Appendix F

Indexology

F.1 General Linear Transformations

F.2 Vectors and Tensors under Rotation

The first definition of a vector that you usually hear is

> A three dimensional object that transforms under rotations in the same way as the position triplet is a vector.

In a three dimensional space, the position is designated by the triplet (x, y, z). This triplet is the position vector, $\vec{X}$, as projected onto a complete orthonormal basis,

$$\vec{X} = x\hat{x} + y\hat{y} + z\hat{z} = \sum_{l=x,y,z} l\hat{l} \tag{F.1}$$

where the three basis vectors satisfy

$$\hat{i} \cdot \hat{j} = \delta_{ij}, \tag{F.2}$$

where

$$\delta_{ij} = \begin{cases} 1 & : \quad i = j \\ 0 & : \quad i \neq j \end{cases} . \tag{F.3}$$

The triplet of numbers (x, y, z) are the coordinates of the position vector.

There are two ways of viewing the directive "under rotation." We can actually rotate the position vector to a new location or we can identify the same position but in a coordinate system that is like the original one but rotated. These two views are in a sense operationally equivalent: the rotated positions components are the same as the components of the same position as viewed by the rotated observer if the sense of the rotation is reversed. These two views are designated the active and passive views of the transfomation respectively. For ease of implementation it is simpler to use the passive view. Occasionally, it is easier to interpret operations in the active view. It should be clear from the context when the active view is being used.

In the passive view, we designate the rotation as the sets of transformation coefficients that identify the relationship between the original and rotated basis vectors,

$$\hat{i}' = \sum_{l=x,y,z} a_{il}\hat{l}, \tag{F.4}$$

where $\hat{i}$ are an orthogonal set of unit vectors in the original configuration and the $\hat{i}'$ are the corresponding unit vectors for the rotated system. In other words, the unit vectors form a rigid frame that is reoriented by the rotation.

The set of nine numbers a_{ij} are a rotation of rigid frames if

$$\sum_{l=x,y,z} a_{il}a_{jl} = \delta_{ij}. \tag{F.5}$$

This condition is a manifestation of the definition of a rotation as a change to an orthogonal three frame, the unit vectors, $\hat{i}$, $i = x, y, z$ to a new set, $\hat{i}'$, $i = x, y, z$ that leave the new set orthogonal and normalized, i. e,

$$\hat{i}' \cdot \hat{j}' = \delta_{ij} \tag{F.6}$$

when the original coordinate unit vectors are themselves orthonormal.

This is only six equations since δ_{ij} is symmetric under the interchange of i and j. Thus from the original nine parameters in a_{ij}, there are three parameters free for the specification of an arbitrary rotation.

Equation F.5 can be recast into matrix form by transposing the indexes of one of the a_{ij}.

$$\delta_{ij} = \sum_{l=x,y,z} a_{il}\tilde{a}_{lj} \tag{F.7}$$

or

$$\mathbf{I} = \mathbf{A}\tilde{\mathbf{A}} \tag{F.8}$$

where $\mathbf{I}$ is the three by three identity matrix, $\mathbf{A}_{ij} = a_{il}$, $\tilde{\mathbf{A}}$ is the transpose of $\mathbf{A}$, and is the inverse of $\mathbf{A}$. An n by n matrix family in which the transpose is the inverse is called an orthogonal set and thus the rotations are usually designated $O(3)$ for the set of 3 by 3 orthogonal matrices. Notice that, since the rotations obviously form a group[1], $\tilde{\mathbf{A}}$ is also a rotation and, since it is the inverse of $\mathbf{A}$, it identifies the meaning of our phrase above of the "reverse" rotation in the discussion of active and passive views of transformations.

In the passive view, the same position, $\vec{X}$, is designated by the original observer as

$$\vec{X} = \sum_{i=x,y,z} i\hat{i} \tag{F.9}$$

and

$$\vec{X} = \sum_{i'=x',y',z'} i'\hat{i}'. \tag{F.10}$$

Using Equation F.5, to invert Equation F.4 and substituting into F.9

$$i' = \sum_{j=x,y,z} a_{i'j} j \tag{F.11}$$

or written in terms of the coordinates

$$\begin{aligned} x' &= a_{xx}x + a_{xy}y + a_{xz}z \\ y' &= a_{yx}x + a_{yy}y + a_{yz}z \\ z' &= a_{zx}x + a_{zy}y + a_{zz}z. \end{aligned} \tag{F.12}$$

Therefore a triplet (V_x, V_y, V_z) is a vector if when recorded by a rotated coordinate system the triplet changes by the rule:

$$V_i' = \sum_{l=x,y,z} a_{il} V_l \tag{F.13}$$

where the a_{il} carry the designations of the rotation in question.

