The Poor Man's Introduction to Tensors

Justin C. Feng

Preface

In the spring semester of 2013, I took a graduate fluid mechanics class taught by Philip J. Morrison. On the first homework assignment, I found that it was easier to solve some of the problems using tensors in the coordinate basis, but I was uncomfortable with it because I wasn't sure if I was allowed to use a formalism that had not been taught in class. Eventually, I decided that if I wrote an appendix on tensors, I would be allowed to use the formalism. Sometime in 2014, I converted[1] the appendix into a set of introductory notes on tensors and posted it on my academic website, figuring that they might be useful for someone interested in learning the topic.

Over the past three years, I had become increasingly dissatisfied with the notes. The biggest issue with these notes was that they were typeset in Mathematica—during my early graduate career, I typeset my homework assignments in Mathematica, and it was easiest to copy and paste the appendix into another Mathematica notebook (Mathematica does have a feature to convert notebooks to TeX, but much of the formatting is lost in the process). I had also become extremely dissatisfied with the *content*—the notes contained many formulas that weren't sufficiently justified, and there were many examples of sloppy reasoning throughout. At one point, I had even considered removing the notes from my website, but after receiving some positive feedback from students who benefited from the notes, I decided to leave it up for the time being, vowing to someday update and rewrite these notes in TeX. My teaching duties and dissertation work have delayed my progress—it has been more than two years since my last update to these notes—but I have finally found the time to properly update these notes.

These notes have been rewritten; this means that I have re-typed the notes from beginning to end, revising and adding to the notes as I went along. What I *did not* do was copy and paste text from my old notes and edit the result—every single word in this document has been re-typed. I do this to ensure that all of the shortcomings of the earlier versions (those that I am aware of, at least) are addressed. All sections have been rewritten and expanded. The overall organization of these notes is the same as it was in the original, except for the addition of a section on index gymnastics and a section at the end which covers surface integrals, the divergence theorem, and Stokes' theorem. Exercises have been added to the end of each section to enhance the reader's understanding of the material. An index has been added as well. Though I have added a lot of content, I have attempted to maintain the succinctness of the original notes, and resisted the temptation to include a full-blown treatment of differential forms and Riemannian geometry—I'll save those for a separate set of notes.

The title, *The Poor Man's Introduction to Tensors*, is a reference to *Gravitation* by Misner, Thorne and Wheeler, which characterizes simplified approaches to a problem as "the poor man's way to do X." Originally, these notes were intended to be a short, informal primer on tensors, and were by no means a substitute for a more formal and complete treatment of the subject. I fear that in my effort to overcome the shortcomings of the previous version, these notes have become too formal and detailed to justify the label "Poor Man's" in the original sense of a simplified, informal treatment of the topic. However, I have tried to rewrite these notes in a manner that is accessible to anyone with a basic training in linear algebra and vector analysis, and I promise to always make these notes freely available on the web—these notes are in this sense *The Poor Man's Introduction to Tensors*.

If you find any errors or have any suggestions for these notes, feel free to contact me at:
`jcfeng@physics.utexas.edu`

Have Fun!

Justin C. Feng
Austin, Texas
December 2017

[1] This involved adding some content, in particular the material on the Levi-Civita tensor and integration, and the removal of frivolous content—the original appendix was full of jokes and humorous remarks, which I often placed in my homework assignments to entertain myself (and also the grader).

The Poor Man's Introduction to Tensors

Justin C. Feng[1]

[1] *Physics Department, The University of Texas at Austin*

(Dated: December 2017)

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 useful because they provide a formalism that is *both* coordinate independent and transparent for performing calculations. In particular, tensors facilitate the transformation of partial differential equations and the formulas of vector calculus to their corresponding forms in curvilinear coordinates. In these notes, I provide an introduction to tensors in Euclidean space for those who are familiar with the basics of linear algebra and vector calculus.

CONTENTS

I. INTRODUCTION

These notes were written for a broad audience—I wrote these notes to be accessible to anyone with a basic knowledge of linear algebra and vector calculus.[1] I have done my best to build up the subject from first principles; the goal of these notes is not to simply teach you the "mechanics" of the formalism[2], but to provide you with a fundamental understanding of what tensors are. Since these notes are intended for a broad audience, I will avoid discussion of General Relativity and non-Euclidean geometry, and focus instead on developing the formalism for ordinary three-dimensional Euclidean space. In addition to providing a fundamental understanding of what tensors are, these notes are intended to provide you with the tools to effortlessly write down explicit expressions for Partial Differential Equations and integrals in a general curvilinear coordinate system.[3]

---

[1] For those who are unfamiliar with these topics and those who need a refresher, I can suggest a few books (and a short symmery). *Linear Algebra: Step by Step* by K. Singh covers all Linear algebra concepts that I assume of the reader. There is also a short 4-page summary in [25], which summarizes the topics covered in the recent (crudely-titled) book *No Bullshit Guide to Linear Algebra* by Ivan Savov. The book *Div, Grad, Curl, and All That* by H. M. Schey [26] provides an excellent informal introduction to vector calculus. I learned the basics from the book *Mathematical Methods in the Physical Sciences* by Mary Boas [4].

[2] In these notes, the word *formalism* is defined as a collection of rules and techniques for manipulating symbols. A good formalism should provide a systematic way of writing down a complicated mathematical operation in a much simpler form. One of my goals in writing these notes is to show you how the formalism of tensors simplify coordinate transformations for PDEs and integrals.

[3] *Curvilinear coordinates* on Euclidean space are defined as coordinate systems in which the coordinate lines are curved.

II. TENSORS CONDENSED

When learning a new topic, I often find it helpful to identify the central ideas and principles 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 essence of 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 trivial fact, one may obtain the main result of tensor analysis: an equation written in tensor form is valid in any coordinate system.

In my attempt to summarize tensor analysis in a single paragraph, I have left out many important details. For instance, the definition for tensors in the first sentence is an incomplete one; in particular, it leaves out the fact that tensors are linear maps, as well as the fact that tensors also "eat" other objects called dual vectors. These details will be discussed in the remainder of these notes.

III. INDEX NOTATION (INDEX PLACEMENT IS IMPORTANT!)

If you are already familiar with indices, it may be tempting to skip this section. However, I emphasize some important points in this section—at the very least, make sure to take note of the boldfaced text.

Indices (the plural of index) provide a useful way to organize a large number of quantities, be they variables, functions, or abstract elements of sets. They are particularly useful when you have a large collection of equations that all have a similar form. Before I tell you what an index is, I'd like to provide a quick motivating example first to sell you on the formalism. Suppose you encounter a physical system which is described by 89 different variables. If you attempt to represent each variable with a single Latin or Greek letter, you will run out of letters before you could write down all the variables for the system!

An index is written as a superscript or a subscript that we attach to a symbol; for instance, the subscript letter i in q_i is an index for the symbol q, as is the superscript letter j in p^j is an index for the symbol p. Indices often represent positive integer values; as an example, for q_i, i can take on the values $i = 1$, $i = 2$, $i = 3$, and so on. In this way, I can represent all 89 variables by simply writing down q_i, with the understanding that i can have any integer value from 1 to 89. In particular, q_i provides a simple way to represent the full list of 89 variables:

$$q_1, q_2, q_3, q_4, q_5, q_6, q_7, q_8, q_9, q_{10}, q_{11}, q_{12}, q_{13}, q_{14}, q_{15}, q_{16}, q_{17}, q_{18}, q_{19}, q_{20}, q_{21}, q_{22}, q_{23}, q_{24}, q_{25}, q_{26}, q_{27}, q_{28}, q_{29}, q_{30}, q_{31}, q_{32}, q_{33}, q_{34}, q_{35}, q_{36}, q_{37}, q_{38}, q_{39}, q_{40}, q_{41}, q_{42}, q_{43}, q_{44}, q_{45}, q_{46}, q_{47}, q_{48}, q_{49}, q_{50}, q_{51}, q_{52}, q_{53}, q_{54}, q_{55}, q_{56}, q_{57}, q_{58}, q_{59}, q_{60}, q_{61}, q_{62}, q_{63}, q_{64}, q_{65}, q_{66}, q_{67}, q_{68}, q_{69}, q_{70}, q_{71}, q_{72}, q_{73}, q_{74}, q_{75}, q_{76}, q_{77}, q_{78}, q_{79}, q_{80}, q_{81}, q_{82}, q_{83}, q_{84}, q_{85}, q_{86}, q_{87}, q_{88}, q_{89}$$

This is a pain to write out by hand—it's much easier to just write q_i, with i representing integer values from 1 to 89.

I now consider a more concrete example. Many problems in physics and engineering are formulated in Cartesian coordinates on three-dimensional Euclidean space. For a three-dimensional Euclidean space, Cartesian coordinates refer to the three variables x, y, and z. Instead of using three different variables x, y, and z to describe the coordinates of a point, I can instead write x^i, where i can be any number in the list $(1, 2, 3)$. Explicitly, I write the following:

$$\begin{aligned} x^1 &= x \\ x^2 &= y \\ x^3 &= z \end{aligned} \tag{3.1}$$

A word of caution—**The superscripts 1, 2, and 3 are NOT exponents!** The superscripts 1, 2, and 3 are simply labels telling you which coordinate you are referring to. To be clear: the "2" in x^2 means that x^2 represents coordinate number 2; it DOES NOT mean $x \cdot x$!

You might be wondering why I choose to represent the coordinates x^i with a superscript, rather than a subscript (for instance, I could have written x_i instead). Though this is partly a matter of convention, the use of superscripts for coordinate indices is widely used. In any case, I feel that I must emphasize this convention:

> **In these notes, coordinate indices will always be superscripts**

This may seem to be overly pedantic, but I'm doing this because I want to emphasize and alert you to the fact that in tensor analysis, **INDEX PLACEMENT IS IMPORTANT!** In case it isn't clear, *Index placement* refers to whether the index is a superscript or a subscript. I'll take this opportunity to introduce you to some terminology: a superscript index, like the j in p^j, is called a *raised index*, and a subscript index, like the i in q_i, is called a *lowered index*.

Indices may also be used to describe a vector in three-dimensional Euclidean space. Typically, we write $\vec{v}$ to represent a vector. In three dimensions, we use three numbers to describe a vector, so that for a vector $\vec{v}$ in Euclidean space (assuming Cartesian coordinates), $\vec{v}$ represents the list of three numbers (v_x, v_y, v_z), with v_x being

the component of the vector in the x-direction, v_y being the component of the vector in the y-direction, and so on. In index notation, I may write $\vec{v}$ as v^i, so that:

$$\begin{aligned} v^1 &= v_x \\ v^2 &= v_y \\ v^3 &= v_z \end{aligned} \tag{3.2}$$

The formulas above (3.2) allow me to identify v^1 as the component of the vector in the x^1-direction, v^2 as the component of the vector in the x^2-direction, and so on. The expression v^i is therefore a compact way express the components of a vector.

