The Poor Man's Introduction to Tensors

Justin C. Feng

Center for Relativity

University of Texas, Austin

Abstract

When solving physical problems, one must often choose between writing formulas in a coordinate independent form or a form in which calculations are transparent. Tensors are important because they allow one to write formulas in a manner that is both coordinate independent and transparent in performing calculations. In these notes, I provide an informal introduction to tensors (in Euclidean space) for those who are familiar with the basics of linear algebra and vector calculus.

Introductory Remarks

These notes are intended for a general audience; the only prerequisites for these notes is a basic knowledge of linear algebra and vector calculus. I also intend to make these notes self-contained (aside from the prerequisites), so I have done my best to build the subject up from first principles. Since this is an introduction to tensor calculus for a broad audience, I will avoid discussing General Relativity and non-Euclidean geometry. Instead, I hope to give you a basic understanding of what tensors are, and to also give you the tools you will need to:

- Write Partial Differential Equations (in Euclidean space) as tensorial equations
- Write coordinate-independent expressions for integrals over Euclidean space
- Efficiently convert Partial Differential Equations and integral expressions between different coordinate systems

If you find any errors or have suggestions for these notes, you may contact me at (jcfeng@physics.utexas.edu).

Enjoy!

Contents

Tensors Condensed

When learning a new topic, I often find it helpful to identify the central idea(s)/principle(s) first; I usually get more out of the topic when I do so. For your convenience, I present to you, in a single paragraph, the central ideas in tensor analysis:

Simply put, a tensor is a mathematical construction that "eats" a bunch of vectors, and "spits out" a scalar. The central principle of tensor analysis lies in the simple, almost trivial fact that scalars are unaffected by coordinate transformations. From this fact, we obtain the main result

of tensor analysis: an equation written in tensor form is valid in any coordinate system.

In my attempt to condense tensor analysis into a single paragraph, I have left out many important details. For instance, my definition in the first paragraph is an incomplete one; the definition leaves out the fact that tensors are linear maps, as well as the fact that they also "eat" other objects called dual vectors. You will find such details in the remainder of these notes.

Indices

An index is simply a way to organize a collection of numbers. For example, consider a vector in three-dimensional Euclidean space, which I'll call $\vec{v}$. In order to express $\vec{v}$ in three dimensions, I need to specify coordinates. For simplicity, I'll use the Cartesian coordinates (x, y, z). The vector can then be expressed with the three numbers (v_x, v_y, v_z), so that:

$$\vec{v} = (v_x, \ v_y, \ v_z) \tag{1}$$

The number v_x is the vector's component in the x-direction, v_y the vector's component in the y-direction, and so on. If I wanted to write a vector in an arbitrary number of dimensions, say 89 dimensions for example, I would run out of letters to label my coordinates, so I'd like to label my coordinates with numbers instead. For example, I can write:

$$
\begin{aligned}
x^1 &= x \\
x^2 &= y \\
x^3 &= z
\end{aligned}
\tag{2}
$$

Be careful; *the numbers* $(1, 2, 3)$ *in the equations above are not exponents*! The numbers $(1, 2, 3)$ are simply labels telling you which coordinate you're talking about--x^2 does not mean $x \cdot x$, the "2" in the superscript simply means that you're talking about coordinate 2.

One can, in a similar fashion, write the components of $\vec{v}$ (or the numbers (v_x, v_y, v_z)) in the following way:

$$
\begin{aligned}
v^1 &= v_x \\
v^2 &= v_y \\
v^3 &= v_z
\end{aligned}
\tag{3}
$$

So that v^1 is now the vector's component in the x^1 direction, v^2 is the vector's component in the x^2 direction, and so on. Suppose I have another vector $\vec{F} = \left(F^1, \ F^2, \ F^3\right)$. When I write the following formula (where b is some positive number):

$$\vec{F} = -b \, \vec{v} \tag{4}$$

I'm actually writing down the following three equations:

$$
\begin{aligned}
F^1 &= -b \, v^1 \\
F^2 &= -b \, v^2 \\
F^3 &= -b \, v^3
\end{aligned}
\tag{5}
$$

One can note that the above three equations look exactly the same. If I introduce a placeholder label i, which can represent any of the numbers $(1, 2, 3)$, then I can write the three equations in a compact form:

$$F^i = -b \, v^i \tag{6}$$

The placeholder label i is called an index.

Indices (the plural of index) can be used to represent vector quantities; instead of writing arrows on top of letters (or bolding the letters), one can stick an index on a letter, so that v^i and F^i represent the vectors $v^i = \vec{v} = \left(v^1, \ v^2, \ v^3\right)$ and $F^i = \vec{F} = \left(F^1, \ F^2, \ F^3\right)$.

At this point, all I've done is invent some notation for vectors. Index notation, however, can also be used to represent matrices. For example, a matrix $\boldsymbol{M}$ in three dimensions can be represented by the following table of numbers:

$$
\boldsymbol{M} =
\begin{pmatrix}
M^{11} & M^{12} & M^{13} \\
M^{21} & M^{22} & M^{23} \\
M^{31} & M^{32} & M^{33}
\end{pmatrix}
\tag{7}
$$

Using index notation, one can write the matrix as M^{ij}. In index notation, the eigenvector formula becomes:

$$\sum_j M^{ij}\, v^j = \lambda\, v^i \tag{8}$$

Now at this point, I want to mention that the proper way to express a matrix in index notation is not M^{ij}, but $M^i{}_j$. I will clarify the importance of index placement a bit in the next section, and more completely when discussing dual vectors and tensors.

Index Placement and Einstein Summation Convention

You may have noticed that I've always written vectors v^i with raised indices. To avoid confusing indices with exponents, it would make more sense to use subscripts instead of superscripts. However, I've done this for a reason (and it isn't because I want to be confusing). The index placement (whether the index is a superscript or subscript) will become extremely important later one, because it is used to distinguish between two types of objects: vectors v^i and dual vectors w_i. By the way, vectors v^i are sometimes called *contravariant vectors*, and dual vectors w_i are sometimes called *covariant vectors* or *one-forms*.

Throughout these notes, I use Einstein Summation Convention. This stuff confused me greatly the first time I saw it, but it just boils down to a simple rule:

Any time you see a pair of indices with the same symbol (one raised and one lowered), a sum is implied.

Explicitly, for some collection of numbers T_{ij}, this means the following:

$$T_{ij}\, v^j = \sum_j T_{ij}\, v^j \tag{9}$$

When working with tensors, the repeated indices (the index j in the equation above) always come in a pair. If they don't, you're either writing down something that doesn't make sense, or you did something wrong in your calculation.

Also, for reasons that I'll discuss later, one of the repeated indices should be raised (like the j in v^j), and the other should be lowered (like the j in T_{ij}).