Beside vectors there are higher order forms that emerge in physics. The angular momentum, $\vec{L}$, is related to the angular velocity, $\vec{\omega}$,

$$\vec{L} = I\vec{\omega}.$$

This simple form is valid only when the rotating body is especially symmetric. In the general case, the angular momentum and the angular velocity are directed in different directions. This requires that I be a bivector, $\overleftrightarrow{I}$. Using,

$$\vec{L} = \overleftrightarrow{I} \cdot \vec{\omega} \Rightarrow L_i = \sum_m I_{im} \omega_m. \tag{F.14}$$

[1] A group is a set of elements with an associative product rule, an identity element, and a requirement for an inverse for each element. There are several texts treating the mathematical theory of groups. Excellent introductions for physicists are [**?**] and [**?**]

In a rotated frame the, the same angular momentum vector and angular velocity have the components

$$L'_i = \sum_{l=x,y,z} a_{il} L_l \tag{F.15}$$

and

$$\omega'_i = \sum_{l=x,y,z} a_{il} \omega_l \tag{F.16}$$

where L' and ω'_m are the rotated coordinates of the angular momentum and angular velocity in the rotated coordinate system. Using Equation F.5, this requires that

$$\begin{aligned} L'_i &= \sum_{l=x,y,z} a_{il} L_l = \sum_{l,l'=x,y,z} a_{i,l} I_{ll'} \omega_{l'} = \sum_{l,l',l''=x,y,z} a_{i,l} I_{ll'} a_{l'l''} \omega_{l''} \\ &= \sum_{l=x.y.z} I'_{il} \omega'_l \Rightarrow I'_{lm} = \sum_{k''k'} a_{lk''} a_{mk'} I_{k''k'} \end{aligned} \tag{F.17}$$

Thus, when $\overleftrightarrow{I}$ is evaluated in a frame that is rotated relative to another one, the nine elements of I transform as the direct product of two vectors and thus the name bivector.

Another simple example of a bivector is the identity bivector, $\overleftrightarrow{\mathbf{I}}$ defined by $I_{ij} \equiv \delta_{ij}$ which is the identity matrix in that notation. In a rotated frame

$$I'_{ij} = \sum_{k''k'} a_{ik''} a_{jk'} \delta_{k''k'} = \sum_{k'} a_{ik'} a_{jk'} = \delta_{ij} = I_{ij}. \tag{F.18}$$

Thus the identity bivector is invariant under rotations; it is the same in all rotated frames.

Another name which is used for objects of this kind, the bivectors, is second rank tensor. In Section F.4, we will motivate, develop, and define a more general definition of the rank of a tensor but for now we can stick with this intuitive identification of a vector as a tensor of the first rank. A scalar, a quantity that is unchanged by a rotation of the coordinate system, is a tensor of zero rank. It is obvious the tensors of rank higher than two can easily be defined and, as we will see, will be needed.

Rotations are a special case of a general transformation between coordinates. Instead of considering the rotation as a reorientation of the basis vector directions, the rotation can be considered the result of a simple linear transformation of the coordinates, i. e. all places in E^3 are labeled by the set (x, y, z) where the values of x, y, z range from $-\infty$ to ∞. A rotated coordinate system would have a similar triplet of labels, (x', y', z') where the same point has the label given by

$$\begin{aligned} x' &= a_{xx}x + a_{xy}y + a_{xz}z \\ y' &= a_{yx}x + a_{yy}y + a_{yz}z \\ z' &= a_{zx}x + a_{zy}y + a_{zz}z. \end{aligned} \tag{F.19}$$

Note the difference in the usage of the symbols x, y, z in the two contexts above. When subscripted on the a_{ij}, they are the values that the indices could take and designate the coordinate labels. Here when they appear not as a subscript, they represent the points on the real line – each varies continuously. There should be no confusion with the use of these same labels for the two purposes.

F.3 Lorentz Transformations

F.3.1 Event Four Vector and Four Velocity

In the previous sections, we developed the formalism for the analysis of rotations in Euclidean space, . In this section, we develop the formalism appropriate to a Minkowski space-time, the manifold of space and time that is the basis of the Theory of Special Relativity. As in Euclidean space, a vector formalism is possible. Given an origin event and inertial observer, a coordinate system can be established. An event is a time and a place, a set of four numbers, $(t, \vec{x})$, that specifies that event in that coordinate system. We can designate the coordinates with an index x^μ with $x^0 \equiv ct$, $x^1 \equiv x$, $x^2 \equiv y$, and $x^3 \equiv z$. In this notation, the Lorentz transformations are expressed as