Note that I also use raised indices (superscripts) for vectors. This is because introductory courses tend to characterize vectors as the difference between two points, which in index notation may be written as Δx^i. Since the index i in Δx^i is raised, vector components defined in this manner should have raised indices.

Index notation may be extended to vector formulas in a straightforward manner. Consider the following well-known formula:

$$\vec{F} = m\,\vec{a}, \tag{3.3}$$

where $\vec{F}$ and $\vec{a}$ are vectors, and m is some positive number. To convert this to index form, I replace the arrows with indices:

$$F^i = m\,a^i \tag{3.4}$$

Equation (3.4) is just a compact way of writing the following three equations:

$$\begin{aligned} F^1 &= m\,a^1 \\ F^2 &= m\,a^2 \\ F^3 &= m\,a^3 \end{aligned} \tag{3.5}$$

Index notation may also be used to describe square matrices. Recall that one may write down a square matrix $\mathbf{M}$ in three dimensions as the following table of numbers:

$$\mathbf{M} = \begin{pmatrix} M_{11} & M_{12} & M_{13} \\ M_{21} & M_{22} & M_{23} \\ M_{31} & M_{32} & M_{33} \end{pmatrix} \tag{3.6}$$

It is tempting to write the components of the matrix as M_{ij}, and this is often done. I will do this for now, but don't get too comfortable with it—I will change my conventions in just a bit. In index notation, I may write the Eigenvector formula $\mathbf{M}\,\vec{v} = \lambda\,\vec{v}$ as:

$$\sum_{j=1}^{3} M_{ij}\,v^j = \lambda\,v^i \tag{3.7}$$

For me, it is uncomfortable to deliberately write this formula down—the feeling is kind of like hearing the sound of fingernails scraping a chalkboard. This is due to a mismatch in the placement of indices. Notice that on the right hand side of (3.7), the index i is raised, but on the left hand side, the index i is lowered. When a matrix acts on a vector, the result must also be a vector, but according to my earlier convention, vectors must have *raised* indices. If the left hand side of (3.7) form the components of a vector, the index i must also be raised.

If I insist that vector indices must be raised, then the "proper" way to express the components of a matrix is $M^i{}_j$, so that the Eigenvector formula (3.7) becomes:

$$\sum_{j=1}^{3} M^i{}_j\,v^j = \lambda\,v^i \tag{3.8}$$

At first, this may seem rather awkward to write, since it suggests that the individual matrix elements be written as $M^1{}_1$, $M^1{}_2$, ... etc. However, I assure you that as you become more comfortable with tensor analysis, equation (3.8) will seem less awkward to write than (3.7).

While I'm on the topic of matrices, I'd like to introduce the *Kronecker delta* δ^i_j, which is just the components of the identity matrix. Specifically, δ^i_j is defined in the following way:

$$\begin{aligned} \delta^i_j &= \begin{cases} 1 & \text{if } \{i = j\} \\ 0 & \text{if } \{i \neq j\} \end{cases} \qquad \Rightarrow \\ \delta^i_j &= \begin{pmatrix} \delta^1_1 & \delta^1_2 & \delta^1_3 \\ \delta^2_1 & \delta^2_2 & \delta^2_3 \\ \delta^3_1 & \delta^3_2 & \delta^3_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \end{aligned} \tag{3.9}$$

The quantities δ^{ij} and δ_{ij} are similarly defined, and are also referred to as Kronecker deltas. To drive home the point that **INDEX PLACEMENT IS IMPORTANT**, it turns out that δ^i_j form the components of a tensor, but δ^{ij} and δ_{ij} do not—after I present the definition for a tensor, I will give you an exercise (exercise **IX.3**) where you show this.

I hope I am not frustrating you with my obsession about index placement. The whole point of this is to alert you to the importance of index placement in tensor analysis. I'll write it again: **INDEX PLACEMENT IS IMPORTANT!** My shouting is especially directed to those who have been exposed to Cartesian tensors[1]. Since I have emphasized this point repeatedly, it is appropriate for me to give you some idea of why index placement is important. The placement of the index is used

[1] Many treatments of tensor analysis begin by studying Cartesian tensors (tensors expressed exclusively in Cartesian coordinates), and when doing so, the distinction between raised and lowered indices is often ignored. One example is the treatment of tensors in [4], which I used as an undergraduate, and also [15]. I understand that the intent in taking this approach is to provide a gentler introduction to tensors, but I feel that this approach obscures the essence of tensors, and can lead to a great deal of confusion when moving on to curvilinear coordinates.

to distinguish two types of objects:[1] vectors v^i and *dual vectors* w_i. This distinction will be critically important in upcoming sections and we will not be able to proceed without it.

Exercise III.1

In your own hand, write down the following sentence three times on a sheet of paper:

INDEX PLACEMENT IS IMPORTANT!

Exercise III.2

Let $\mathbf{M}$, $\mathbf{P}$, and $\mathbf{Q}$ be $n \times n$ matrices, with the respective components $M^i{}_j$, $P^i{}_k$, and $Q^k{}_j$ (I choose the letters for the indices on the components to help you out). Rewrite the formula $\mathbf{M} = \mathbf{PQ}$ (which involves matrix multiplication) using index notation. Your result should contain a sum over the values of one index.

IV. EINSTEIN SUMMATION CONVENTION

I've been told on more than one occasion that Albert Einstein's greatest contribution to physics and mathematics[2] is his invention of the Einstein summation convention, which is the following rule:

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

For instance, given the matrix components $M^i{}_j$ and vector components v^i, Einstein summation convention states that when you see an expression of the form $M^i{}_j\, v^j$, there is a sum over the index j, since the letter j appears twice. More explicitly, Einstein summation convention states that the following expression:

$$M^i{}_j\, v^j \tag{4.1}$$

is equivalent to the explicit sum[3]:

$$\sum_j M^i{}_j\, v^j \tag{4.2}$$

I state this again: Einstein summation convention states that when I write down the expression $M^i{}_j\, v^j$, I should automatically assume that there is a sum over any index that appears twice (again, one must be raised and the other lowered).

Einstein invented this convention after noting that the sums which appear in calculations involving matrix and vector products always occur over pairs of indices. At first, Einstein summation convention seems like a potentially dangerous thing to do; if you think about it, we're basically taking a sum like the one in equation (4.2) and erasing the summation symbol $\sum_j$. You might imagine that erasing summation symbols $\sum_j$ will produce a bunch of ambiguities in the formalism. However, Einstein summation convention works (in the sense that it is unambiguous) because when performing a tensor calculation, the indices you sum over always come in a pair—one raised and one lowered. If you encounter more than two repeated indices or a pair of indices that are *both* raised or *both* lowered, you have either written down a nonsensical expression, or you have made a mistake.

Summation convention does have one limitation. If you want to refer to a *single* term in the sum, but you don't want to specify which one, you have to state that there is no sum implied. One way to get around this (though it is not standard[4]) is to underline the index pairs that you *do not* sum over, for instance: $M^i{}_{\underline{a}}\, v^{\underline{a}}$.

Exercise IV.1

If you did Exercise **III.2** properly, you would have noticed that your sum is consistent with Einstein's observation: the symbol for the indices that you sum over should appear twice in your expression, with one of the indices raised and the other lowered. If your result is not consistent, fix this. Write out the matrix multiplication $\mathbf{MPQ}$ in index notation using explicit sums (again, assume that $\mathbf{M}$, $\mathbf{P}$, and $\mathbf{Q}$ are $n \times n$ matrices). This time, you should get two sums, and—if you do this correctly—you will find that Einstein's observation[5] holds in this case as well.

---

[1] I'll take this opportunity to introduce some terminology, which I will repeat later on. Vectors v^i are often called *contravariant vectors*, and dual vectors w_i are called *covariant vectors*.

[2] I've also been told that Einstein himself made a remark to this effect.

[3] To be clear, summation convention DOES NOT apply to formula (4.2); by including the summation symbol $\sum_j$, I have made the summation explicit.

[4] What is usually done is to explicitly state that there is no sum (see, for instance, page 9 of [17]).

[5] Namely, the observation that sums in matrix and vector products occur over pairs of indices.

V. VECTORS

Before I provide the definition of a tensor, I must first provide a definition for a vector; in fact, you will later see that a vector is in fact an example of a tensor.

A vector is simply a directional derivative. Before you write me off as a nut, examine the directional derivative for some function $f(x^a)$:

$$\vec{v}\cdot\vec{\nabla}f(x^a) = v^i\,\frac{\partial f}{\partial x^i} \tag{5.1}$$

where x^a represent Cartesian coordinates on Euclidean space.[1] I use Einstein summation convention in equation (5.1)—**in Einstein summation convention, the index i for partial derivatives $\frac{\partial}{\partial x^i}$ is treated as if it were a lowered index.** Now I remove the function from equation (5.1) to obtain the directional derivative operator:

$$\vec{v}\cdot\vec{\nabla} = v^i\,\frac{\partial}{\partial x^i} \;\Leftrightarrow\; \vec{v}\cdot\vec{\nabla} = v^1\,\frac{\partial}{\partial x^1}+v^2\,\frac{\partial}{\partial x^2}+v^3\,\frac{\partial}{\partial x^3} \tag{5.2}$$

Now compare this with the vector $\vec{v}$ written out in terms of the orthogonal unit vectors $\hat{e}_a$ (if you are more familiar with the unit vectors $\hat{i}$, $\hat{j}$, and $\hat{k}$, then you can imagine that $\hat{e}_1=\hat{i}$, $\hat{e}_2=\hat{j}$, and $\hat{e}_3=\hat{k}$):

$$\vec{v} = v^i\,\hat{e}_i \qquad\Leftrightarrow\qquad \vec{v} = v^1\,\hat{e}_1+v^2\,\hat{e}_2+v^3\,\hat{e}_3 \tag{5.3}$$

Side-by-side, equations (5.2) and (5.3) suggest that **partial derivatives $\partial/\partial x^i$ can be thought of as basis vectors**! This is essentially what I'm claiming when I say that a vector is simply a directional derivative. The basis of partial derivatives, by the way, is called the *coordinate basis*.