When a matrix acts on a vector, the result must also be a vector, not a dual vector. This fact, combined with Einstein summation convention, means that the components of matrices $M^i{}_j$ should be written with one index raised (i) and one index lowered (j). Then, using Einstein summation convention, equation (8) can be written:

$$M^i{}_j\, v^j = \lambda\, v^i \tag{10}$$

Vectors

Before I can tell you what a tensor is, I must tell you what a vector really is; in fact, you will later see that a vector is a type of tensor.

A vector is simply a directional derivative. Before you write me off as a nut, take a look at the directional derivative of some scalar function $f(x^j)$:

$$v \cdot \nabla f(x^j) = v^i\, \frac{\partial}{\partial x^i} f(x^j) \tag{11}$$

Where x^j are my coordinates. If I remove the function, my directional derivative operator looks like:

$$v \cdot \nabla = v^i\, \frac{\partial}{\partial x^i} \tag{12}$$

In short, I'm claiming that I can write a vector in the basis of partial derivatives with respect to the coordinates (this basis is called the coordinate basis). In other words, I claim that the directional derivative contains the same information as the the components v^i. To see whether this claim is true, suppose I want to get the v^1 component of the vector. To do this, I'll feed the following function into the directional derivative operator:

$$f(x^j) = x^1 \tag{13}$$

I then get:

$$v \cdot \nabla x^1 = v^i \frac{\partial x^1}{\partial x^i} \tag{14}$$

Since the coordinates are independent of each other, $\frac{\partial x^a}{\partial x^b} = 0$ if $a \neq b$, and $\frac{\partial x^a}{\partial x^b} = 1$ if $a = b$. Then, one can write:

$$\frac{\partial x^a}{\partial x^b} = \delta_b^a \tag{15}$$

Where δ_b^a is the *Kronecker delta*, often defined by the following property (which implies that the Kronecker delta is basically the identity matrix):

$$
\begin{array}{ll}
\delta_b^a = 1 & \text{If } a = b \\
\delta_b^a = 0 & \text{If } a \neq b
\end{array}
\quad\Rightarrow\quad
\delta_b^a = \begin{pmatrix} \delta_1^1 & \delta_2^1 & \delta_3^1 \\ \delta_1^2 & \delta_2^2 & \delta_3^2 \\ \delta_1^3 & \delta_2^3 & \delta_3^3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{16}
$$

Equation (14) can then be written as:

$$v \cdot \nabla x^1 = v^i \delta_i^1 = v^1 \tag{17}$$

This means that all I have to do to pick out a component of a vector is to feed the corresponding coordinate into the directional derivative operator. More generally:

$$v^i = v \cdot \nabla x^i \tag{18}$$

Some people get rid of the function f altogether, and write the following:

$$v = v^i \frac{\partial}{\partial x^i} \tag{19}$$

with the components are defined as:

$$v^i = v(x^i) \tag{20}$$

where the right hand side $v(x^i)$ is to be interpreted as the directional derivative in equation (19) acting on the coordinate x^i.

Comments for the more mathematically minded:

Vectors can be defined similarly in non-Euclidean geometries, such as curved surfaces. In non-Euclidean geometries, you can't move vectors around freely like you can in Euclidean space with Cartesian coordinates, so you have to define your vectors at a point:

$$v_p(f) = \left(v^i \frac{\partial f}{\partial x^i} \right)_p \tag{21}$$

Where p denotes the point at where the vector is defined (the right hand side of the equation is to be evaluated at p).

In Cartesian coordinates, one can get away with defining a vector without reference to a point (all you need is a magnitude and a direction, right?), but this ambiguity is no longer acceptable when working in coordinates where the basis vectors $\frac{\partial}{\partial x^i}$ change direction from point to point.

Furthermore, the notion of a single vector (or direction, for that matter) being defined for an entire space breaks down in spaces with non-Euclidean geometries because the geometry itself can change from point to point (even if the geometry of the space is uniform, the topology of the space can ruin this; try to globally define a vector for a Mobius strip or the surface of a 2-sphere--you'll run into contradictions). The closest thing you can have in a non-Euclidean space is a vector field.

More precisely, vectors are objects that live in the *tangent space* of a point *p* in a *differentiable manifold*. A *manifold M* is some *n*-dimensional space that has that topological properties of Euclidean space $\mathbb{R}^n$ near a point (this is what is meant by the term: "locally Euclidean"). A *differentiable manifold* is a manifold that has all the properties you need in order to do calculus on it. The *tangent space* of a point *p* (often denoted $T_p M$) is the space of all directional derivative operators at *p* (equation 21).

If you want to see a better explanation of this stuff, here are some references:

- Bernard F. Schutz, *Geometrical methods of mathematical physics*, Cambridge University Press 1980
- John C. Baez & Javier P. Muniain, *Gauge Fields, Knots, and Gravity*, World Scientific 1994
- Theodore Frankel, *The Geometry of Physics: An Introduction*, 3rd Ed., Cambridge University Press 2011

The Metric as a Generalized Dot Product

One can define an inner product (or dot product) between two vectors *u* and *v* using something called a *metric*;

$$\langle u, v \rangle = g_{ij} u^i v^j \tag{22}$$

Where g_{ij} are the components of the metric.

One can define an *inverse metric* g^{ij} using the following formula:

$$g^{ij} g_{jk} = \delta^i_k \tag{23}$$

Recall that the Kronecker delta δ^a_b defined in Equation (16), is just the identity matrix; this justifies the term "inverse metric" for g^{ij}. To confuse others, some people use the term "metric" for both.

By the way, in Cartesian coordinates on Euclidean space, the metric can be written as the following table (it's not exactly a matrix, since both indices are lowered):

$$g_{ij} = \begin{pmatrix} g_{11} & g_{12} & g_{13} \\ g_{21} & g_{22} & g_{23} \\ g_{31} & g_{32} & g_{33} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{24}$$

When this metric is used, you have your standard dot product for vectors in Euclidean space. Since this is the identity, the inverse has exactly the same form; the only difference is that the indices are raised.

In the literature, one often finds that the metric tensor is given as a line element:

$$ds^2 = g_{ij} \, dx^i \, dx^j$$

which can be thought of as the norm $\langle dx, dx \rangle$ of some infinitesimal displacement vector dx^i. Line elements provide nice geometric interpretation for the metric, and are useful for certain problems (like calculating the length of a curve, for instance). However, I don't want to spend too much time discussing line elements, since I feel that they distract the reader from the main ideas in these notes.