$$x'^\alpha = \sum_{\mu=0}^{3} \Lambda^\alpha_{\ \mu} x^\mu \tag{F.20}$$

with

$$\Lambda^\alpha_{\ \mu} = \begin{pmatrix} \Lambda^0_{\ 0} & \Lambda^0_{\ 1} & \Lambda^0_{\ 2} & \Lambda^0_{\ 3} \\ \Lambda^1_{\ 0} & \Lambda^1_{\ 1} & \Lambda^1_{\ 2} & \Lambda^1_{\ 3} \\ \Lambda^2_{\ 0} & \Lambda^2_{\ 1} & \Lambda^2_{\ 2} & \Lambda^2_{\ 3} \\ \Lambda^3_{\ 0} & \Lambda^3_{\ 1} & \Lambda^3_{\ 2} & \Lambda^3_{\ 3} \end{pmatrix}$$

$$= \begin{pmatrix} \frac{1}{\sqrt{1-\frac{v^2}{c^2}}} & \frac{-\frac{v}{c}}{\sqrt{1-\frac{v^2}{c^2}}} & 0 & 0 \\ \frac{-\frac{v}{c}}{\sqrt{1-\frac{v^2}{c^2}}} & \frac{1}{\sqrt{1-\frac{v^2}{c^2}}} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{F.21}$$

for a Lorentz transformation along the positive x direction with speed v. Other Lorentz transformations are implemented similarly. The rotations which are a subgroup of the Lorentz transformations are the usual rotation elements operating in the bottom three by three spaces in this four by four object. There is some ambiguity in the identification of matrix elements and the two indexed objects, Λ^{μ}_{ν}. The location of indexes will be clarified later but for now the interpretation of Equation F.21 and Equation F.20 is consistent with Equation **??** and the Einstein convention. There is a broadly accepted convention that simplifies the notation considerably called the Einstein convention which eliminates the summation symbol for cases in which the same index appears up and down in the same term of an equation. In this notation, Equation F.20 appears simply as

$$x'^{\alpha} = \Lambda^{\alpha}_{\mu} x^{\mu}. \tag{F.22}$$

In fact, it might appear that the lower second index on Λ^{μ}_{ν} is placed there only to accommodate the Einstein convention. We will find that there is a more important significance to the placement of an index in the upper or lower position.

Given two events we can talk about the interval between them. In this language, there is a four vector interval

$$s^{\mu} \equiv \Delta x^{\mu} = \left(c\left(t_2 - t_1\right), \left(x_2 - x_3\right), \left(y_2 - y_3\right), \left(z_2 - z_3\right)\right). \tag{F.23}$$

The invariant interval squared, Equation **??**, is now expressed as

$$\Delta x^2 = s^{\alpha} g_{\alpha\mu} s^{\mu}, \tag{F.24}$$

where

$$g_{\alpha\mu} \equiv \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{F.25}$$

In this case, the identification of the matrix elements with the first index being the matrix row designator and the second index the column indicator.

This is more direct than in Equation F.21 for the Lorentz transformation. $g_{\alpha\mu}$ is called the metric in Minkowski space. This terminology and the index operations are a special case of the general formalism called indexology in this Appendix F. We will not need the full power of the indexology formalism until we get to transformations that involve the event labels such as gravitation or in the discussion of generalized coordinates. For our purposes now, it will be sufficient to deal with four vectors defined as any quartet of numbers that transform under Lorentz transformations in the same pattern as Equation F.22. Lorentz scalars are quadratic forms of four vectors such as Equation F.24 that are invariant under Lorentz transformations. We will also have use of higher order forms made from the four vectors such as second rank tensors. There are many examples of four vector quantities such as a four velocity, defined below, and the relativistic energy and momentum defined in Chapter **??**. In this section, for quantities other than the event coordinates, we will take the transformation properties of these objects for granted and leave for the later definitions the proof of their transformation properties.

The condition that the interval squared be an invariant places a condition on the form of the Lorentz transformations,

$$\Delta x^2 = \Delta x'^2$$

which implies

$$\Delta x^\alpha g_{\alpha\mu} \Delta x^\mu = \Delta x'^\rho g_{\rho\gamma} \Delta x'^\gamma = \Delta x^\delta \Lambda^\rho{}_\delta g_{\rho\gamma} \Lambda^\gamma{}_\omega \Delta x^\omega$$

for all Δx^μ. This relationship is also true for any four vector. The reader must realize that all the indices in this expression are summed and are thus dummies and can take any convenient label as long as the set runs through the values 0,1,2,3.