Of course, in order for me to say that a vector is a directional derivative operator, I must show that **the directional derivative operator contains the same information as the information contained in the explicit components** (v^1,v^2,v^3). I can do one better—I can show you how to extract the components (v^1,v^2,v^3) from the directional derivative operator $\vec{v}\cdot\vec{\nabla}$. Let's see what happens when I feed the trivial function $f=x^3$ into the directional derivative operator (here, the "3" in x^3 is an *index*, NOT an exponent!):

$$\vec{v}\cdot\vec{\nabla}x^3 = v^i\,\frac{\partial x^3}{\partial x^i} = v^1\,\frac{\partial x^3}{\partial x^1}+v^2\,\frac{\partial x^3}{\partial x^2}+v^3\,\frac{\partial x^3}{\partial x^3} \tag{5.4}$$

Coordinates are independent of each other, so $\frac{\partial x^i}{\partial x^j}=0$ if $i\neq j$ and $\frac{\partial x^i}{\partial x^j}=1$ if $i=j$. A compact way of writing this is:

$$\frac{\partial x^i}{\partial x^j} = \delta^i_j \tag{5.5}$$

where δ^i_j is the Kronecker delta defined in (3.9). Equation (5.4) is then:

$$\vec{v}\cdot\vec{\nabla}x^3 = v^i\,\delta^3_i = v^3 \tag{5.6}$$

This result (equation (5.6)) means that all I have to do to pick out a component v^i of a vector is to feed the corresponding coordinate x^i into the directional derivative operator. In fact, you can even *define* the components v^i of the vector $\vec{v}$ this way:

$$v^i := \vec{v}\cdot\vec{\nabla}x^i \tag{5.7}$$

The whole point of this discussion is to motivate (in an operational sense) the definition of a vector as the following operator:

$$v(\cdot) := \vec{v}\cdot\vec{\nabla}(\cdot) = v^i\,\frac{\partial}{\partial x^i}(\cdot) \qquad\Rightarrow\qquad v = v^i\,\frac{\partial}{\partial x^i} \tag{5.8}$$

I drop[2] the arrow on the vector v to indicate that it is an operator now (with $(\cdot)$ being the place where you insert a function). Given the definition (5.8), the components (5.7) may be rewritten as:[3]

$$v^i := v(x^i) \tag{5.9}$$

where the right-hand side is to be interpreted as the (directional derivative) operator $v(\cdot)$ in (5.8) acting on the coordinate x^i.

To give you some intuition for the definition (5.8), I'll relate it to an example that may be more familiar to you. Consider a curve in Euclidean space. I may parameterize the curve by the functions $x^i(t)$; in particular, I write the coordinates of points that lie along the curve as functions of a parameter t. The notation here is meant to be suggestive; you might imagine that $x^i(t)$ describes the motion of a particle with respect to time. If I take the derivative of $x^i(t)$ with respect to t, I obtain a vector v^i that is tangent to the curve $x^i(t)$ (the *tangent vector* to the curve):

$$v^i = \frac{dx^i}{dt} \tag{5.10}$$

Note that this is a local expression; it can be evaluated at a single point in space. To see how this relates to the directional derivative definition (5.8) for a vector, consider

[1] Often, indices are suppressed in the arguments of a function: it is typical to write $f(x)$ rather than $f(x^a)$. However, there are some situations in which it is convenient to use the same symbol for two (or more) different quantities, with the distinction provided by the number of indices—for instance, one might write p^i for a vector and p^i_j for a matrix.

[2] Another reason for doing this is to get you comfortable with notation that is commonly used in the literature.

[3] Vectors that have components v^i with raised indices are sometimes called *contravariant vectors*.

a function[1] $\phi(x^a(t))$. Using the chain rule, the derivative of $\phi(x^a(t))$ with respect to t is:

$$\frac{d\phi(x^a(t))}{dt} = \frac{dx^i}{dt}\frac{\partial\phi}{\partial x^i} \qquad \Rightarrow \qquad \frac{d\phi}{dt} = v^i\frac{\partial\phi}{\partial x^i} \quad (5.11)$$

You might recognize the expression on the right-hand side of (5.11) as the directional derivative operator acting on ϕ. If v^i is the tangent vector to a curve parameterized by t, then the directional derivative operator is another way of writing derivatives with respect to t.

Before I move on, I must mention an assumption that I will make for the remainder of these notes. **I will always assume that v^i is a function of the coordinates x^a.** In particular, whenever I talk about vectors (dual vectors and tensors too, for that matter), I am actually referring to *vector fields*, which is a construction that assigns a vector to each point in space.[2]

Manifolds and Tangent Spaces

I have a few remarks for those interested in the applications of this formalism beyond that of Euclidean space, since the formalism here is designed to be flexible enough to describe spaces which have non-Euclidean geometries, such as curved surfaces. A generalization of Euclidean space that is often used is that of a *manifold*, denoted $\mathcal{M}$, which can be thought of as a set that has the (topological) properties of Euclidean space near a point (we say that manifolds are *locally Euclidean*). A *differentiable manifold* is a manifold that has all the properties needed to do calculus on it.

The intuition behind manifolds and differentiable manifolds is that if you look very closely at a point p on a curved 2d surface (like a globe), the immediate area surrounding p (called the *neighborhood* of p) looks like a small piece of a flat 2d plane (which is a 2d Euclidean space). In fact, I can imagine attaching a 2d plane to a curved 2d surface, so that in the neighborhood of p, it only touches the curved 2d surface at the point p; we say that the plane is *tangent* to the surface. It turns out that if you draw a path/curve on the curved 2d surface that passes through p, all the tangent vectors to that path at p will lie along (or are parallel to) a plane tangent to the surface. In other words, given a point p on a curved surface, all of the tangent vectors I can possibly construct from curves passing through p will lie along the same plane (called the tangent plane).

The example in the previous section motivates the concept of a *tangent space* of a manifold at point p (often denoted $T_p\mathcal{M}$), which is defined as the set of all directional derivative operators (5.8) evaluated at p, which I write as $v_p(\cdot) = v(\cdot)|_p$. In particular, these are operators that act on functions f in the following way:

$$v_p(f) = v(f)|_p = v^i\left.\frac{\partial f}{\partial x^i}\right|_p . \quad (5.12)$$

You can find a discussion of manifolds and their mathematical underpinnings in [2], [27], [20], [11], and [19]. Some General Relativity textbooks, such as [31] and [5], will also introduce and discuss the concept of a manifold.

[1] If it helps, you could imagine that ϕ is the potential energy for a particle, with a corresponding force $\vec{F} = \vec{\nabla}\phi$. If $x^a(t)$ is the trajectory of the particle, then (5.11) computes the power $P = dW/dt$ (W being work) applied to the particle by the force $\vec{F}$: $P = \vec{F}\cdot\vec{v}$.

[2] When working with tensors, you need to get rid of the idea that a single vector can be defined everywhere in space, if you haven't already. This idea only makes sense in Euclidean space, and it is only useful in Cartesian coordinates. Though these notes will focus on curvilinear coordinates in Euclidean space, the study of non-Euclidean spaces is the main motivation for doing tensor analysis, particularly in General Relativity.

Exercise V.1

In this section, I used Einstein summation convention throughout. Identify all the equations in this section that use Einstein summation convention, and re-insert the summation symbols Σ_i.

Exercise V.2

In linear algebra, a vector is defined as an element of a vector space. Show that the directional derivative operator (5.8) is indeed an element of a vector space.

Exercise V.3

You can describe a circle of radius R in Euclidean space as a curve $x^i(\theta)$ parameterized by the parameter θ, with the coordinates $x^i(\theta)$ explicitly given by the following formulas:

$$x^1(\theta) = R\cos(\theta)$$
$$x^2(\theta) = R\sin(\theta)$$
$$x^3(\theta) = 0$$

Use these formulas to obtain the components of the tangent vector $v^i(\theta) = \frac{dx^i}{d\theta}$. Draw a circle (use a compass!) on a sheet of graph paper, and draw arrows representing the vector v^i for various values of θ, with the tail ending at the point of the circle corresponding to each value of θ. In doing so, convince yourself that v^i is indeed a vector tangent to the circle.

VI. THE METRIC GENERALIZES THE DOT PRODUCT

You should be familiar with the dot product $u \cdot v$ for two vectors u and v. Note that I have again dropped the arrow notation for vectors; for the remainder of these notes, the symbols u and v will be used exclusively for vectors. In index notation, the dot product can be written as (assuming Cartesian coordinates on Euclidean space):

$$u \cdot v = \delta_{ij}\, u^i\, v^j \tag{6.1}$$

where δ_{ij} is the Kronecker delta with lower indices ($\delta_{ij} = 1$ for $i = j$ and $\delta_{ij} = 0$ for $i \neq j$).

The dot product is an example of an inner product[1] for vectors $\langle u, v\rangle$, which is a generalization of the dot product. The inner product may be written in terms of a quantity g_{ij} called the *metric*:

$$\langle u, v\rangle = g_{ij}\, u^i\, v^j \tag{6.2}$$

Since inner products are symmetric $\langle u, v\rangle = \langle v, u\rangle$, the metric has the following symmetry:

$$g_{ij} = g_{ji} \tag{6.3}$$

We require the existence of an *inverse metric*[2] g^{ij}, defined as the solution to the following equation:

$$g^{ik}\, g_{kj} = \delta^i_j \tag{6.4}$$

Recall that the Kronecker delta δ^i_j, defined in equation (3.9), is just the components of the identity matrix; this justifies the term "inverse metric" for g^{ij}.

For dot products in Cartesian coordinates on Euclidean space ($\langle u, v\rangle = u \cdot v$), we see that the metric is $g_{ij} = \delta_{ij}$. Explicitly, the metric may be written as the following table (the metric is not exactly a matrix, since both indices are lowered):

$$g_{ij} = \begin{pmatrix} g_{11} & g_{12} & g_{13} \\ g_{21} & g_{22} & g_{23} \\ g_{31} & g_{32} & g_{33} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{6.5}$$

It should not be too hard to infer (see exercise **VI.1** below) that the components of the inverse metric are $g^{ij} = \delta^{ij}$.