Dual Vectors

A *dual vector* (also called a 1-form* or a covector/covariant vector) is simply a mathematical object is the that "eats" a vector and "spits out" a scalar (or a real number). Basically, it does the following:

$$w(v) = w_i v^i \tag{25}$$

Where w_i are the components of the dual vector. Note that in general, dual vectors are only defined at the same point *p* that the vector is defined. Furthermore, vectors are also "dual" to the dual vectors; vectors also "eat" dual vectors and "spit out" scalars in exactly the same way:**

$$v(w) = v^i w_i \tag{26}$$

The basis elements for the dual vector are the coordinate differentials dx^i, which are related to the coordinate basis vectors by the following expression:

$$dx^j \left(\frac{\partial}{\partial x^i} \right) = \frac{\partial}{\partial x^i} (dx^j) = \delta_i^j \tag{27}$$

Remember, dual vectors "eat" vectors and spit out scalars (recall equation (25)); the formula above states that when the basis element dx^j "eats" a basis vector $\frac{\partial}{\partial x^i}$, it spits out 1 or 0 depending on whether $i = j$.

Thus, the dual vector can be written (compare with equation (19)):

$$w = w_i \, dx^i \tag{28}$$

Using the metric g_{ij}, one can turn a vector v^i into a dual vector v_j:

$$v_j = g_{ij} \, v^i \tag{29}$$

One can see that feeding the vector v to this dual vector gives the norm of the vector:

$$v_i \, v^i = \langle v, \, v \rangle = |v|^2 \tag{30}$$

The metric takes vectors and turns them into dual vectors. It would be nice to invert the process; I want to take a dual vector, and turn it into a normal vector. One way to do this is to define the inverse of the metric (metrics are invertible):

$$g^{ij} \, g_{jk} = \delta_k^i \tag{31}$$

The Kronecker delta is basically the identity matrix, if you prefer to think about two-indexed objects as matrices. Anyway, the inverse metric g^{ij} can be used to turn a dual vector w_j into a vector w^i:

$$w^i = w_j \, g^{ij} \tag{32}$$

The process of using metrics and inverse metrics to change vectors into dual vectors and vice versa, is referred to as raising and lowering of indices.

If you recall linear algebra, it might seem that all we're doing here is converting column vectors into row vectors and vice versa. What we're doing here is similar, but its best to keep in mind that vectors and dual vectors are two different things--one can only relate the two if there is a metric.†

Finally, to give you some intuition for what a dual vector does, consider the differential of some function $f(x^i)$:

$$df = \frac{\partial f}{\partial x^i} \, dx^i \tag{33}$$

The differential df is an example of a dual vector, with components $\frac{\partial f}{\partial x^i}$ (compare the above with equation (28)). Now feed it some vector v^i:

$$df(v) = \frac{\partial f}{\partial x^i} \, v^i \tag{34}$$

Now pretend for a moment that v^i represents the components of some displacement vector, so that $v^i = \Delta x^i$ (we momentarily forget the discussion surrounding equation (21)). Then, one can interpret $df(v)$ as the change in the value of the function f between the tip and the tail of the displacement vector $v^i = \Delta x^i$: §

$$df(v) \sim f(\text{tip of } v) - f(\text{tail of } v) \tag{35}$$

----

*Dual vectors are often referred to as 1-forms, since they are an example of a differential form. Differential forms are useful in defining integration on manifolds, but I won't discuss differential forms in detail in these notes (I'll save that for another set of notes).

** Mathematicians may consider this sloppy; the dual of a dual vector is a double dual, not a vector. However, the (vector) space of double duals is isomorphic to the tangent space that the vector lives in, so it posesses the same algebraic structure. This is why it's okay for physicists to treat double duals as if they're vectors.

† Recall the form of the metric for Euclidean space in Cartesian coordinates (Equation (24)). In this case, the components of the dual vector v_j are exactly the same as the components of the vector v^i. This is why many people don't care about index placement (whether indices are lowered or raised) or the distinction between vectors and their duals, because it doesn't matter in Cartesian coordinates on Euclidean space. However, it becomes important when you work in curvilinear coordinates, and even more so when you work in curved spaces.

§ If you tried to learn about differential forms from *Gravitation* by Misner, Thorne & Wheeler, this is what the whole "bongs of a bell" discussion is all about.

Coordinate Invariance and Tensors

A vector is a geometric object. In preschool, we're told that vectors are quantities that have a magnitude and direction. How we represent the magnitude and direction depends on coordinates. But vectors, being geometric objects, don't care about the coordinates. Coordinates are just labels or "names" that we give to points in space.

This means that coordinate transformations cannot affect things like directional derivatives and inner products. In other words, if I evaluate the following scalar quantities at some point p, the numerical values should be unaffected by coordinate transformations:

$$v(f) = v \cdot \nabla f \tag{36}$$

$$\langle u, v \rangle = g_{ij}\, u^i\, v^i \tag{37}$$

Dual vectors are also geometric objects. If v is a vector and w is a dual vector, then $w(v)$, being a scalar, can't be affected by a coordinate transformation either:

$$w(v) = w_i\, v^i \tag{38}$$

Now suppose I perform a coordinate transformation on v^i. To do this, I write:*

$$v^\alpha = \frac{\partial y^\alpha}{\partial x^i}\, v^i \tag{39}$$

Where $y^\alpha = y^\alpha(x^i)$ are my new coordinates, and x^i are my old coordinates. In order for $w(v)$ to remain unchanged by the coordinate transformation, it must transform in the ***opposite*** way:

$$w_\alpha = \frac{\partial x^i}{\partial y^\alpha}\, w_i \tag{40}$$

Where $x^i = x^i(y^\alpha)$.** Plugging the transformed w and transformed v into $w(v)$:

$$w(v) = w_\alpha\, v^\alpha = \frac{\partial x^j}{\partial y^\alpha}\, w_j\, \frac{\partial y^\alpha}{\partial x^i}\, v^i = \frac{\partial x^j}{\partial y^\alpha}\, \frac{\partial y^\alpha}{\partial x^i}\, w_j\, v^i \tag{41}$$

I'll apply the chain rule:

$$\frac{\partial x^j}{\partial y^\alpha}\, \frac{\partial y^\alpha}{\partial x^i} = \frac{\partial x^j}{\partial x^i} = \delta_i^j \tag{42}$$

The last step comes from the fact that coordinates are independent of each other (otherwise you'll have constraints). Also, recall the definition of the Kronecker delta (Equation 16).

Equation (38) becomes:

$$w(v) = \delta_i^j\, w_j\, v^i = w_i\, v^i \tag{43}$$

What I've just shown is that if w_i transforms in a way that's "opposite" to how v^i transforms under a coordinate change, then the transformations will cancel each other out, so that the scalar $w(v)$ will be unchanged by coordinate transformations.

