spacetime.
 - v. Λ^4 is the space of 4-forms; in general it has dimension $\binom{n}{4}$. For 4-dimensional spacetime, this is a 1-dimensional space; i.e., every 4-form is proportional to every other; we refer to some particular choice of basis for this space as *the volume form*. (In general n dimensions, the volume form is always an n -form.)
 - vi. Over n -dimensional spacetime, it is impossible to have more than n things skew all at once; therefore, the volume form is always the last in the sequence of basis sets for p -forms. So, in 4 dimensions, there is no Λ^p for $p \geq 5$.
- b.) Working in the usual (local) Minkowski coordinates, where it is reasonable to choose $\{dx, dy, dz, dt\}$ as a basis for 1-forms, we choose the particular 4-form

$$\mathcal{V} \equiv dx \wedge dy \wedge dz \wedge dt \quad \text{the (standard) volume form} \quad (0.24)$$

as our choice of a volume form.

If $\{\omega^\alpha\}_1^4$ is an arbitrary basis for 1-forms, we may define the very important quantity $\eta^{\alpha\beta\gamma\delta}$, which gives the “components” of the volume form relative to this choice of basis:

$$\omega^\alpha \wedge \omega^\beta \wedge \omega^\gamma \wedge \omega^\delta \equiv \eta^{\alpha\beta\gamma\delta} \mathcal{V} \quad , \quad (0.25)$$

$$\mathcal{V} = \frac{1}{4!} \eta_{\alpha\beta\gamma\delta} \omega^\alpha \wedge \omega^\beta \wedge \omega^\gamma \wedge \omega^\delta \equiv \frac{1}{4!} \{g_{\alpha\rho} g_{\beta\sigma} g_{\gamma\tau} g_{\delta\varphi} \eta^{\rho\sigma\tau\varphi}\} \omega^\alpha \wedge \omega^\beta \wedge \omega^\gamma \wedge \omega^\delta \quad . \quad (0.26)$$

This tensor is completely skew-symmetric, i.e., it changes sign when any two indices are interchanged, and so must be proportional to the *Levi-Civita symbol*, $\epsilon^{\alpha\beta\gamma\delta}$, used for determinants. One verifies that the following defines tensors of type [4,0] and [0,4], respectively, related as usual by raising/lowering of indices via the metric tensor:

$$\eta^{\alpha\beta\gamma\delta} = \frac{1}{m} \epsilon^{\alpha\beta\gamma\delta} \quad , \quad \eta_{\alpha\beta\gamma\delta} = (-1)^s m \epsilon_{\alpha\beta\gamma\delta} \quad ,$$

where $m \equiv \det(M)$, $G = M^T H M$, (0.27)

$$g \equiv \det G = m^2 \det(H) = (-1)^s m^2 \quad , \quad s = 0 \text{ or } 1 .$$

Of course the values of $\eta^{\alpha\beta\gamma\delta}$ depend on the basis chosen; however, for an orthonormal tetrad, where the metric components are just $\eta_{\mu\nu}$, we have that $\eta^{1234} = +1 = -\eta_{1234}$. On the other hand, in ordinary, Cartesian 3-dimensional space, in an orthonormal frame, we simply have $\eta^{123} = +1 = \eta_{123}$, as expected.

The following rather complicated relations are worth writing down, since they are occasionally of considerable use:

$$\eta^{a_1 \dots a_n} \eta_{b_1 \dots b_n} = (-1)^s \begin{vmatrix} \delta_{b_1}^{a_1} & \dots & \delta_{b_n}^{a_1} \\ \vdots & \ddots & \vdots \\ \delta_{b_1}^{a_n} & \dots & \delta_{b_n}^{a_n} \end{vmatrix} \equiv (-1)^s \delta_{b_1 \dots b_n}^{a_1 \dots a_n} \quad , \quad (0.28)$$

where the vertical bars imply a determinant, and we regularly intend to denote a $q \times q$ determinant made with Kronecker delta entries, as above, with the *generalized Kronecker delta symbol* indicated above, that has q upper and q lower indices.