I'll take a moment to explain *meaning* of the metric components g_{ij}. Recall that a three-dimensional vector may be written in terms of three linearly independent basis vectors e_i in the following way:

$$v = v^i\, e_i \tag{6.6}$$

Here, I *do not* assume that e_i are orthogonal unit vectors; they may not be of unit length and they may not be orthogonal to each other. Since e_i are themselves vectors, I can take inner products of the basis vectors: $\langle e_i, e_j\rangle$. Using the properties[3] of inner products, you can show (see exercise **VI.2**) that $\langle e_i, e_j\rangle$ form the components of the metric tensor:

$$g_{ij} = \langle e_i, e_j\rangle \tag{6.7}$$

Thus, the metric components g_{ij} are just inner products between basis vectors. In the previous section, I introduced the idea that partial derivatives $\frac{\partial}{\partial x^i}$ are basis vectors. Equation (6.7) in turn suggests that the metric components g_{ij} define inner products for partial derivatives:

$$g_{ij} = \left\langle \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \right\rangle \tag{6.8}$$

In the literature, you may encounter statements to the effect that the metric provides a way to measure distances in space. This is because the metric can be used to construct the *line element*:

$$ds^2 = g_{ij}\, dx^i\, dx^j \tag{6.9}$$

which can be thought of as the norm $\langle dx, dx\rangle$ of an infinitesimal displacement vector dx^i. In the usual x-y-z variables for Cartesian coordinates, line element may also be written as (the superscript 2 in (6.10) is an *exponent*, not an index):

$$ds^2 = dx^2 + dy^2 + dz^2 \tag{6.10}$$

Given a curve $x^i(s)$ parameterized by a parameter t, the line element can be used to measure distances along the curve. The line element (6.9) can be thought of as the square of the infinitesimal distance ds along the curve, and the distance Δs between any two points $x^i_1 = x^i(t_1)$ and $x^i_2 = x^i(t_2)$ is given by the following integral:

$$\Delta s = \int_{t_1}^{t_2} \sqrt{g_{ij}\, \frac{dx^i}{dt}\, \frac{dx^j}{dt}}\, dt \tag{6.11}$$

---

[1] In relativity, we often use a definition for inner products that may be slightly different from those used in your linear algebra classes (see definition 4.1 in [29])). In particular, we define an inner product by the following four properties for vectors p, u, v and scalar α:

1. $\langle u + p, v\rangle = \langle u, v\rangle + \langle p, v\rangle$
2. $\langle \alpha\, u, v\rangle = \alpha \langle u, v\rangle$
3. $\langle u, v\rangle = \langle v, u\rangle$
4. Nondegeneracy: There exists no nonzero vector u such that $\langle u, v\rangle = 0$ holds for *all* vectors v

The difference is in property **4**; in linear algebra classes, the positive-definite condition ($\langle v, v\rangle \geq 0$ for all v) is often used in place of property **4**. In special relativity, we use property **4** because special relativity requires a notion of inner product for which $\langle v, v\rangle$ can have any value for nonzero vectors v.

[2] Confusingly, some authors refer to *both* g_{ij} and g^{ij} as the "metric." You can get away with this, since the index placement in g^{ij} is used to indicate that it is the inverse of g_{ij}, and g_{ij} and g^{ij} can be thought of as two different ways of writing down the same information (since in principle, you can get one from the other).

[3] See footnote 1.

You can derive (6.11) from (6.9) by taking the square root of (6.9) to get a formula for ds, and using the physicist's trick of multiplying and dividing by differentials:

$$ds = \sqrt{g_{ij}\, dx^i\, dx^j} = \sqrt{g_{ij}\, dx^i\, dx^j}\, \frac{dt}{dt} = \sqrt{g_{ij}\, \frac{dx^i}{dt}\, \frac{dx^j}{dt}}\, dt \tag{6.12}$$

One can then integrate equation (6.12) to obtain (6.11).

I must make one last remark before concluding the present discussion. **The metric/inner product can be thought of as a (linear[1]) "machine" that "eats" two vectors and "spits out" a scalar.** We will later see that tensors can be defined by a similar characterization.

Exercise VI.1

Show (or argue[2]) that in Cartesian coordinates (where $g_{ij} = \delta_{ij}$), the components of the inverse metric g^{ij} are given by $g^{ij} = \delta^{ij}$, where $\delta^{ij} = 1$ for $i = j$ and $\delta^{ij} = 0$ for $i \neq j$.

Exercise VI.2

Show that the inner products between basis vectors e_i form the components of the metric g_{ij}; in particular, show that $g_{ij} = \langle e_i, e_j \rangle$. You can do this by writing out $\langle u, v \rangle$ as the explicit sum (dropping Einstein summation convention):

$$\langle u, v \rangle = \sum_{i=1}^{3} \sum_{j=1}^{3} \langle u^i\, e_i, v^j\, e_j \rangle$$

Here, it is appropriate to think of u^i and v^i as "scalars" in the sense that they are just coefficients in front of the basis vectors e_i. In particular, show that:

$$\sum_{i=1}^{3} \sum_{j=1}^{3} \langle u^i\, e_i, v^j\, e_j \rangle = \sum_{i=1}^{3} \sum_{j=1}^{3} u^i\, v^j\, \langle e_i, e_j \rangle$$

$$\Rightarrow \qquad \langle u, v \rangle = \sum_{i=1}^{3} \sum_{j=1}^{3} u^i\, v^j\, \langle e_i, e_j \rangle$$

by expanding the sums into nine terms and using the property $\langle \alpha\, u, v \rangle = \alpha\, \langle u, v \rangle$ (where α is a scalar) on each of the nine terms. Finally, compare the result (specifically the formula to the right of the arrow symbol "$\Rightarrow$") to equation (6.2).

Exercise VI.3

Consider a circle of radius R in Euclidean space. Recall the formulas for the parameterization of the circle $x^i(\theta)$ from exercise **V.3**. Using the metric $g_{ij} = \delta_{ij}$, show that the integral (6.11) yields the correct value for the circumference of the circle.

---

[1] In the sense of properties **1.** and **2.** of footnote 1 on page 9.

[2] Hint: This might be easy once you recognize that (6.4) has the same form of as a matrix multiplication formula.

VII. DUAL VECTORS

In this section, I wish to introduce a new type of vector, which I call a *dual vector* (also called a *one-form*[1], or a *covariant vector*). Simply put, a *dual vector* is a quantity that "eats" a vector and "spits out" a scalar (or a real number). Explicitly, a dual vector $w(\cdot)$ is an operator that does the following:

$$w(v) = w_i\, v^i \tag{7.1}$$

where w_i are the components of the dual vector (note the lowered index!). In a similar manner, you can think of vectors v as being "dual" to dual vectors; vectors can be thought of as operators that "eat" dual vectors and "spit out" scalars:[2]

$$v(w) = v^i\, w_i \tag{7.2}$$

I will take the opportunity here to introduce some notation. It is standard to write expressions like (7.1) and (7.2) using the following "inner product" notation:

$$\langle w, v\rangle = \langle v, w\rangle = w_i\, v^i = v^i\, w_i \tag{7.3}$$

This differs from the inner product in equation (6.2) in that one of the arguments in $\langle\,,\rangle$ is a dual vector. There is no ambiguity here; if both of the arguments are vectors, $\langle\,,\rangle$ is given by equation (6.2), and if one of the arguments is a vector and one is a dual vector, $\langle\,,\rangle$ is given by equation (7.3). Later, I'll give you an exercise where you show that $\langle\,,\rangle$ as defined in (7.3) satisfies the same linearity properties as the usual inner product. In particular, for vectors v, p, dual vectors w, q, and scalars α, β, we have the properties:

$$\begin{aligned}\langle v + \alpha\, p, w + \beta\, q\rangle &= \langle v, w + \beta\, q\rangle + \alpha\, \langle p, w + \beta\, q\rangle \\ &= \langle v + \alpha\, p, w\rangle + \beta\, \langle v + \alpha\, p, q\rangle \\ &= \langle v, w\rangle + \beta\, \langle v, q\rangle + \alpha\, \langle p, w\rangle + \alpha\, \beta\, \langle p, q\rangle\end{aligned} \tag{7.4}$$

The notion that vectors and dual vectors are "dual" to each other (in the sense of equations (7.1), (7.2) and (7.3)) is central to tensor calculus—we will return to this in the next section. For now, I must say a few more things about dual vectors.

A natural set of basis elements (or "basis dual vectors") for the dual vector are the coordinate differentials dx^i, so that a dual vector may be written as:[3]

$$w = w_i\, dx^i \tag{7.5}$$

This is in contrast to vectors, which have partial derivatives $\frac{\partial}{\partial x^i}$ as basis elements (or basis vectors). The sense in which coordinate differentials dx^i form a natural set of basis elements comes from the differential of a function, which is an example of a dual vector. The differential df of a function $f(x^a)$ is defined as:

$$df := \frac{\partial f}{\partial x^i}\, dx^i \tag{7.6}$$

The components of the differential are just the components of gradient of the function. To simplify things, I'll use the symbol $\mathcal{G}_i$ to represent the components of the gradient:

$$\mathcal{G}_i := \frac{\partial f}{\partial x^i} \qquad \Rightarrow \qquad df := \mathcal{G}_i\, dx^i \tag{7.7}$$

The index i in $\mathcal{G}_i$ is lowered because as stated earlier, the index i in partial derivatives $\frac{\partial}{\partial x^i}$ are treated as lowered indices in Einstein summation convention. To see that $\mathcal{G}_i$ do indeed form the components of a dual vector (in the sense of equation(7.1)), apply the directional derivative operator v as defined in (5.8) to the function f. Using (7.7), I can write the following expression for $v(f)$:

$$v(f) = v^i\, \frac{\partial f}{\partial x^i} = v^i\, \mathcal{G}_i \tag{7.8}$$

Upon comparing the above with equation (7.1) and recalling that $v(f)$ is a scalar, we see that the $\mathcal{G}_i$ can be used to form the components of a dual vector $\mathcal{G} = df$ that "eats" v and "spits out" a scalar.

The basis elements dx^i are themselves dual vectors, in the same way that basis vectors are themselves vectors, be they partial derivatives $\frac{\partial}{\partial x^i}$ or unit vectors $\hat{e}_i$. There must also be a sense in which dx^i "eats" vectors and spits out scalars. Equation (7.3), combined with the expressions $v = v^i\, \frac{\partial}{\partial x^i}$ and $w = w_i\, dx^i$ can be used to infer the following formula for the basis elements:

$$\left\langle dx^j, \frac{\partial}{\partial x^i}\right\rangle = \left\langle \frac{\partial}{\partial x^i}, dx^j\right\rangle = \delta^j_i \tag{7.9}$$

The formula above states that when dx^j "eats" a basis vector $\frac{\partial}{\partial x^i}$, it "spits out" 1 or 0 depending on whether $i =$

---

[1] Dual vectors are referred to as one-forms because they are an example of a class of tensors called *differential forms*, which are beyond the scope of these notes.

[2] I'm being a bit sloppy here, since the dual of a dual vector is technically a *double dual*, not a vector. However, the vector space formed by double duals is isomorphic to the vector space that the vector v lives in, so it possesses the same algebraic structure. This is why we can get away with treating double duals as if they are vectors.