Note that this also places a restriction on how the metric g_{ij} transforms. One can show that both indices of g_{ij} must be transformed like the components of w. This means that:

$$g_{\alpha\beta} = \frac{\partial x^i}{\partial y^\alpha} \frac{\partial x^j}{\partial y^\beta} g_{ij} \tag{44}$$

It turns out that the indices of the inverse metric transform like those of vectors:

$$g^{\alpha\beta} = \frac{\partial y^\alpha}{\partial x^i} \frac{\partial y^\beta}{\partial x^j} g^{ij} \tag{45}$$

This leads us to a defining property of *tensors*:

Raised indices of tensors transform like vectors under a coordinate change
Lowered indices of tensors transform like dual vectors under a coordinate change

Any multi-indexed object that transforms according to the rules above is a tensor. For example, a tensor with components $T^i{}_j$ will transform like:

$$T^\alpha{}_\beta = \frac{\partial y^\alpha}{\partial x^i} \frac{\partial x^j}{\partial y^\beta} T^i{}_j \tag{46}$$

The complete definition of tensor is (I'll explain the term "linear map" in just a bit):

A tensor is a linear map that "eats" vectors and/or dual vectors and "spits out" a scalar

For example, a tensor that eats one vector v and one dual vector w can be written:

$$T(w, v) = T^i{}_j w_i v^j \tag{47}$$

Or, for a 5-indexed tensor that eats vectors v, u, q, and dual vectors w, z, onw can write:

$$S(w, z, v, u, q) = S^{ij}{}_{klm} w_i z_j v^k u^l q^m \tag{48}$$

Where $S^{ij}{}_{klm}$ are the components the tensor S.

Linearity, by the way, is what allows us to write tensors as multi-indexed objects; one can only write equations (47) and (48) if T and S are linear maps. In particular, the condition that R be a linear map means R must satisfy the following properties:

$$R(..., \alpha u + \beta v, ...) = \alpha R(..., u, ...) + \beta R(..., v, ...) \tag{49}$$

$$R(..., \sigma w + \tau z, ...) = \sigma R(..., w, ...) + \tau R(..., z, ...) \tag{50}$$

for vectors u, v, dual vectors w, z, and real numbers α, β, σ, τ. Thus, if the mathematical construction R satisfies equations (49) & (50), we call it a *linear map*. Apologies for using the term before I define it.

Some terminology: the number of vectors and dual vectors that a tensor "eats" is called the *rank* of the tensor. The two-indexed tensor $T^i{}_j$ is a rank 2 tensor (to be more precise, it is a rank (1,1) tensor, since it eats one vector and one dual vector), and the 5-indexed tensor $S^{ij}{}_{klm}$ is a rank 5 tensor (or a rank (2,3) tensor).

Note that vectors and dual vectors are themselves tensors; a vector eats a single dual vector and spits out a scalar, and a dual vector eats a single

vector and spits out scalar:

$$v(w) = v^i w_i \tag{51}$$

$$w(v) = w_i v^i \tag{52}$$

By the way, vectors and dual vectors are rank 1 tensors (or rank (1,0) and rank (0,1) tensors, respectively). For completeness, scalar functions are rank (0,0) tensors.

Now, I'll get to the punchline of these notes. We like to use tensors because:

Tensor equations look the same in all coordinate systems.

For example, consider the following equation:

$$G_{ij} = \kappa \, T_{ij} \tag{53}$$

Where G_{ij} and T_{ij} are tensors and κ is some constant.

If I perform a coordinate change:

$$G_{\alpha\beta} = \frac{\partial x^i}{\partial y^\alpha} \frac{\partial x^j}{\partial y^\beta} G_{ij} \tag{54}$$

$$T_{\alpha\beta} = \frac{\partial x^i}{\partial y^\alpha} \frac{\partial x^j}{\partial y^\beta} T_{ij} \tag{55}$$

The equation becomes:

$$\frac{\partial x^i}{\partial y^\alpha} \frac{\partial x^j}{\partial y^\beta} G_{ij} = \kappa \, \frac{\partial x^i}{\partial y^\alpha} \frac{\partial x^j}{\partial y^\beta} T_{ij} \quad \longrightarrow \quad G_{\alpha\beta} = \kappa \, T_{\alpha\beta} \tag{56}$$

So in the new coordinates, the formula looks just like it does in the old ones!

----

*   When you encounter a quantity like $\frac{\partial y^\alpha}{\partial x^i}$, the index i is treated as a lowered index. So I can write my transformation as: $\Lambda^\alpha_i = \frac{\partial y^\alpha}{\partial x^i}$.

**   I should mention that all of this depends on the invertibility of the functions $y^\alpha(x^i)$, in particular the assumption that you can invert $y^\alpha(x^i)$ to get $x^i = x^i(y^\alpha)$.

Transforming the Metric / Unit Vectors as Non-Coordinate Basis Vectors

Since the metric is a tensor, one can find its expression in different coordinates by using the tensor transformation rule:

$$g_{\alpha\beta} = \frac{\partial x^i}{\partial y^\alpha} \frac{\partial x^j}{\partial y^\beta} g_{ij} \tag{57}$$

Let's transform the metric to spherical coordinates (r, θ, ϕ). Let $y^\alpha = (r, \theta, \phi)$, and:

$$x^1 = r \cos \phi \sin \theta$$
$$x^2 = r \sin \phi \sin \theta \tag{58}$$
$$x^3 = r \cos \theta$$

Using the formulas above, one can show that the metric tensor takes the following form:

$$g_{\alpha\beta} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & r^2 \sin^2\theta \end{pmatrix} \tag{59}$$

If one chooses a different set of coordinates, the coordinate basis vectors may not be unit vectors. The components of the metric can be defined in terms of the basis vectors e_i:

$$g_{ij} = \langle e_i, e_j \rangle \tag{60}$$

In the coordinate basis, $e_i = \frac{\partial}{\partial x^i}$, so that:

$$g_{ij} = \left\langle \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \right\rangle \tag{61}$$

Comparing (56) with (54), one finds that:

$$\left\langle \frac{\partial}{\partial r}, \frac{\partial}{\partial r} \right\rangle = g_{rr} = 1$$
$$\left\langle \frac{\partial}{\partial \theta}, \frac{\partial}{\partial \theta} \right\rangle = g_{\theta\theta} = r^2 \tag{62}$$
$$\left\langle \frac{\partial}{\partial \phi}, \frac{\partial}{\partial \phi} \right\rangle = g_{\phi\phi} = r^2 \sin^2\theta$$

With all other inner products between the coordinate basis vectors vanishing. This implies that the basis vectors $\left\{ \frac{\partial}{\partial r}, \frac{\partial}{\partial \theta}, \frac{\partial}{\partial \phi} \right\}$ are orthogonal, and that only $\frac{\partial}{\partial r}$ is a unit vector. This illustrates an important point:

Coordinate basis vectors $\frac{\partial}{\partial x^i}$ are <u>not</u> unit vectors, and unit vectors may not be coordinate basis vectors!