Writing $p + p' \equiv n$, it then follows that

$$\eta^{a_1 \dots a_p c_1 \dots c_{p'}} \eta_{b_1 \dots b_p c_1 \dots c_{p'}} = (-1)^s \delta_{b_1 \dots b_p c_1 \dots c_{p'}}^{a_1 \dots a_p c_1 \dots c_{p'}} = (p')! (-1)^s \delta_{b_1 \dots b_p}^{a_1 \dots a_p} \quad . \quad (0.29)$$

c.) **the (Hodge) dual**, $*$: $\Lambda^p \longrightarrow \Lambda^{n-p}$

Let \mathcal{Q} be an arbitrary p -form; then we denote the (Hodge) dual by $^*\mathcal{Q}$, an $(n-p)$ -form.

They are related as follows:

$$\begin{aligned} \mathcal{Q} &= \frac{1}{p!} \alpha_{\mu_1 \dots \mu_p} \varpi^{\mu_1} \wedge \dots \wedge \varpi^{\mu_p} \quad , \\ ^*\mathcal{Q} &\equiv \frac{i^{pp'+s}}{p!(p')!} \alpha^{b_1 \dots b_p} \eta_{b_1 \dots b_p c_1 \dots c_{p'}} \varpi^{c_1} \wedge \dots \wedge \varpi^{c_{p'}} \equiv \frac{1}{(p')!} (^*\alpha)_{c_1 \dots c_{p'}} \varpi^{c_1} \wedge \dots \wedge \varpi^{c_{p'}} \quad . \end{aligned} \quad (0.30)$$

The factors of $i \equiv \sqrt{-1}$ have been inserted in **just such a way that** the dual of the dual brings one back to where she started:

$$^*\{^*\mathcal{Q}\} = \mathcal{Q} \quad . \quad (0.31)$$

There are various conventions concerning the i 's in the definition. My convention, using the factors of i , allows for eigen-2-forms of the $*$ operator, since Eq. (0.31) obviously tells us that the eigenvalues of $*$ are ± 1 .

- d.) Since the definition of (Hodge) duality appears quite complicated, it is worthwhile writing it down for **all** plausible exemplars that may occur, in our 4-dimensional spacetime. We do this for the standard Minkowski tetrad, $\{dx, dy, dz, dt\}$, and the bases of each Λ^p :

$$\begin{aligned} \Lambda^1 \leftrightarrow \Lambda^3 : * \begin{pmatrix} dx \\ dy \\ dz \\ dt \end{pmatrix} &= - \begin{pmatrix} dy \wedge dz \wedge dt \\ dz \wedge dx \wedge dt \\ dx \wedge dy \wedge dt \\ dx \wedge dy \wedge dz \end{pmatrix}, & \Lambda^0 \leftrightarrow \Lambda^4 : *1 &= -i dx \wedge dy \wedge dz \wedge dt, \\ \Lambda^2 \leftrightarrow \Lambda^2 : * \begin{pmatrix} dx \wedge dy \\ dy \wedge dz \\ dz \wedge dx \end{pmatrix} &= -i \begin{pmatrix} dz \wedge dt \\ dx \wedge dt \\ dy \wedge dt \end{pmatrix}, \quad . \end{aligned} \quad (0.32)$$

As an example, consider the electromagnetic 2-form, in Eqs. (0.8), from which we have:

$$\underline{F} \implies F_{\mu\nu} = \begin{pmatrix} 0 & B^z & -B^y & E_x \\ -B^z & 0 & B^x & E_y \\ B^y & -B^x & 0 & E_z \\ -E_x & -E_y & -E_z & 0 \end{pmatrix}, \quad (0.32)$$

$$* \underline{F} \implies (*F)_{\mu\nu} = -i \begin{pmatrix} 0 & -E^z & E^y & B_x \\ E^z & 0 & -E^x & B_y \\ -E^y & E^x & 0 & B_z \\ -B_x & 0 & -B_y & -B_z \end{pmatrix}.$$

Note that the map from \underline{F} to $i^* \underline{F}$ is accomplished by sending $\vec{B} \rightarrow -\vec{E}$ and $\vec{E} \rightarrow +\vec{B}$.

- e. To complete the picture, we also give details for **3-dimensional, Euclidean space**, with Cartesian basis, $\{dx, dy, dz\}$:

$$\Lambda^1 \leftrightarrow \Lambda^2 : * \begin{pmatrix} dx \\ dy \\ dz \end{pmatrix} = - \begin{pmatrix} dy \wedge dz \\ dz \wedge dx \\ dx \wedge dy \end{pmatrix}, \quad \Lambda^0 \leftrightarrow \Lambda^3 : *1 = dx \wedge dy \wedge dz \quad .$$