[3] Given equation (7.5), you might wonder if a dual vector w can always be written as a differential of some scalar function f. Equivalently, you might ask whether there always exists a function f such that $w_i = \frac{\partial f}{\partial x^i}$. It turns out that in general, you can't; this is only possible if w_i satisfies the (Frobenius) integrability condition: $\frac{\partial w_i}{\partial x^j} = \frac{\partial w_j}{\partial x^i}$ (as an exercise, show that the gradient components $\mathcal{G}_i$ satisfy this property). This integrability condition is a special case of the Frobenius theorem, the details of which can be found in [3, 11, 16, 27].

j. It also states[1] that when the basis vector $\frac{\partial}{\partial x^j}$ "eats" dx^i, it also "spits out" 1 or 0 depending on whether $i = j$.

When the Differential df Eats a Vector

To give you some intuition for dual vectors, imagine that at some point p, the components of the vector v^i form the coordinates for some 3d Euclidean space, which I'll call $T_p\mathcal{M}$ (which is a fancy name for the tangent space of a manifold—recall the remarks at the end of section **V**). We can construct scalar functions $\phi(v^a)$ on $T_p\mathcal{M}$, just like we can in ordinary Euclidean space. A dual vector w is just a linear function $\phi(v^a) = w_i\, v^i$ on the Euclidean space $T_p\mathcal{M}$. It is sometimes helpful to visualize what's going on: $\phi(v^a)$ is a linear function, so surfaces defined by constant values for ϕ are 2d planes in the 3d Euclidean space $T_p\mathcal{M}$. If you imagine the vector v as an arrow in $T_p\mathcal{M}$ (with the tail at the origin), the tip of the vector lies on the 2d plane corresponding to the value of ϕ at the tip ($w(v) = \phi(v^a)$).[2]

If you would like a less abstract (but sloppier) example, recall that the differential df is a dual vector, which eats a vector v^i and spits out a scalar:

$$df(v) = \mathcal{G}_i\, v^i \tag{7.10}$$

In Cartesian coordinates on Euclidean space, displacement vectors Δx^i make sense. Now pretend for a moment that v is a small displacement vector $v^i = \Delta x^i$. If the magnitude of v is small enough, then can then interpret $df(v)$ as the change in value of the function f between the tip and tail of the displacement vector $v^i = \Delta x^i$:

$$df(v) \approx f(\text{tip of } v) - f(\text{tail of } v)$$

If you think about it, this isn't terribly surprising, since this is virtually the same statement as the definition of the differential (7.6): $df = \mathcal{G}_i\, dx^i$, if you imagine that dx^i is a small displacement vector.

Exercise VII.1

Show that $\langle v, w\rangle = v^i\, w_i$ (equation (7.3) satisfies the linearity properties (7.4).

Exercise VII.2

Convince yourself of (7.9). You can do this by writing out $\langle v, w\rangle$ explicitly:

$$\begin{aligned}\langle v, w\rangle &= \left\langle v^i \frac{\partial}{\partial x^i}, w_j\, dx^j \right\rangle \\ &= \left\langle v^1 \frac{\partial}{\partial x^1} + v^2 \frac{\partial}{\partial x^2} + v^3 \frac{\partial}{\partial x^3}\,, \right. \\ &\qquad \left. w_1\, dx^1 + w_2\, dx^2 + w_3\, dx^3 \right\rangle\end{aligned}$$

Then, use linearity properties (7.4) to further expand the above into an expression containing nine terms of the form $\alpha\, \langle dx^j, \frac{\partial}{\partial x^i}\rangle$, where α is a coefficient of the form $v^a\, w_b$. Finally, ask yourself what conditions $\langle dx^j, \frac{\partial}{\partial x^i}\rangle$ must satisfy in order to obtain the desired result (7.3): $\langle v, w\rangle = v^i\, w_i = v^1\, w_1 + v^2\, w_2 + v^3\, w_3$

---

[1] In this context, you should think of $\frac{\partial}{\partial x^j}$ as an abstract basis vector, rather than a partial derivative.

[2] If you tried to learn about differential forms from [18], this is essentially what the whole "bongs of a bell" example is all about. I recall being bewildered by the explanation in [18], but my good friend Luis Suazo eventually clarified it for me.

VIII. SOME BASIC INDEX GYMNASTICS

In section III, I repeatedly shouted the statement: **INDEX PLACEMENT IS IMPORTANT!** If you have had exposure to Cartesian tensors, you may wonder how so many authors can get away with ignoring the issue of index placement. This section will reveal the reason why you can get away with this in Cartesian coordinates.

If you have a metric g_{ij}, you can convert vectors to dual vectors. Explicitly, this may be done by performing the following operation on the components v^i:

$$v_j = g_{ij}\, v^i \tag{8.1}$$

The quantities v_j form the components of a dual vector $\tilde{v}$, which may be written as:

$$\tilde{v}(\cdot) = \langle v, \cdot \rangle \tag{8.2}$$

where the argument $(\cdot)$ requires a vector. If we feed a vector u into the dual vector $\tilde{v}$, we recover the inner product (6.2) between vectors u and v:

$$\tilde{v}(u) = \langle v, u \rangle = g_{ij}\, v^i\, u^j \tag{8.3}$$

Note that the notation v_j introduced in (8.1) provides a slick way to write the inner product: $\langle v, u \rangle = v_j\, u^j$.

We see that the metric can turn vectors into dual vectors. It would be nice to do the reverse; I'd like to take a dual vector and turn it into a vector. To do this, recall that we require the existence of an inverse metric g^{ij}, defined as the solution to equation (6.4): $g^{ik} g_{kj} = \delta^i_j$. The inverse metric g^{ij} can then be used to turn the components of a dual vector w_j into the components of a vector w^i:

$$w^i = g^{ij}\, w_j \tag{8.4}$$

The process of using metrics and inverse metrics to convert vectors to dual vectors and vice versa is called the *lowering* and *raising of indices*. In Cartesian coordinates, the metric is just the Kronecker delta $g_{ij} = \delta_{ij}$, as is the inverse metric $g^{ij} = \delta^{ij}$. As a result, the raising and lowering of indices will not change the value of the components: in Cartesian coordinates, v_1 will have the same value as v^1, v_2 will have the same value as v^2, and so on. This is why one can get away with neglecting index placement when working with Cartesian tensors. In curvilinear coordinates, however, the metric no longer has this simple form, and index placement becomes important.[1]

The raising and lowering of indices is part of a formalism for manipulating indices called *index gymnastics*. I won't go through all of the techniques in the formalism here, but you will encounter most of the rule in one way or another in these notes.[2] On the other hand, I would like to introduce a few of them before moving on to the next section.

An important technique is the *contraction of indices*, which is a generalization of the trace operation for matrices. Given an object with two indices or more, index contraction is the process of relabeling (and the repositioning) of an index so that it has the same symbol as another index. Einstein summation convention then implies that there is a sum over the indices. For a matrix $\mathbf{M}$ with components $M^i{}_j$, the contraction of indices is just the trace:

$$M^i{}_i = M^1{}_1 + M^2{}_2 + M^3{}_3 \tag{8.5}$$

Now consider a quantity with three indices: Q^{ijk} (don't worry about visualizing it; just think about it as a collection of $3 \times 3 \times 3 = 27$ variables). If I wish to contract the indices i and k, I can use the metric g_{ia} to lower the index i to obtain a quantity $Q_a{}^{jk}$:

$$Q_a{}^{jk} = g_{ia}\, Q^{ijk} \tag{8.6}$$

I then relabel the index a: explicitly, I replace the symbol a with the symbol k to get $Q_k{}^{jk}$. The result is Q^j, which are the components of a vector:

$$Q^j = Q_k{}^{jk} = g_{ik}\, Q^{ijk} \tag{8.7}$$

Index contraction provides a way to reduce the number of indices for a quantity. There is an operation, called the *tensor product*, which can be use to construct quantities with more indices. The tensor product is a straightforward operation: given the components of a vector v^i and a dual vector w_i, multiply them to form the components of a matrix[3] $K^i{}_j$:

$$K^i{}_j := v^i\, w_j \tag{8.8}$$

Tensor products can be performed on quantities with more indices. For instance, you could multiply the quantity Q^{ijk} with the matrix components $M^i{}_j$ to form the components of a big 5-indexed quantity $B^{ijkl}{}_m$:

$$B^{ijkl}{}_m := Q^{ijk}\, M^l{}_m \tag{8.9}$$

A very useful tool for index manipulation is the *relabeling of dummy indices*. *Dummy indices* refer to the indices that I sum over, for instance the indices i and k in the expression $g_{ik}\, Q^{ijk}$ (8.7). They are called dummy indices because I can change the symbols that are being

---

[1] A professor in one of my undergraduate classes once made a remark to the following effect: *You can spend the rest of your life working in Cartesian coordinates, but it would be a very miserable life!*

[2] If you want, see page 85 of [18] for a list of the techniques that make up index gymnastics.

[3] A warning: In general, it is *not* always possible to write matrix components as a product of vector components. One example is the identity matrix/Kronecker delta δ^i_j; there exist no vectors that can generate δ^i_j by way of formula (8.8).

summed over without affecting the meaning of the expression. For instance, the expressions $w_i\, v^i$ and $w_j\, v^j$ have the same exact meaning:

$$w_i\, v^i = w_j\, v^j \tag{8.10}$$

If the above isn't clear, it may help to write the above (8.10) as an explicit sum (we drop Einstein summation convention here):

$$\sum_{i=1}^{3} w_i\, v^i = \sum_{j=1}^{3} w_j\, v^j = w_1\, v^1 + w_2\, v^2 + w_3\, v^3 \tag{8.11}$$

A more illustrative example is the expression $g_{ik}\, Q^{ijk}$ in (8.7). I can change the label i to a and k to s in the expression $g_{ik}\, Q^{ijk}$ to obtain $g_{as}\, Q^{ajs}$; both $g_{ik}\, Q^{ijk}$ and $g_{as}\, Q^{ajs}$ have the same meaning:

$$g_{ik}\, Q^{ijk} = g_{as}\, Q^{ajs} \tag{8.12}$$

If you need more convincing, write both sides of (8.12) as explicit sums and expand—you should see that in each case, you get the same result (exercise **VIII.4**). The relabeling of dummy indices can be an extremely useful tool when dealing with long expressions—it is particularly useful for identifying quantities that are equivalent to each other (see exercise **VIII.5**).