Once the norm of the coordinate basis vector is known, it's easy to convert between coordinate basis vectors and unit vectors. For spherical coordinates, you simply divide by the square root of the norm to get the unit vector you're familiar with in vector calculus:

$$\hat{\mathbf{r}} = \frac{1}{\sqrt{g_{rr}}} \frac{\partial}{\partial r} = \frac{\partial}{\partial r}$$
$$\hat{\boldsymbol{\theta}} = \frac{1}{\sqrt{g_{\theta\theta}}} \frac{\partial}{\partial \theta} = \frac{1}{r} \frac{\partial}{\partial \theta} \tag{63}$$
$$\hat{\boldsymbol{\phi}} = \frac{1}{\sqrt{g_{\phi\phi}}} \frac{\partial}{\partial \phi} = \frac{1}{r \sin\theta} \frac{\partial}{\partial \phi}$$

The above relations allow you to convert vectors and tensors from coordinate basis vectors to unit vectors in spherical coordinates. For an excercise, try to derive the same relations for cylindrical coordinates.

The Derivatives of Tensors

Many physical problems involve solving Partial Differential Equations (PDEs). Often, one finds that the difficulty of solving a PDE depends on the coordinate system used to express it; it may be much easier, for instance, to solve boundary value problem for a PDE in spherical coordinates than cylindrical coordinates.

By definition, PDEs contain partial derivatives; unfortunately, partial derivatives of tensors don't always transform like tensors, so we'll have to find a way to take derivatives so that the resulting quantity transforms like a tensor.

However, in the simplest case, the partial derivative does transform like a tensor. Let's start with the gradient of some function $f(x)$ (which is a rank (0,0) tensor--recall the discussion after equation 54):

$$(\nabla f(x^j))_k = \frac{\partial}{\partial x^k} f(x^j) \tag{64}$$

It turns out that the gradient transforms as a dual vector (or a rank (0,1) tensor); to see this, simply use the chain rule:

$$(\nabla f(x^j))_\alpha = \frac{\partial x^k}{\partial y^\alpha} \frac{\partial}{\partial x^k} f(x^j) \tag{65}$$

Since the gradient of a function transforms as a dual vector, so gradients are tensorial quantities.

The gradient of a vector, on the other hand, does not transform as a tensor. Consider the quantity:

$$A^i_j := \frac{\partial}{\partial x^j} v^i \tag{66}$$

If A^i_k were a tensor, it should transform like the following:

$$A^\alpha_\beta \overset{w}{=} \frac{\partial y^\alpha}{\partial x^i} \frac{\partial x^j}{\partial y^\beta} A^i_j \tag{67}$$

The $\overset{w}{=}$ indicates that we want the equality to hold. However, this doesn't work if you transform the vector before taking the derivative:

$$A^\alpha_\beta = \frac{\partial x^j}{\partial y^\beta} \frac{\partial}{\partial x^j} \left(\frac{\partial y^\alpha}{\partial x^i} v^i \right) = \frac{\partial y^\alpha}{\partial x^i} \frac{\partial x^j}{\partial y^\beta} A^i_j + \underline{\frac{\partial x^j}{\partial y^\beta} \frac{\partial}{\partial x^j} \left(\frac{\partial y^\alpha}{\partial x^i} \right) v^i} \tag{68}$$

The extra term (which I've underlined) prevents A^i_j from being a tensor.

However, it almost works. If it weren't for the underlined term, equation (68) would be satisfied.

Perhaps we should modify the definition of the derivative to deal with the second piece. The simplest modification that one can make to the partial derivative is to add a correction term that is linear in v^i:

$$\nabla_j v^i = \frac{\partial}{\partial x^j} v^i + \Gamma^i_{aj} v^a \tag{69}$$

If the coefficients Γ^i_{ab} are not tensors, but transform in the following way:

$$\Gamma^\gamma_{\alpha\beta} = \frac{\partial y^\gamma}{\partial x^i} \frac{\partial x^a}{\partial y^\alpha} \frac{\partial x^b}{\partial y^\beta} \Gamma^i_{ab} - \frac{\partial x^a}{\partial y^\alpha} \frac{\partial x^b}{\partial y^\beta} \frac{\partial}{\partial x^a} \left(\frac{\partial y^\gamma}{\partial x^b} \right) \tag{70}$$

Then after some work, you can show that the new derivative transforms as:

$$\nabla_\beta v^\alpha = \frac{\partial x^j}{\partial y^\beta} \nabla_j \left(\frac{\partial y^\alpha}{\partial x^i} v^i \right) = \frac{\partial y^\alpha}{\partial x^i} \frac{\partial x^j}{\partial y^\beta} \nabla_j v^i \tag{71}$$

The quantity $\nabla_j v^i$ transforms like a tensor! The derivative operator ∇_j is called a _covariant derivative_.

This can be extended to rank 2 tensors; the gradient of the tensor B^{ij} is:

$$\nabla_k B^{ij} = \frac{\partial}{\partial x^k} B^{ij} + \Gamma^i_{ak} B^{aj} + \Gamma^j_{ak} B^{ia} \tag{72}$$

Since B^{ij} has two indices, one needs two terms to "correct" the transformation properties. At this point, you might notice a pattern. Let's list the coordinate invariant gradients we have so far:

$$\nabla_k f = \frac{\partial f}{\partial x^k} \tag{73}$$

$$\nabla_k v^i = \frac{\partial}{\partial x^k} v^i + \Gamma^i_{ak} v^a \tag{74}$$

$$\nabla_k B^{ij} = \frac{\partial}{\partial x^k} B^{ij} + \Gamma^i_{ak} B^{aj} + \Gamma^j_{ak} B^{ia} \tag{75}$$

The number of correction terms we need depends on the rank (the number of indices) of the tensor.

The covariant derivative operator ∇_j has the following properties:

1. Additivity:

   For tensors A^{ij} and B^{mn},

   $$\nabla_k(A^{ij} + B^{ij}) = \nabla_k(A^{ij}) + \nabla_k(B^{ij}) \tag{76}$$

2. Product (Leibniz) Rule:

   For tensors A^{ij} and C^{lmn},

   $$\nabla_k\left(A^{ij}\,C^{lmn}\right) = C^{lmn}\,\nabla_k(A^{ij}) + A^{ij}\,\nabla_k\left(C^{lmn}\right) \tag{77}$$

These properties can be extended to tensors of any rank (including vectors and scalar functions) and tensors with lowered indices.