Thus we have the condition that

$$g_{\alpha\rho} = \Lambda^\mu{}_\alpha g_{\mu\gamma} \Lambda^\gamma{}_\rho. \tag{F.26}$$

This condition can be interpreted in several ways. It can be used as the defining equation for the Lorentz transformations. The sixteen numbers, $\Lambda^\mu{}_\nu$, are a Lorentz transformation if they satisfy Equation F.26. Equation F.26 is not sixteen equations since $g_{\mu\nu}$ is symmetric. This is ten independent equations which leaves six free parameters. That is just what is needed – three parameters to label a velocity and three parameters to label rotations in a three space. Another interpretation of Equation F.26 is that all Lorentz observers have the same metric. In other words, Equation F.26 is a condition on how the object $g_{\mu\nu}$ transforms; it transforms into itself. It is an invariant tensor.

An alternative notational approach to invariant forms is to define a new set of coordinate four vectors defined by

$$x_\mu \equiv g_{\mu\nu} x^\nu . \tag{F.27}$$

Using this four vector the invariant takes the form $\Delta x_\mu \Delta x^\mu$. This is just the statement that for any linear vector space there exists a dual vector and it is linearly related to the original vector space. To facilitate the manipulation of both the upper and lower cased index objects, it is relevant to introduce the inverse metric tensor,

$$\left(g^{-1}\right)_{\mu\nu} \equiv g^{\mu\nu}, \tag{F.28}$$

or

$$g^{\mu\nu} g_{\nu\mu'} = \delta^\mu_{\ \mu'} . \tag{F.29}$$

where

$$\delta^\mu_{\ \mu'} \equiv \begin{cases} 1 & : \quad \mu = \mu' \\ 0 & : \quad \text{otherwise} \end{cases} . \tag{F.30}$$

For our case of the Minkowski metric,

$$\left(g^{-1}\right)_{\mu\nu} \equiv g^{\mu\nu} = g_{\mu\nu} . \tag{F.31}$$

This greatly simplifies the interpretation of the lower indexed four vector entities. It will not be the case when we deal with relativity as geometry.

The x_μ transform as a covariant four vector which is indicated as $\binom{0}{1}$ and the x^μ transform as a $\binom{1}{0}$. The metric indicated as $g_{\mu\nu}$ is a $\binom{0}{2}$ and $g^{\mu\nu}$ is a $\binom{2}{0}$. The Lorentz transform, $\Lambda^\mu_{\ \alpha}$ is a $\binom{1}{1}$.

For a time-like trajectory, we can define a four vector velocity by using the proper time over the trajectory to calculate a rate of change. In other words, a trajectory which is a connected set of events which would be coordinatized by some inertial observer as $(\vec{x}(t), t)$, can be parametrized by the elapsed proper time of the time-like trajectory, $(\vec{x}(\tau), t(\tau))$, where

$$\begin{aligned} \tau\left[trajectory : (\vec{x}_0, t_0; \vec{x}, t)\right] &\equiv \int_{traj.,\vec{x}_0,t_0}^{\vec{x},t} d\tau \\ &= \int_{traj.,\vec{x}_0,t_0}^{\vec{x},t} \sqrt{dt^2 - \frac{d\vec{x}\cdot d\vec{x}}{c^2}} \end{aligned}$$

$$= \int_{traj.,\vec{x}_0,t_0}^{\vec{x},t} \sqrt{1 - \frac{\frac{d\vec{x}}{dt} \cdot \frac{d\vec{x}}{dt}}{c^2}} dt$$

$$= \int_{traj.,\vec{x}_0,t_0}^{\vec{x},t} \sqrt{1 - \frac{\vec{v} \cdot \vec{v}}{c^2}} dt. \qquad \text{(F.32)}$$

where $\vec{v} \equiv \frac{d\vec{x}}{dt}$ is called the coordinate velocity and is the usual definition of velocity. This may look like a rather complex object but this construction is much like the parametrizing of a curve in two space with distance along the curve, see Section **??**. The notation is also the same as that in Section **??**. The elapsed proper time is a functional of the trajectory but is a function of the labels of the events at the end points of the integral. Since it is a function of the time on the trajectory we can derive a differential form for Equation F.32. Differentiating with respect to the time of the event on the worldline,

$$\frac{d\tau}{dt} = \sqrt{1 - \frac{\vec{v} \cdot \vec{v}}{c^2}} \qquad \text{(F.33)}$$

Putting all this together, we can construct a four vector velocity,

$$u^\mu \equiv \frac{dx^\mu}{d\tau} = \frac{\frac{dx^\mu}{dt}}{\frac{d\tau}{dt}}, \qquad \text{(F.34)}$$

which, since the Lorentz transforms are linear and constant, transforms the same way as x^μ in Equation F.22. In addition, note that all for components of u^μ have the dimensions of a velocity.