Another set of techniques is the *symmetrization* and *antisymmetrization* of indices. A quantity S_{ij} is said to be *symmetric* in the indices i and j if it satisfies the property $S_{ij} = S_{ji}$. Given some quantity P_{ij} that is *not* symmetric in i and j, I can *symmetrize* the indices i and j by performing the following operation:

$$P_{(ij)} := \frac{1}{2}\,(P_{ij} + P_{ji}) \tag{8.13}$$

where $P_{(ij)}$ is called the *symmetric part* of P_{ij}. A quantity A_{ij} is said to be *antisymmetric* in the indices i and j if it satisfies the property $A_{ij} = -A_{ji}$. If P_{ij} that is *not* antisymmetric in i and j, I can *antisymmetrize* the indices i and j by performing the following operation:

$$P_{[ij]} := \frac{1}{2}\,(P_{ij} - P_{ji}) \tag{8.14}$$

where $P_{[ij]}$ is called the *antisymmetric part* of P_{ij}.

Exercise VIII.1

Let u_i and v_j be dual vector components obtained from the respective vector components u^i and v^j. Show that the inner product $\langle u, v\rangle$ can be written as: $u_i\, v^i$. Also show that $u^i\, v_i$ is equivalent to $u_i\, v^i$.

Exercise VIII.2

Take the trace of δ^i_j, assuming 3d Euclidean space. Now do the same for 2d Euclidean space (in 2d Euclidean space, the indices i and j only take on two values: 1 and 2). Do the same for 4d Euclidean space. You should be able to deduce the result for the trace in n-dimensional Euclidean space. Now contract the indices of δ_{ij} with those of δ^{ij} (compute $\delta_{ij}\,\delta^{ij}$); how does your result compare with the trace of δ^i_j?

Exercise VIII.3

Consider the quantity Z_{ijkl}. Contract the indices k and l, and raise an index to obtain a matrix $S^i{}_j$. Write down the expression for $S^i{}_j$ in terms of the original quantity Z_{ijkl} and the inverse metric g^{ij}. Contract the indices of $S^i{}_j$, and write down the result in terms of Z_{ijkl} and g^{ij}.

Exercise VIII.4

Write $g_{ik}\, Q^{ijk}$ and $g_{as}\, Q^{ajs}$ as explicit sums. Now expand the result (you should have nine terms), and in doing so, show that equation (8.12) is valid: $g_{ik}\, Q^{ijk} = g_{as}\, Q^{ajs}$.

Exercise VIII.5

Relabel dummy indices to show that the following three quantities (which are all scalars) are equivalent to $g_{jl}\, Q^{ijk}\, M^l{}_i\, v_k$:

$$g_{ki}\, Q^{jkl}\, M^i{}_j\, v_l$$
$$Q^{kil}\, v_l\, M^j{}_k\, g_{ij}$$
$$v_d\, M^b{}_c\, g_{ab}\, Q^{cad}$$

It is quite possible to do this by inspection, but if you wish to show this explicitly, you can do this by writing out the expression each time you relabel a pair of indices, proceeding until you have the expression $g_{jl}\, Q^{ijk}\, M^l{}_i\, v_k$.

Exercise VIII.6

Let A_{ij} be an antisymmetric quantity, so that $A_{ij} = -A_{ji}$, and S_{ij} be a symmetric quantity, so that $S_{ij} = S_{ji}$. Show that if you raise the indices of both quantities to obtain A^{ij} and S^{ij}, they satisfy the same properties: $A^{ij} = -A^{ji}$ and $S^{ij} = S^{ji}$. Use this result to show that $A^{ij}\, S_{ij} = 0$ and $A_{ij}\, S^{ij} = 0$.

Exercise VIII.7

Show that if you symmetrize an antisymmetric quantity A_{ij} ($A_{ij} = -A_{ji}$), you get zero: $A_{(ij)} = 0$. Likewise, show that if you antisymmetrize a symmetric quantity S_{ij} ($S_{ij} = S_{ji}$), you also get zero: $S_{[ij]} = 0$. Finally, show that if you have a quantity P_{ij} that is *neither* symmetric or antisymmetric ($P_{ij} \neq P_{ji}$ and $P_{ij} \neq -P_{ji}$), then you can decompose it into its antisymmetric parts; in other words, show that $P_{ij} = P_{(ij)} + P_{[ij]}$.

IX. COORDINATE INVARIANCE AND TENSORS

In this section, I will finally reveal to you the precise definition of a tensor, and show how it follows from the principle that the *value* of a scalar function at a *point* is unaffected by coordinate transformations. This is because coordinates are just labels or "names" that we give to points in space.[1] Now I must make an important distinction here: the principle I mentioned is only valid if I think of scalar functions as functions of points, *not* coordinates. From this section on, I will always define scalar functions so that **scalar functions are functions of *points*, not coordinates.** It follows that a coordinate transformation can change how a scalar function depends on *coordinates*, but a coordinate transformation does NOT change how a scalar function depends on *points*.[2]

A vector is a geometric quantity. In kindergarten,[3] we are taught that vectors consist of a magnitude and direction. How we represent the magnitude and the direction of the vector depends on the coordinates. Again, coordinates are just labels or "names" that we give to points in space, so they should not alter the geometric properties of a vector (its magnitude, for instance); geometry does not care about the way we choose to "name" points in space.

Intuitively, this means that coordinate transformations cannot change the meaning of an inner product or a directional derivative. More precisely, both inner products and directional derivatives acting on a function yield scalars, and the value of a scalar at a point should not depend on the "name" (coordinates) we give to that point. In other words, if I evaluate the following scalar quantities at some point p, the numerical values should be unaffected by coordinate transformations:

$$\langle u, v\rangle = g_{ij}\, u^i\, v^j \tag{9.1}$$

$$v(f) = v \cdot \nabla f \tag{9.2}$$

On the other hand, the values for the vector components u^i, v^i and the values for the gradient components $\mathcal{G}_i = \frac{\partial f}{\partial x^i}$ *do* change under coordinate transformations. To proceed, I need to know the transformation law for vectors; in particular, I want to know how the components v^i of the vectors change under a coordinate transformation.

Before I derive the transformation law for vectors, let me first describe the coordinate transformation I wish to

---

[1] See chapter 1 of [18] to see how points in space can have a physical meaning independent of coordinate labels.

[2] For this reason, all formulas in this section should be thought of as being evaluated at the same *point*, irrespective of the coordinate values for that point.

[3] This is a reference to Steven Weinberg, who says this when recalling an elementary concept in his lectures.

perform. Consider a coordinate transformation $y^\alpha(x^a)$, where y^α are my new coordinates, and x^i are my old coordinates. Lowercase Greek[1] letters $\alpha, \beta, \gamma, ...$ will be used for indices corresponding to new coordinates y^α, and lowercase Latin indices $a, b, c, ...i, j, k, ...$ will be used for indices corresponding to the old coordinates x^i. I assume that the functions $y^\alpha(x^a)$ are invertible, so that I can obtain from it the functions $x^i(y^\alpha)$ (and vice-versa). I can take derivatives of the functions $y^\alpha(x^a)$ and $x^i(y^\alpha)$, to obtain the transformation matrices—the quantities $\frac{\partial x^j}{\partial y^\beta}$ and $\frac{\partial y^\beta}{\partial x^j}$ form the components of the transformation matrices. The chain rule tells me that the derivatives/transformation matrices $\frac{\partial x^j}{\partial y^\beta}$ and $\frac{\partial y^\beta}{\partial x^j}$ are inverses of each other in the following way:

$$\begin{aligned} \frac{\partial x^i}{\partial y^\alpha}\frac{\partial y^\alpha}{\partial x^j} &= \frac{\partial x^i}{\partial x^j} = \delta^i_j \\ \frac{\partial y^\alpha}{\partial x^i}\frac{\partial x^i}{\partial y^\beta} &= \frac{\partial y^\alpha}{\partial y^\beta} = \delta^\alpha_\beta \end{aligned} \tag{9.3}$$

The last equality comes from the assumption that the coordinates are independent of each other (cf. (5.5)).

The chain rule also tells us how the components of the gradient transforms:

$$\frac{\partial f}{\partial y^\beta} = \frac{\partial x^j}{\partial y^\beta}\frac{\partial f}{\partial x^j} \tag{9.4}$$

As discussed in section VII, the components of the gradient $\mathcal{G}_i = \frac{\partial f}{\partial x^i}$ are the components of a dual vector, which suggests the following transformation law for the components of dual vectors:

$$w_\beta = \frac{\partial x^j}{\partial y^\beta} w_j \tag{9.5}$$

If v is a vector and w is a dual vector, then the value of $w(v) = w_i\, v^i$, being a scalar, cannot be affected by a coordinate transformation. For the value of $w(v)$ to remain unchanged under a coordinate transformation, the transformation law for v^i must be ***opposite***[2] to that of w_i:

$$v^\alpha = \frac{\partial y^\alpha}{\partial x^i} v^i \tag{9.6}$$

Under a coordinate transformation, I can then write the following expression for $w(v)$:

$$\begin{aligned} w(v) = w_\alpha\, v^\alpha &= \frac{\partial x^j}{\partial y^\alpha} w_j \frac{\partial y^\alpha}{\partial x^i} v^i = \frac{\partial x^j}{\partial y^\alpha}\frac{\partial y^\alpha}{\partial x^i} w_j\, v^i \\ &= \delta^j_i\, w_j\, v^i = w_i\, v^i \\ \Rightarrow \qquad & w_\alpha\, v^\alpha = w_i\, v^i \end{aligned} \tag{9.7}$$

This computation (9.7) demonstrates that the transformation laws (9.5) and (9.6), combined with (9.3), guarantee that $w(v)$ is unchanged under coordinate transformations. In short, I have shown that:

> **If v^i transforms in a manner that is "opposite" to the transformation of w_i under a change of coordinates, then the transformations in $w(v) = w_i\, v^i$ cancel out, ensuring that the value of the scalar $w(v)$ remains unchanged under coordinate transformations.**

Inner products are invariant under coordinate transformations. This demand establishes the transformation law for the metric tensor. If you examine the expression (9.1) for the inner product $\langle u, v\rangle = g_{ij}\, u^i\, v^j$, you may realize that the indices of the metric tensor must acquire transformation matrices that cancel out the transformation matrices that the vector components u^i and v^i acquire under a coordinate transformation. The metric tensor therefore satisfies the following transformation law:

$$g_{\alpha\beta} = \frac{\partial x^i}{\partial y^\alpha}\frac{\partial x^j}{\partial y^\beta}\, g_{ij} \tag{9.8}$$

Since the inner product can also be written as $\langle u, v\rangle = g^{ij}\, u_i\, v_j$ (see exercise **VIII.1**), the inverse metric can be similarly shown to satisfy the following transformation law:

$$g^{\alpha\beta} = \frac{\partial y^\alpha}{\partial x^i}\frac{\partial y^\beta}{\partial x^j}\, g^{ij} \tag{9.9}$$

Now recall the characterization of the metric at the end of section VI as a (linear[3]) "machine" that "eats" two vectors and "spits out" a scalar. One might imagine a more general construction, which "eats" any number of vectors and dual vectors, and "spits out" a scalar. This construction is what we call a *tensor*, which is defined by the statements:

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

---

[1] If you plan to study relativity, I must alert you to the fact that this notation is nonstandard (usually, primes are put on indices). In relativity, Greek indices are typically reserved for coordinates on spacetime, and Latin indices either used to denote spatial indices, spacetime indices in the orthonormal basis (see for instance Appendix J in [5]), or as "abstract" indices, which are used to keep track of tensor arguments (the latter is called abstract index notation, the details of which are discussed in [31]).