Note that I haven't written down the gradient of objects with lowered indices yet.

For a dual vector, the covariant derivative is:

$$\nabla_j w_i = \frac{\partial}{\partial x^j} w_i - \Gamma^a{}_{ij}\,w_a \tag{78}$$

The negative sign in the last term is not a mistake. This comes from the fact that $w(v) = w_i\,v^i$ is a scalar, and the fact that ∇_k satisfies the product rule. When taking the derivative, one obtains:

$$\nabla_k(w_i\,v^i) = \frac{\partial}{\partial x^k}(w_i\,v^i) = v^i\,\frac{\partial w_i}{\partial x^k} + w_i\,\frac{\partial v^i}{\partial x^k} \tag{79}$$

The product rule, on the other hand, implies that:

$$\nabla_k(w_i\,v^i) = v^i\,\nabla_k(w_i) + w_i\,\nabla_k(v^i) \tag{80}$$

The negative sign in equation (78) is needed in order for equations (79) and (80) to be consistent with each other; the terms containing $\Gamma^i{}_{ab}$ must cancel each other out.

Thus, for a tensor with raised and lowered indices, the covariant derivative becomes:

$$\nabla_k T^i{}_j = \frac{\partial}{\partial x^k} T^i{}_j + \Gamma^i{}_{ak}\,T^a{}_j - \Gamma^a{}_{jk}\,T^i{}_a \tag{81}$$

At this point, the coefficients $\Gamma^i{}_{ab}$ are undetermined; all we know are its transformation properties. In fact, the transformation properties do not uniquely determine the coefficients. However, there are two conditions that we can impose on ∇_k to uniquely determine the coefficients:

1. The Torsion-Free condition:*

   For a scalar function f,

   $$\nabla_i \nabla_j f = \nabla_j \nabla_i f \quad \Rightarrow \quad \Gamma^a{}_{ij} = \Gamma^a{}_{ji} \tag{82}$$

2. Metric compatibility:

   $$\nabla_k g_{ij} = 0 \tag{83}$$

If these two conditions are satisfied, the coefficients $\Gamma^i{}_{ab}$ become:

$$\boxed{\Gamma^i{}_{ab} = \frac{1}{2} g^{ic}\left(\frac{\partial g_{ac}}{\partial x^b} + \frac{\partial g_{cb}}{\partial x^a} - \frac{\partial g_{ab}}{\partial x^c}\right)} \tag{84}$$

These are called *Christoffel symbols*. In Cartesian coordinates in Euclidean space, $\Gamma^i{}_{ab} = 0$.

Some terminology: Mathematicians call covariant derivatives *connections*, and $\Gamma^i{}_{ab}$ are *connection coefficients*. Their terminology is more general, and the covariant derivative that I have defined is called the *Levi-Civita connection* (not to be confused with the Levi-Civita symbol, which I will discuss later).*

The punchline of this section is that the following gradients transform as tensors:

$$\boxed{\nabla_j v^i = \frac{\partial}{\partial x^j} v^i + \Gamma^i{}_{aj}\,v^a} \tag{85}$$

$$\boxed{\nabla_j w_i = \frac{\partial}{\partial x^j} w_i - \Gamma^a{}_{ij}\, w_a} \tag{86}$$

$$\boxed{\nabla_k T^i{}_j = \frac{\partial}{\partial x^k} T^i{}_j + \Gamma^i{}_{ak}\, T^a{}_j - \Gamma^a{}_{jk}\, T^i{}_a} \tag{87}$$

So, if you have a gradient in Cartesian coordinates on Euclidean space, all you have to do to convert it into a tensor expression is to replace your partial derivatives with these fancy new covariant derivatives:**

$$\frac{\partial}{\partial x^k} \to \nabla_k \tag{88}$$

Once you do that, you'll have a gradient that's valid in any coordinate system!

In fact, you can write all derivative operators in terms of tensor gradients, so this trick allows you to rewrite PDEs in any coordinate system! Keep in mind that this replacement works only if the PDE is originally written in Cartesian coordinates on Euclidean space; this trick is valid because $\Gamma^i{}_{ab}$ vanishes in Cartesian coordinates on Euclidean space.

This material can be found in most General Relativity textbooks, but they often focus on non-Euclidean geometries, parallel transport in non-Euclidean geometries, and curvature. If you're interested in these things, I recommend four General Relativity textbooks:

- Bernard F. Schutz, *A First Course in General Relativity*, Cambridge University Press 2009
- Sean M. Carroll, *Spacetime and Geometry: An Introduction to General Relativity*, Addison-Wesley 2003
- Robert M. Wald, *General Relativity*, University Of Chicago Press 1984
- Charles W. Misner, Kip S. Thorne, & John A. Wheeler, *Gravitation*, W.H. Freeman 1973

* One can define a class of geometries that include Euclidean and non-Euclidean geometries by specifying a nondegenerate metric (the metric produces a unique dual vector for each vector) and a torsion-free, metric-compatible covariant derivative. These geometries are called semi-Riemannian geometries, which are used in Special and General Relativity. If the metric is positive definite, or that all vectors have positive norms, one has a Riemannian geometry, which can be used to describe the geometry of curved surfaces.

** This is called the "comma to semicolon rule" in General Relativity. The commas and semicolons refer to the notation for partial and covariant derivatives often used in General relativity, where $w_{i,j} = \partial_j w_i$ and $w_{i;j} = \nabla_j w_i$. However, it should be mentioned that subtleties arise in applying this rule in non-Euclidean geometries; if you're interested, see Ch. 4 of *General Relativity* by Robert Wald, and Box 16.1 in *Gravitation* by Misner, Thorne and Wheeler.

Divergences and Laplacians

PDEs are often expressed in terms of Divergences and Laplacians. At this point, we have all the tools necessary to write these down.

In vector calculus, the divergence of a vector function $\vec{V}$ is written in the following way:

$$\vec{\nabla} \cdot \vec{V} \tag{89}$$

The most natural coordinate-invariant expression that one can construct is the following:

$$\boxed{\vec{\nabla} \cdot \vec{V} = \nabla_i V^i} \tag{90}$$

However, if we take equation (89) literally, $\vec{\nabla}$ is a vector operator, which has the components:

$$(\vec{\nabla})^i = g^{ij}\, \nabla_j \tag{91}$$

What equation (89) is really saying is the following (recall that the dot product becomes the metric product):

$$\vec{\nabla} \cdot \vec{V} = g_{ab}(\vec{\nabla})^a \, V^b = g_{ab} \, g^{aj} \, \nabla_j V^b = \delta_b^j \, \nabla_j V^b = \nabla_j V^j \tag{92}$$

The last three equalities demonstrate that equation (89) is equivalent to equation (90).