The construction of other kinematic four vectors such as a four acceleration follows the same pattern,

$$a^\mu \equiv \frac{du^\mu}{d\tau} = \frac{d^2 x^\mu}{d\tau^2}. \qquad \text{(F.35)}$$

By construction of the proper time, it follows that the four velocity vector length is always the same,

$$u^\mu g_{\mu\nu} u^\nu = \frac{\frac{dx^\mu}{dt} g_{\mu\nu} \frac{dx^\nu}{dt}}{\left(\frac{d\tau}{dt}\right)^2} = \frac{\left(c^2 - \vec{v} \cdot \vec{v}\right)}{\left(1 - \frac{\vec{v}\cdot\vec{v}}{c^2}\right)} = c^2. \qquad \text{(F.36)}$$

Any four vector with a positive length squared such as the four velocity is called time-like four vector; there always exists a Lorentz frame in which the four vector takes the form $\left(c, \vec{0}\right)$. In the general frame, the four velocity takes the form

$$u^\mu = \left(\frac{c}{\sqrt{1-\frac{v^2}{c^2}}}, \frac{\vec{v}}{c\sqrt{1-\frac{v^2}{c^2}}}\right). \tag{F.37}$$

Differentiating Equation F.36,

$$\frac{d}{d\tau}\left(u^\mu g_{\mu\nu} u^\nu\right) = 0 = 2a^\mu g_{\mu\nu} u^\nu. \tag{F.38}$$

In the frame in which $u^\mu = \left(c, \vec{0}\right)$, the four acceleration must take the form $a^\mu = (0, \vec{a})$, where $\vec{a}$ is the usual acceleration. We derive this result below in F.40.

Since the length squared is a Lorentz invariant,

$$a^\mu g_{\mu\nu} a^\nu > 0. \tag{F.39}$$

The acceleration is a space-like four vector. Differentiating Equation F.37, the general form of the acceleration four vector is

$$\begin{aligned}
\frac{du^\mu}{d\tau} &= \left(\frac{\vec{v}\cdot\frac{d\vec{v}}{d\tau}}{\left\{1-\frac{v^2}{c^2}\right\}^{\frac{3}{2}}}, \frac{\frac{d\vec{v}}{d\tau}}{\sqrt{1-\frac{v^2}{c^2}}} + \vec{v}\frac{\vec{v}\cdot\frac{d\vec{v}}{d\tau}}{\left\{1-\frac{v^2}{c^2}\right\}^{\frac{3}{2}}}\right) \\
&= \left(\frac{\vec{v}\cdot\frac{d\vec{v}}{dt}}{\left\{1-\frac{v^2}{c^2}\right\}^{\frac{3}{2}}}, \frac{\frac{d\vec{v}}{dt}}{\sqrt{1-\frac{v^2}{c^2}}} + \vec{v}\frac{\vec{v}\cdot\frac{d\vec{v}}{dt}}{\left\{1-\frac{v^2}{c^2}\right\}^{\frac{3}{2}}}\right)\frac{dt}{d\tau} \\
&= \left(\frac{\frac{\vec{v}}{c}\cdot\frac{d\vec{v}}{dt}}{\left(1-\frac{v^2}{c^2}\right)^2}, \frac{\frac{d\vec{v}}{dt}}{\left(1-\frac{v^2}{c^2}\right)} + \frac{\vec{v}}{c}\frac{\vec{v}\cdot\frac{d\vec{v}}{dt}}{\left(1-\frac{v^2}{c^2}\right)^2}\right)
\end{aligned} \tag{F.40}$$

In the frame in which $u^\mu = \left(c, \vec{0}\right)$, i. e. $\vec{v} \to 0$, the acceleration four vector is $(0, \vec{a})$, where $\vec{a} = \frac{d\vec{v}}{dt}$ is the coordinate acceleration as measured by a comoving inertial observer. This is the acceleration of Newtonian physics.

F.4 General Coordinates Labels

Our example of rotations is a special case, a linear transformation, of the generalization

$$q_1 = f_{q_1}(x, y, z)$$

$$\begin{aligned} q_2 &= f_{q_2}(x,y,z) \\ q_3 &= f_{q_3}(x,y,z) \end{aligned} \tag{F.41}$$

where the triplet (q_1, q_2, q_3) is the new designation of the points in E^3. These may have different ranges of values from the (x, y, z) triplet and may only cover a portion of E^3. These more general transformations are the basis of curvilinear coordinates.