[2] The transformation is opposite in that that it uses the transformation matrix $\frac{\partial y^\alpha}{\partial x^i}$ instead of $\frac{\partial x^j}{\partial y^\beta}$; recall that from equation (9.3), $\frac{\partial y^\alpha}{\partial x^i}$ is the *inverse* of $\frac{\partial x^j}{\partial y^\beta}$ and vice versa.

[3] In the sense of properties **1.** and **2.** of footnote 1 on page 9.

By *linear map*,[1] I mean that a tensor must be a linear function[2] of the things it "eats" and also that a tensor vanishes when it "*eats zero*", meaning that if you feed a tensor a zero vector $v^i = 0$ or a dual vector with all vanishing components $w_i = 0$, it automatically returns zero.[3] These conditions imply that a tensor that eats one vector v and one dual vector w must have the following form:

$$T(w, v) = T^i{}_j \, w_i \, v^j \tag{9.10}$$

Note that the above formula is linear[4] in both w_i and v^i, and vanishes if *either* $w_i = 0$ or $v^i = 0$ (as per the "eats zero" property). For $T(w, v)$ to be a scalar, the components $T^i{}_j$ must transform in the following way:

$$T^\alpha{}_\beta = \frac{\partial y^\alpha}{\partial x^i} \frac{\partial x^j}{\partial y^\beta} \, T^i{}_j \tag{9.11}$$

At this point, you may notice a pattern in equations (9.8), (9.9) and (9.11). In all cases, raised indices transform just like vector indices, and lowered indices transform like dual vector indices. This leads to the following properties of tensor components:

Raised indices of tensor components transform like vector indices

Lowered indices of tensor components transform like dual vector indices

These properties are often used as *defining* properties of tensors; in much of the literature, tensors are defined by components which transform according to the two properties listed above.

Tensors can in general "eat" an arbitrary number of vectors and dual vectors. The number of vectors and dual vectors a tensor "eats" is called the *rank* of the tensor. The two indexed tensor $T^i{}_j$ defined in (9.10) is called a rank 2 tensor (to be more precise, we say that it is a rank $(1, 1)$ tensor, since it eats one vector and one dual vector). An example of a rank 5 tensor (or rank (3,2) tensor), which eats the vectors v, u, q and the dual vectors w, p, is the following:

$$B(w, p, v, u, q) = B^{ij}{}_{klm} \, w_i \, p_j \, v^k \, u^l \, q^m \tag{9.12}$$

Note that the above expression is linear in the individual components w_i, p_j, v^k, u^l, and q^m, and vanishes if any one of the vectors or dual vectors is zero (the "eats zero" property).

The metric g_{ij} and its inverse g^{ij} are examples of tensors; they are both linear maps, and respectively "eat" vectors and dual vectors, and spit out scalars. Vectors and dual vectors are also themselves examples of tensors (they are both rank-1 tensors); a vector is a linear map that "eats" a single dual vector and spits out a scalar, and a dual vector is a linear map that "eats" a single vector and spits out a scalar:

$$\begin{aligned} v(w) &= v^i \, w_i \\ w(v) &= w_i \, v^i \end{aligned} \tag{9.13}$$

I've given you the definition for a tensor. Now, I'll deliver the punchline for these notes:

Tensors are useful because tensor equations look the same in all coordinate systems.

As an example, consider the following tensor equation:

$$G^i{}_j = \kappa \, T^i{}_j \tag{9.14}$$

where $G^i{}_j$ and $T^i{}_j$ are (rank-2) tensors, and κ is some constant. Under a coordinate transformation, $G^i{}_j$ and $T^i{}_j$ transform as:

$$\begin{aligned} T^\alpha{}_\beta &= \frac{\partial y^\alpha}{\partial x^i} \frac{\partial x^j}{\partial y^\beta} \, T^i{}_j \\ G^\alpha{}_\beta &= \frac{\partial y^\alpha}{\partial x^i} \frac{\partial x^j}{\partial y^\beta} \, G^i{}_j \end{aligned} \tag{9.15}$$

Equation (9.14) transforms as:

$$\frac{\partial y^\alpha}{\partial x^i} \frac{\partial x^j}{\partial y^\beta} \, G^i{}_j = \kappa \frac{\partial y^\alpha}{\partial x^i} \frac{\partial x^j}{\partial y^\beta} \, T^i{}_j \;\; \Rightarrow \;\; G^\alpha{}_\beta = \kappa \, T^\alpha{}_\beta \tag{9.16}$$

We see that in the new coordinates, equation (9.16) has the same form as it did in the old coordinates (9.14)!

Exercise IX.1

Use the transformation laws (9.4) and (9.6) to show that the directional derivative $v(f) = V^i \frac{\partial f}{\partial x^i}$ is invariant under coordinate transformations.

Exercise IX.2

Write out the transformation law for the components $S^{ij}{}_{klm}$ of the tensor S defined in (9.12).

---

[1] A more formal definition for the term "linear map" is given in exercise **IX.7**.

[2] By this I mean that if a tensor "eats" a vector, it is a linear function of the vector components if everything else it "eats" is held fixed. Another way to put it is that a tensor is linear in each of its arguments.

[3] The "eats zero" property is another way of saying that the tensor is a *homogeneous function* (of degree 1).

[4] $T(w, v)$ is linear in the sense that it is linear in each of its arguments.

Exercise IX.3

Show that the Kronecker delta (3.9) with one raised index and one lowered index, δ^i_j, retains its identity in all coordinate systems. In particular, show that[1] if I define the matrix $D^\alpha{}_\beta$ as the coordinate transformation of the Kronecker delta: $D^\alpha{}_\beta = \frac{\partial y^\alpha}{\partial x^i}\frac{\partial x^j}{\partial y^\beta}\delta^i_j$, then $D^\alpha{}_\beta = \delta^\alpha_\beta$. Note that this is not in general true for the lowered index Kronecker delta δ_{ij} or the raised index Kronecker delta δ^{ij}.

Exercise IX.4

Show that if $B^{ij}{}_{klm}$ form the components of a rank-5 tensor, then performing a contraction on any pair of indices yields a rank-3 tensor. In particular, convince yourself that $B^{ij}{}_{ilm}$ is a rank-3 tensor, $B^{ijk}{}_{km} = g^{kl}\,B^{ij}{}_{klm}$ is a rank-3 tensor, and so on. To show this formally, first perform a coordinate transformation on $B^{ij}{}_{klm}$, and then perform one contraction on the indices of the transformed components. The result should be the transformation law for a rank-3 tensor.

Exercise IX.5

The *tensor product* for tensors is simply the multiplication of tensor components to form a tensor of higher rank (recall the discussion of the tensor product in section VIII). Show that if Q^{ijk} and P_{ij} form the components of a tensor, then the tensor product $Q^{ijk}\,P_{lm}$ transforms as the components of a tensor.

Exercise IX.6

Show that if the tensor components S_{ij} are symmetric in i and j, meaning that $S_{ij} = S_{ji}$, then the transformed tensor components are also symmetric: $S_{\alpha\beta} = S_{\beta\alpha}$. Also show that if the tensor components A_{ij} are antisymmetric in i and j, meaning that $A_{ij} = -A_{ji}$, then the transformed tensor components are also antisymmetric: $A_{\alpha\beta} = -A_{\beta\alpha}$

[1] Hint: Recall the chain rule (9.3).

Exercise IX.7

In this exercise, I give a more formal definition for a linear map. A *linear map* R is defined by the following properties in each of its arguments:

$$\text{Additivity:} \qquad R(y+z) = R(y)+R(z)$$

$$\text{Homogeneity}^2 \text{ of degree 1:} \quad R(\alpha\, y) = \alpha\, R(y)$$

where y, z represent vectors or dual vectors, and α is a scalar. Show that vectors $v(\cdot)$ and dual vectors $w(\cdot)$ satisfy this property.

For an arbitrary number of arguments, a linear map R satisfies the following:

$$\begin{aligned} &R(\alpha\, u + \beta\, v, ..., \sigma\, w + \tau\, p, ...) \\ &= \alpha\, R(u, ..., \sigma\, w + \tau\, p, ...) + \beta\, R(v, ..., \sigma\, w + \tau\, p, ...) \\ &= \sigma\, R(\alpha\, u + \beta\, v, ..., w, ...) + \tau\, R(\alpha\, u + \beta\, v, ..., p, ...) \end{aligned}$$

where u, v are vectors, w, p are dual vectors, and α, β, σ, τ are scalars. Show that the tensor T in (9.10) and the tensor B in (9.12) both satisfy the above formulas.

[2] I add the phrase "degree 1" because in general Homogeneity refers to the property $f(\alpha\, y) = \alpha^k\, f(y)$ for some integer k (which is an *exponent*, not an index). In this case, we say $f(y)$ is Homogeneous of degree k.