The Laplacian of a function is defined as:

$$\nabla^2 f := \vec{\nabla} \cdot \vec{\nabla} f \tag{93}$$

Using equation (91), this becomes:

$$\vec{\nabla} \cdot \vec{\nabla} f = g_{ab}(\vec{\nabla})^a \, (\vec{\nabla})^b \, f = g_{ab} \, g^{ai} \, \nabla_i \left(g^{bj} \, \nabla_j f \right) \tag{94}$$

Recall that the covariant derivative ∇_i satisfies metric compatibility ($\nabla_i g_{ab} = \nabla_i g^{ab} = 0$) and the product rule, so that:

$$\vec{\nabla} \cdot \vec{\nabla} f = g_{ab} \, g^{ai} \, g^{bj} \, \nabla_i \nabla_j f = \delta_b^i \, g^{bj} \, \nabla_i \nabla_j f = g^{ij} \, \nabla_i \nabla_j f \tag{95}$$

The Laplacian is then:

$$\boxed{\nabla^2 f = g^{ij} \, \nabla_i \nabla_j f} \tag{96}$$

The Levi-Civita Tensor: Cross Products, Curls and Volume Integrals

In vector analysis and vector calculus, one often comes across expressions involving cross products and curls:

$$\vec{A} \times \vec{B} \tag{97}$$

$$\vec{\nabla} \times \vec{V} \tag{98}$$

In Cartesian coordinates, one can use the permutation symbol* ϵ_{ijk} to write the components of these expressions. The permutation symbol is defined as:

$$\epsilon_{ijk} := \begin{cases} +1 & \text{if } (i,\, j,\, k) \text{ is } a \text{ permutation of } (1,\, 2,\, 3) \\ -1 & \text{if } (i,\, j,\, k) \text{ is } a \text{ permutation of } (1,\, 3,\, 2) \\ 0 & \text{if } i = j, \text{ or } j = k, \text{ or } i = k \end{cases} \tag{99}$$

Explicitly, this means that the components of ϵ_{ijk} are:

$$\epsilon_{123} = \epsilon_{312} = \epsilon_{231} = 1$$
$$\epsilon_{132} = \epsilon_{321} = \epsilon_{213} = -1 \tag{100}$$

with all others vanishing.

As mentioned before, one can use ϵ_{ijk} to write the components of $\vec{A} \times \vec{B}$ and $\vec{\nabla} \times \vec{V}$ in Cartesian coordinates:**

$$\left(\vec{A} \times \vec{B} \right)^i = g^{ia} \, \epsilon_{ajk} \, A^j \, B^k \tag{101}$$

$$\left(\vec{\nabla} \times \vec{V} \right)^i = g^{ia} \, \epsilon_{ajk}(\vec{\nabla})^j \, V^k = g^{ia} \, g^{jb} \, \epsilon_{ajk} \, \nabla_b V^k \tag{102}$$

The inverse metric g^{ia} is used to raise the free index (the one that is not summed over); this is because the expression $\epsilon_{ajk} \, A^j \, B^k$, having a lowered index a, is a dual vector.

Equations (101) and (102) are *not* tensor equations; they are only valid in Cartesian coordinates!

This is because ϵ_{ijk} is not a tensor. Since ϵ_{ijk} is defined by equation (99), its components don't change under a coordinate transformation.

To fix this, consider the following expression in Cartesian coordinates:

$$\vec{A} \cdot \left(\vec{B} \times \vec{C} \right) = \epsilon_{ijk} \, A^i \, B^j \, C^k \tag{103}$$

This is not a coordinate-invariant expression. Under a coordinate change, the right-hand side of equation (102) becomes:

$$\epsilon_{ijk} \, \frac{\partial x^i}{\partial y^\alpha} \, \frac{\partial x^j}{\partial y^\beta} \, \frac{\partial x^k}{\partial y^\gamma} A^\alpha \, B^\beta \, C^\gamma \neq \epsilon_{\alpha\beta\gamma} \, A^\alpha \, B^\beta \, C^\gamma \tag{104}$$

which is not equal to the same expression in the new coordinates. However, the quantity that appears on the left-hand side of the equation above is related to an expression for the Jacobian determinant of coordinate transformation matrix $\frac{\partial y^\alpha}{\partial x^i}$:

$$\epsilon_{\alpha\beta\gamma} \det\left(\frac{\partial x^i}{\partial y^\alpha}\right) = \epsilon_{ijk} \frac{\partial x^i}{\partial y^\alpha} \frac{\partial x^j}{\partial y^\beta} \frac{\partial x^k}{\partial y^\gamma} \tag{105}$$

Furthermore $\det\left(\frac{\partial x^i}{\partial y^\alpha}\right)$ is related to the determinant of the metric tensor, which is defined as:

$$\det(g_{ij}) = \epsilon^{ijk} g_{i1} g_{j2} g_{k3} \tag{106}$$

Where ϵ^{ijk} is defined the same way as ϵ_{ijk}; the only difference is that it now "eats" dual vectors instead of vectors.†

The determinant of a coordinate transformed metric is:

$$\det(g_{\alpha\beta}) = \det\left(\frac{\partial x^i}{\partial y^\alpha} \frac{\partial x^j}{\partial y^\beta} g_{ij}\right) = \left(\det\left(\frac{\partial x^i}{\partial y^\alpha}\right)\right)^2 \det(g_{ij}) \tag{107}$$

Where I've used the property of determinants for matrix multiplication in the last equality.‡ If g_{ij} the metric for Cartesian coordinates, then $\det(g_{ij}) = 1$, so the coordinate-transformed metric is the Jacobian determinant:

$$\boxed{|g| := \det(g_{\alpha\beta}) = \left(\det\left(\frac{\partial x^i}{\partial y^\alpha}\right)\right)^2} \tag{108}$$

One can use plug this back into equation (105) to obtain:

$$\sqrt{|g|} \ \epsilon_{\alpha\beta\gamma} = \epsilon_{ijk} \frac{\partial x^i}{\partial y^\alpha} \frac{\partial x^j}{\partial y^\beta} \frac{\partial x^k}{\partial y^\gamma} \tag{109}$$

Equation (104) becomes:

$$\epsilon_{ijk} \frac{\partial x^i}{\partial y^\alpha} \frac{\partial x^j}{\partial y^\beta} \frac{\partial x^k}{\partial y^\gamma} A^\alpha B^\beta C^\gamma = \sqrt{|g|} \ \epsilon_{\alpha\beta\gamma} A^\alpha B^\beta C^\gamma \tag{110}$$

Or, more suggestively:

$$\epsilon_{ijk} A^i B^j C^k = \sqrt{|g|} \ \epsilon_{\alpha\beta\gamma} A^\alpha B^\beta C^\gamma \tag{111}$$

Where the left hand side is written in Cartesian coordinates.