The coordinates (x, y, z) is a complete coordinate system on E^3; in a sense, it defines E^3. We can use the coordinates in such a way that they produce an orthogonal coordinate system at each place. By moving along any of the coordinate directions you can generate a unit vector and with the three coordinate directions you can construct an orthogonal frame such as that which was used in Section F.3.1. This construction is especially easy in E^3 with (x, y, z) since the coordinates are lengths and, not only are the orientation of the vectors determined by the coordinates, but the lengths are also. Later we will make a more detail analysis of the idea of lengths but again for now the intuitive ideas will suffice. The important point for now is to note that a tangent frame at any point can be constructed using the (x, y, z) coordinates. More general tangent systems can be constructed using transformations such as Equation F.41 to construct other tangent frames at each point. For instance, In the case of E^3 using the rotated coordinates of Equation F.19 we can generate an infinity of coordinate frames such that at each point you can generate at each point is the same as the one that would be generated at the origin in the sense that they are all simple translates of each other.

Even in the simple space E^3, it is possible to find other orthogonal coordinate systems that may have some advantages in different contexts. An example is the well known spherical polar coordinate system. With some minor ambiguities, this coordinate system covers E^3 nicely. The relationship between it and the rectangular coordinate system is

$$\begin{aligned} r &= \sqrt{x^2+y^2+z^2} \\ \theta &= \tan^{-1}\left(\frac{z}{\sqrt{x^2+y^2}}\right) \\ \phi &= \tan^{-1}\left(\frac{y}{x}\right). \end{aligned} \tag{F.42}$$

The inverse transformations are

$$x = r\cos(\phi)\sin(\theta)$$

$$\begin{aligned} y &= r\sin(\phi)\sin(\theta) \\ z &= r\cos(\theta). \end{aligned} \tag{F.43}$$

In both these sets of equations, Equations F.42 and F.43, r has the range $0 < r < \infty$, θ has the range $0 < \theta < \pi$, and ϕ has the range $0 < \phi < 2\pi$.

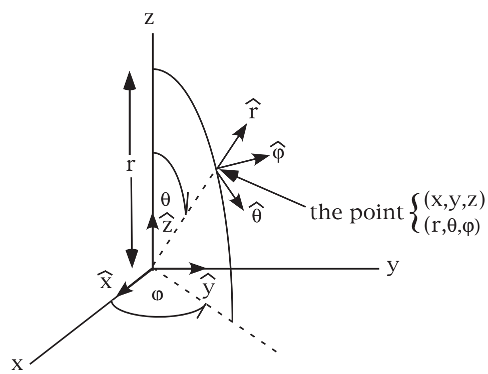

Figure F.1: At each point in E^3, you can locate a point with either the coordinates (x, y, z) or (r, θ, ϕ) related by Equations F.42 and F.43. At a point labeled (x, y, z) or (r, θ, ϕ), you can construct an orthogonal coordinate frame. For clarity, the frame constructed using (x, y, z) is shown at the origin. All the frames constructed by (x, y, z) are translates of each other.

The spherical polar coordinate system generates an orthogonal set of unit vectors at each point, see Figure F.1. For a general point (r, θ, ϕ), we construct the unit vectors directly by incrementing each of the these coordinates. The vectors generated by this action are

$$\hat{r} \equiv \frac{\Delta x_{\Delta r}\hat{x} + \Delta y_{\Delta r}\hat{y} + \Delta z_{\Delta r}\hat{z}}{\sqrt{(\Delta x_{\Delta r})^2 + (\Delta y_{\Delta r})^2 + (\Delta z_{\Delta r})^2}} \tag{F.44}$$

where $\Delta x_{\Delta r}$ is the increment of x for an increment of r and thus:

$$\begin{aligned} \Delta x_{\Delta r} &\equiv \cos(\phi)\sin(\theta)\,\Delta r \\ \Delta y_{\Delta r} &\equiv \sin(\phi)\sin(\theta)\,\Delta r \\ \Delta z_{\Delta r} &\equiv \cos(\theta)\,\Delta r, \end{aligned} \tag{F.45}$$

and

$$\hat{\theta} \equiv \frac{\Delta x_{\Delta\theta}\hat{x} + \Delta y_{\Delta\theta}\hat{y} + \Delta z_{\Delta\theta}\hat{z}}{\sqrt{(\Delta x_{\Delta\theta})^2 + (\Delta y_{\Delta\theta})^2 + (\Delta z_{\Delta\theta})^2}} \tag{F.46}$$

where we now increment θ:

$$\begin{aligned}\Delta x_{\Delta\theta} &\equiv r\cos(\phi)\cos(\theta)\,\Delta\theta \\ \Delta y_{\Delta\theta} &\equiv r\sin(\phi)\cos(\theta)\,\Delta\theta \\ \Delta z_{\Delta\theta} &\equiv -r\sin(\theta)\,\Delta\theta, \end{aligned} \tag{F.47}$$

and

$$\hat{\phi} \equiv \frac{\Delta x_{\Delta\phi}\hat{x} + \Delta y_{\Delta\phi}\hat{y} + \Delta z_{\Delta\phi}\hat{z}}{\sqrt{(\Delta x_{\Delta\phi})^2 + (\Delta y_{\Delta\phi})^2 + (\Delta z_{\Delta\phi})^2}} \tag{F.48}$$

where the increment is in ϕ:

$$\begin{aligned}\Delta x_{\Delta\phi} &\equiv -r\sin(\phi)\sin(\theta)\,\Delta\phi \\ \Delta y_{\Delta\phi} &\equiv r\cos(\phi)\sin(\theta)\,\Delta\phi \\ \Delta z_{\Delta\phi} &\equiv o. \end{aligned} \tag{F.49}$$

Thus the relationship between the orthogonal frame generated by (x, y, z) and the (r, θ, ϕ) at the same point is

$$\begin{aligned}\hat{r} &= \cos\phi\sin\theta\hat{x} + \sin\phi\sin\theta\hat{y} + \cos\theta\hat{z} \\ \hat{\theta} &= \cos\phi\cos\theta\hat{x} + \sin\phi\cos\theta\hat{y} - \sin\theta\hat{z} \\ \hat{\phi} &= -\sin\phi\hat{x} + \cos\phi\hat{y} \end{aligned} \tag{F.50}$$

Note that the although the orthogonal triplet formed at a point (r, θ, ϕ) are normalized, the length of a vector in this resolution is

$$\begin{aligned}\Delta\vec{r} &= \Delta\vec{r}_{\Delta r}\hat{r} + \Delta\vec{r}_{\Delta\theta}\hat{\theta} + \Delta\vec{r}_{\Delta\phi}\hat{\phi} \\ &= \Delta r\hat{r} + r\Delta\theta\hat{\theta} + r\sin(\theta)\,\Delta\phi\hat{\phi} \end{aligned} \tag{F.51}$$

or

$$\Delta\vec{r}\cdot\Delta\vec{r} = (\Delta r)^2 + r^2(\Delta\theta)^2 + r^2\sin^2(\theta)(\Delta\phi)^2. \tag{F.52}$$

This introduces the idea of a metric on the coordinate system. In the case of the general coordinate system, the 6 coefficients of a symmetric second rank tensor $g_{ij}(l, m, n)$ determine the length via

$$\Delta\vec{r}\cdot\Delta\vec{r} = \sum_{i,j} g_{ij}(l, m, n)(\Delta i)(\Delta j). \tag{F.53}$$

Here the complexity of the notation of using the indices i, j, l. m, and n as both coordinate label and continuous variable that is the coordinate value is manifest.

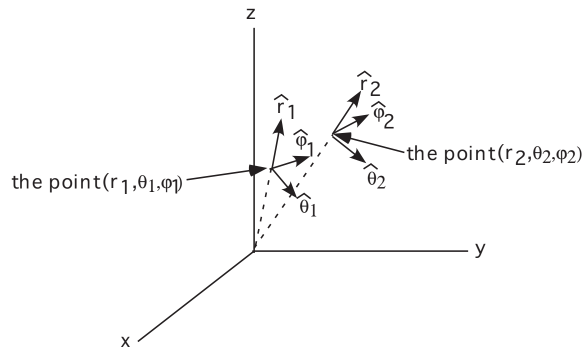

Figure F.2: The orthogonal frames generated at different points using a spherical polar coordinate system are not translates of each other. A frame generated at the point (r_1, θ_1, ϕ_1) is translated and rotated relative to the frame generated at the point (r_2, θ_2, ϕ_2).

More significantly, orthogonal frames at different points (r_1, θ_1, ϕ_1) and (r_2, θ_2, ϕ_2) are not translates of each other. Figure F.2 shows the two orthogonal frames at two points generated by the (r, θ, ϕ) coordinate system. **The frames generated by translation of the coordinates are not translates of each other.** At this stage we can see the heart of the problem. In a given coordinate system, keeping track of the directional information required for a vector field is complicated by the reorientation of the coordinate basis for different points.