It turns out that the Levi-Civita symbol can be turned into a tensor-like quantity if one multiplies it by $\sqrt{|g|}$, where $|g|$ is the determinant of the metric tensor. One can define the Levi-Civita tensor $\tilde{\epsilon}_{ijk}$ as follows:§

$$\tilde{\epsilon}_{ijk} := \sqrt{|g|} \ \epsilon_{ijk} \tag{112}$$

Which, under a coordinate transformation satisfies:

$$\tilde{\epsilon}_{ijk} A^i B^j C^k = \tilde{\epsilon}_{\alpha\beta\gamma} A^\alpha B^\beta C^\gamma \tag{113}$$

Explicitly, the formula above can be written:

$$\sqrt{|g|} \ \epsilon_{ijk} A^i B^j C^k = \sqrt{|g'|} \ \epsilon_{\alpha\beta\gamma} A^\alpha B^\beta C^\gamma \tag{114}$$

Where the prime denotes the coordinate-transformed metric.

Finally, one can write the cross-product and the curl in a coordinate-invariant way:

$$\boxed{\left(\vec{A} \times \vec{B}\right)^i = g^{ia} \tilde{\epsilon}_{ajk} A^j B^k} \tag{115}$$

$$\boxed{\left(\vec{\nabla} \times \vec{V}\right)^i = g^{ia} g^{jb} \tilde{\epsilon}_{ajk} \nabla_b V^k} \tag{116}$$

Where:

$$\boxed{\tilde{\epsilon}_{ijk} := \sqrt{|g|} \ \epsilon_{ijk}} \tag{117}$$

One neat byproduct of the preceding discussion is that we can now write volume integrals in a coordinate invariant way. Recall that the volume element $d^3 x = dx^1 \, dx^2 \, dx^3$ picks up a factor of the Jacobian determinant under a coordinate transformation. Since $\sqrt{|g|}$ is equal to the Jacobian determinant, one can write the following:

$$d^3 x = \sqrt{|g|} \, dx^1 \, dx^2 \, dx^3 \tag{118}$$

This expression is coordinate-invariant--it applies to all coordinate systems. Try computing this for spherical coordinates, and verify that it produces the expected volume element.

For completeness, one can write the volume integral for the function $f(x^i)$:

$$\int_V f(x^i) \, d^3 x = \int_V f(x^i) \sqrt{|g|} \, dx^1 \, dx^2 \, dx^3 \tag{119}$$

----

*This is also referred to as the Levi-Civita symbol, not to be confused with the Levi-Civita connection, which is defined by equation (84).

**You should convince yourself that this is equivalent to the ordinary definition of the cross product and the curl.

†Depending on the expressions involved, one may find it necessary to introduce mixed-index permutation symbols (like $\epsilon^i{}_{jk}$, for example), which are also defined in the same way.

‡If you haven't already, you should convince yourself that for two matrices A^i_a and B^b_l, the expression $A^i_a B^b_l$ is just matrix multiplication. You should also convince yourself that g_{ij} can be effectively treated as a matrix in this case.

§The Levi-Civita tensor $\tilde{\epsilon}_{ijk}$ is sometimes called a pseudotensor because it is not an actual tensor in the strictest sense. This is because unlike a tensor, it picks up a negative sign when coordinates are reversed. This also applies to any "vector" (pseudovector) formed out of cross-products.

Further Reading

As I mentioned in the introductory remarks, these notes are meant to be self-contained, and I have done my best to build the subject up from first principles. However, I understand that my notes may be unsatisfactory to some; they may not be pedagogical enough, not rigorous enough, or too informal to be taken seriously.* To address these potential shortcomings, I've attempted to compile a list of references that may better explain the material.

Though I have included a few references throughout these notes, I have avoided them as much as possible, since these notes are intended for a broad audience and I did not want to inundate the reader with a bunch of General Relativity texts (although I fear I already have). Still, any list of references I provide will consist of GR textbooks and GR-minded mathematical physics books, since I exclusively consulted such books when learning tensor calculus.

- D. F. Lawden, *Introduction to Tensor Calculus, Relativity and Cosmology*, Dover 2003
- Bernard F. Schutz, *A First Course in General Relativity*, Cambridge University Press 2009
- Sean M. Carroll, *Spacetime and Geometry: An Introduction to General Relativity*, Addison-Wesley 2003
- Robert M. Wald, *General Relativity*, University Of Chicago Press 1984
- Charles W. Misner, Kip S. Thorne, & John A. Wheeler, *Gravitation*, W.H. Freeman 1973
- Ignazio Ciufolini & John A. Wheeler, *Gravitation and Inertia*, Princeton University Press 1995
- Bernard F. Schutz, *Geometrical methods of mathematical physics*, Cambridge University Press 1980
- John C. Baez & Javier P. Muniain, *Gauge Fields, Knots, and Gravity*, World Scientific 1994
- Theodore Frankel, *The Geometry of Physics: An Introduction*, 3rd Ed., Cambridge University Press 2011

----

*   These notes grew out of an appendix to a homework assignment, and were originally full of frivolous remarks (they were meant to be humorous, but I have a strange sense of humor when working on homework assignments at 3 am). I have done my best to remove these remarks while still maintaining an informal tone; this may explain the varying tone in these notes. Let me know if you find any inappropriately frivolous comments.

Some Exercises

I conclude these notes by leaving you with a few simple exercises:

1. Rewrite the theorems of vector calculus (Gradient theorem, Green's Theorem, Stokes Theorem, Divergence Theorem) using tensorial quantities. For a challenge, express area elements in terms of coordinates on the boundary.

2. Rewrite the heat and wave equations (or alternately, Laplace's equation) using tensorial quantities.

3. Rewrite Maxwell's Equations in both differential and integral form using tensorial quantities.

4. Rewrite the Navier-Stokes Equations in both differential and integral form using tensorial quantities.

5. Rewrite all of the above in cylindrical coordinates and spherical coordinates (in a coordinate basis)

6. Convert the equations in problem **5.** to an orthonormal (unit vector) basis.

7. Show, using equation (85), that the commutator of covariant derivatives acting on a vector is proportional to the vector itself. In particular, show that, for a vector v^b, $(\nabla_i \nabla_j - \nabla_j \nabla_i) v^a = R^a{}_{bij} v^b$, where $R^a{}_{bij}$ depends only on connection coefficients $\Gamma^i{}_{aj}$ and their derivatives. Write down your expression for $R^a{}_{bij}$. $R^a{}_{bij}$, by the way, is called the Riemann curvature tensor; it vanishes if the geometry is Euclidean, but is nonzero for non-Euclidean geometries.