The relationship between the orthogonal frames at two points is

$$\begin{aligned}
\hat{r_2} &= (\cos(\phi_2 - \phi_1)\sin\theta_2\sin\theta_1 + \cos\theta_2\cos\theta_1)\,\hat{r_1} \\
&\quad + (\cos(\phi_2 - \phi_1)\sin\theta_2\cos\theta_1 - \cos\theta_2\sin\theta_1)\,\hat{\theta_1} \\
&\quad + \sin(\phi_2 - \phi_1)\sin\theta_2\hat{\phi_1} \\
\hat{\theta_2} &= (\cos(\phi_2 - \phi_1)\cos\theta_2\sin\theta_1 - \sin\theta_2\cos\theta_1)\,\hat{r_1} \\
&\quad + (\cos(\phi_2 - \phi_1)\cos\theta_2\cos\theta_1 + \sin\theta_2\sin\theta_1)\,\hat{\theta_1} \\
&\quad + \sin(\phi_2 - \phi_1)\cos\theta_2\hat{\phi_1} \\
\hat{\phi_2} &= -\sin(\phi_2 - \phi_1)\sin\theta_1\hat{r_1} \\
&\quad - \sin(\phi_2 - \phi_1)\cos\theta_1\hat{\theta_1} \\
&\quad + \cos(\phi_2 - \phi_1)\,\hat{\phi_1}
\end{aligned} \tag{F.54}$$

The reader should check that this is, in fact, a rotation as described in Section F.3.1 in Equation F.5 where the labels i' and l in Equation F.4 range over the sets $i' = r_2, \theta_2, \phi_2$ and $l = r_1, \theta_1, \phi_1$.

We will use this example to motivate our study of vector and tensor fields. It should be clear that this is a choice of convenience and that there are many other coordinate choices.

F.5 Fields and Manifolds

A field is a very fundamental concept in physics. It is an entity that is defined over a manifold. Think of our usual three space, E^3. At every place there is quantity, say the temperature, that is measured. This collection of temperatures is a field. If this quantity that is measured over the points of the manifold is a vector or tensor, we have a vector or tensor field. In this case, how do you construct the directional information required for vectors and tensors at different places?

We find our way through the manifold by assigning coordinates to all the points. For example in E^3, three independent labels are required. One immediately thinks of the triplet (x, y, z), three R^1s.

The first point to note is that, if all you have are the coordinates, and you did not know that the field had been defined on an E^3 which has that nice covering made of the three independent R^1s all of whose frames are parallel, you would have to be careful dealing with questions of parallel vectors and so forth because the directions of the orthogonal frames change at different points on the manifold. Also note the the lengths of the coordinate displacement vectors vary depending on where you are in the coordinate system, see Equations F.51, and F.52. When is a vector field at one point parallel to a vector field at another point? When is the field the same?

In order to proceed to analyze these questions, we will need the idea of the geodesic. A geodesic is a path in the space that is a straight as possible. For any two points in the space, there is an infinity of paths connecting the points. We can find the path that minimizes the distance traveled. Consider a path parameterized by a running variable t. Then the length of the path between two points is

$$L = \int_a^b \sqrt{\frac{d\vec{r}}{dt} \cdot \frac{d\vec{r}}{dt}} dt = \int_a^b \sqrt{\sum_{i,j} g_{ij}\left(l, m, n\right) \left(\frac{di}{dt}\right) \left(\frac{dj}{dt}\right)} dt \tag{F.55}$$

where again the i and j play the role of being an index label and the coordinate value where appropriate. Although it is obvious, if we know the distances that are generated by coordinate increments, we can find the length of any path. This is our first notice of the important role of the metric equation, Equation F.53. You can find the shortest, straight, path from a construction that deals only with the metric. It will turn out that all of the geometry of the system can be realize using only information that is contained in the metric, see Section **??**.

F.5.1 Tensor Fields and General Transformations

A transformation from one coordinate system to another will change both the orientation and scale of the directed quantities.

F.5.2 Nomenclature

The First fundamental Form – the metric– is

$$g_{ij} \equiv \vec{x}_i \cdot \vec{x}_j. \tag{F.56}$$

where $\vec{x}_i \equiv \frac{\partial \vec{x}}{\partial q^i}$.

The second fundamental form – b_{ij} – is defined by the relations

$$b_{ij} \equiv \vec{x}_{ij} \cdot \hat{n} \tag{F.57}$$

where $\vec{x}_{ij} \equiv \frac{\partial^2 \vec{x}}{\partial q^i \partial q^j}$. It is the components of the second derivatives of the position with respect to the coordinates in the normal direction.

$$\hat{n}_i = \frac{\partial \hat{n}}{\partial q^i} = -b_i^j \vec{x}_j \tag{F.58}$$

and $b_{ij} = b_i^l g_{lj}$. This is possible since the unit vector, $\hat{n}$ must have its derivative in the tangent plane and thus be expandable in the tangent vectors.