X. TRANSFORMATIONS OF THE METRIC AND THE UNIT VECTOR BASIS

The metric is a tensor; you can find the expression for the components g_{ij} in a different set of coordinates using the tensor transformation law (9.8). In Cartesian coordinates on Euclidean space, the components of the metric are just the lowered index Kronecker delta δ_{ij}. In curvilinear coordinates, the components of the metric $g_{\alpha\beta}$ are:

$$g_{\alpha\beta} = \frac{\partial x^i}{\partial y^\alpha}\frac{\partial x^j}{\partial y^\beta}\,\delta_{ij} \tag{10.1}$$

Let's transform the metric to spherical coordinates[1] (r, θ, ϕ). In particular, I choose $y^1 = r$, $y^2 = \theta$, and $y^3 = \phi$. I write the original Cartesian coordinates x^i as functions of the spherical coordinates r, θ, and ϕ (or y^α):

$$\begin{aligned} x^1 &= r\,\cos\,\phi\,\sin\,\theta \\ x^2 &= r\,\sin\,\phi\,\sin\,\theta \\ x^3 &= r\,\cos\,\theta \end{aligned} \tag{10.2}$$

Using the formulas above (exercise **X.2**), you can show that the components of the metric tensor take the following form:

$$g_{\alpha\beta} = \begin{pmatrix} g_{11} & g_{12} & g_{13} \\ g_{21} & g_{22} & g_{23} \\ g_{31} & g_{32} & g_{33} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & r^2\,(\sin\theta)^2 \end{pmatrix} \tag{10.3}$$

The above formula (10.3) for the metric components $g_{\alpha\beta}$ is useful in two respects. First, it tells you how to measure distances along curves in spherical coordinates, via the line element (cf. equations (6.9) and (6.10)):

$$ds^2 = g_{\alpha\beta}\,dy^\alpha\,dy^\beta = dr^2 + r^2\,d\theta^2 + r^2\,(\sin\,\theta)^2\,d\phi^2 \tag{10.4}$$

where the superscripts after the *second* equality are *exponents*, not indices. Second, the metric components $g_{\alpha\beta}$ in (10.3) tells you how to take the dot product, given the components of vectors u^α and v^α in spherical coordinates:

$$u \cdot v = g_{\alpha\beta}\,u^\alpha\,v^\beta \tag{10.5}$$

However, there is a catch. If you have prior exposure to vectors in spherical coordinates, the vector components u^α and v^α in spherical coordinates may not be the same as those you are familiar with. This is because:

In general, coordinate basis vectors $\frac{\partial}{\partial x^i}$ are <u>*not*</u> unit vectors, and unit vectors are not in general coordinate basis vectors!

Recall that a *unit vector* is a vector of unit norm, meaning that a unit vector $\hat{u}$ satisfies the condition that $\hat{u}\cdot\hat{u} = 1$. Also recall that in section VI, the components of the metric g_{ij} are actually inner products (in this case, dot products) between the basis vectors e_i—in particular, recall equation (6.7) $g_{ij} = \langle e_i, e_j\rangle$—and also equation (6.8), which I rewrite here:

$$g_{ij} = \left\langle \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \right\rangle \tag{10.6}$$

In order for $\frac{\partial}{\partial y^\alpha}$ to be unit vectors, the diagonal elements of the metric, g_{11}, g_{22}, and g_{33} must all be equal to 1. This is clearly not the case for the metric in spherical coordinate (10.3); upon comparing (10.6) with (10.3), I find that:

$$\begin{aligned} \left\langle \frac{\partial}{\partial r}, \frac{\partial}{\partial r} \right\rangle &= g_{11} = 1 \\ \left\langle \frac{\partial}{\partial \theta}, \frac{\partial}{\partial \theta} \right\rangle &= g_{22} = r^2 \\ \left\langle \frac{\partial}{\partial \phi}, \frac{\partial}{\partial \phi} \right\rangle &= g_{33} = r^2\,(\sin\,\theta)^2 \end{aligned} \tag{10.7}$$

with all other inner products between the coordinate basis vectors vanishing. Only the basis element $\frac{\partial}{\partial r}$ is a unit vector.

Note that if the metric is diagonal, meaning that $g_{ij} = \langle e_i, e_j\rangle = 0$ if $i \neq j$, then the basis elements are orthogonal. If the basis vectors $\frac{\partial}{\partial y^\alpha}$ are orthogonal, then the norms (10.7) can be used to obtain expressions for the corresponding unit vectors in a straightforward manner; in spherical coordinates, simply divide by the square root of the norm to get the unit vectors:

$$\begin{aligned} \hat{r} &= \frac{1}{\sqrt{g_{11}}}\frac{\partial}{\partial y^1} = \frac{\partial}{\partial r} \\ \hat{\theta} &= \frac{1}{\sqrt{g_{22}}}\frac{\partial}{\partial y^2} = \frac{1}{r}\frac{\partial}{\partial \theta} \\ \hat{\phi} &= \frac{1}{\sqrt{g_{33}}}\frac{\partial}{\partial y^3} = \frac{1}{r\,\sin\,\theta}\frac{\partial}{\partial \phi} \end{aligned} \tag{10.8}$$

The above formulas allow you to convert vectors from the basis of partial derivatives (called the coordinate basis) and the basis of unit vectors in spherical coordinates.

Exercise X.1

For simplicity, let's work in two dimensions for this problem. Consider the coordinate functions relating Cartesian coordinates x^i to the polar coordinates $y^1 = r$ and $y^2 = \phi$:

$$\begin{aligned} x^1 &= r\,\cos\,\phi \\ x^2 &= r\,\sin\,\phi \end{aligned} \tag{10.9}$$

[1] I use the physicist's convention, where ϕ is the azimuthal angle that runs from 0 to 2π. Incidentally, I find it odd that so many of my physics colleagues still use the angle θ for polar coordinates in 2 dimensions (though I admit, I still do this on occasion); even more oddly, I've seen many of them switch to using ϕ when working in cylindrical coordinates!

Use the tensor transformation law (10.1) to obtain the components $g_{\alpha\beta}$ for the metric in polar coordinates. Can you infer the components of the inverse metric $g^{\alpha\beta}$?

Exercise X.2

Obtain the components $g_{\alpha\beta}$ for the metric in spherical coordinates from the coordinate functions (10.3) and the transformation law (10.1) for the metric. Use your result to infer the components of the inverse metric $g^{\alpha\beta}$.

Exercise X.3

In Cartesian coordinates, the simplest nonvanishing vector field you can write down is the constant vector field:

$$v^1 = 1$$
$$v^2 = 0$$
$$v^3 = 0$$

Lower the indices to get the components of the dual vector v_i. Obtain the components v_α for the dual vector field in spherical coordinates, and use the components of the inverse metric $g^{\alpha\beta}$ (see exercise **X.2**) to obtain the components for the vector field v^α in spherical coordinates. Finally, obtain the components for the vector in the unit-vector basis: $\hat{r}$, $\hat{\theta}$, $\hat{\phi}$.

Exercise X.4

Use the coordinate functions (10.9) in exercise **X.1** to infer the coordinate functions for cylindrical coordinates. Use the transformation law (10.1) to obtain an expression for the metric in cylindrical coordinates. Work out the norm for each basis vector $\frac{\partial}{\partial y^\alpha}$, and obtain expressions for the corresponding unit vectors.

XI. DERIVATIVES OF TENSORS

Ultimately, the goal here is to construct a formalism for writing down Partial Differential Equations (PDEs) in a manner that is transparent[1] for doing calculations, but also coordinate-invariant. Often, we wish to express PDEs in coordinates adapted to a particular problem, especially when symmetry is involved. For instance, when using a PDE to model a spherically symmetric system, it is far more appropriate to express the PDE in spherical coordinates than cylindrical coordinates.

By definition, PDEs contain partial derivatives. Unfortunately, it turns out that partial derivatives of tensors do not transform like tensors. The purpose of this section is to develop a formalism for the derivatives that preserves the transformation properties of tensors.

The partial derivative of a scalar function $f(x^a)$, the gradient $\mathcal{G}_i = \frac{\partial f}{\partial x^i}$, transforms as a tensor—a dual vector in fact. Actually, if you recall the logic in section IX, it is the other way around—these notes derive the transformation properties of the dual vector w_i (and *all* the transformation properties of tensors) from the transformation of the gradient components $\mathcal{G}_i$. On the other hand, the partial derivative of the components v^i of a vector field, do not transform as a tensor. Consider the following quantity:

$$A^i{}_j := \frac{\partial v^i}{\partial x^j} \tag{11.1}$$

and its corresponding expression in the coordinates y^α:

$$A^\alpha{}_\beta := \frac{\partial v^\alpha}{\partial y^\beta} \tag{11.2}$$

If $A^i{}_j$ and $A^\alpha{}_\beta$ are components of a tensor, they would be related to each other by a tensor transformation law (9.11), which I rewrite here for $T^i{}_j$ and $T^\alpha{}_\beta$:

$$T^\alpha{}_\beta = \frac{\partial y^\alpha}{\partial x^i}\frac{\partial x^j}{\partial y^\beta}\, T^i{}_j \tag{11.3}$$

Unfortunately, $A^i{}_j$ and $A^\alpha{}_\beta$ *do not* satisfy a tensor transformation law of the form (11.3). To see this, I insert the formula $v^\alpha = \frac{\partial y^\alpha}{\partial x^i}\, v^i$ into equation (11.2) for $A^\alpha{}_\beta$:

$$A^\alpha{}_\beta = \frac{\partial}{\partial y^\beta}\left(\frac{\partial y^\alpha}{\partial x^i}\, v^i\right) = \frac{\partial x^j}{\partial y^\beta}\frac{\partial}{\partial x^j}\left(\frac{\partial y^\alpha}{\partial x^i}\, v^i\right) \tag{11.4}$$

---

[1] This is the disadvantage of the standard Gibbs-Heaviside formalism for vector analysis that you may be accustomed to (see the article [6] for a historical discussion of the Gibbs-Heaviside formalism). Abstract expressions such as $\nabla \cdot v$ and $\nabla \times v$ are in some sense coordinate-invariant, but they give little indication of the methods for computing them in an arbitrary set of curvilinear coordinates. You need to first *specify* the meaning of $\nabla \cdot v$ and $\nabla \times v$ in some "natural" coordinate system (Cartesian coordinates), then derive expressions for $\nabla \cdot v$ and $\nabla \times v$ in another coordinate system by way of coordinate transformations.

where I have applied the chain rule in the second equality. Upon applying the product rule and recognizing a factor of $A^i{}_j$, I obtain the result:

$$A^\alpha{}_\beta = \frac{\partial x^j}{\partial y^\beta}\frac{\partial y^\alpha}{\partial x^i}\,A^i{}_j + \underline{\frac{\partial x^j}{\partial y^\beta}\frac{\partial^2 y^\alpha}{\partial x^j \partial x^i}\,v^i} \tag{11.5}$$

If I cover up the underlined term with my hand, the equation above (11.5) looks like the tensor transformation law. However, the presence of the underlined term means that the quantities $A^i{}_j$ and $A^\alpha{}_\beta$ *do not* form the components of a tensor. If we can somehow get rid of the underlined term in (11.5) without placing constraints[1] on the functions $y^\alpha(x^a)$, then we could recover the transformation law for a tensor.

One way to do this is to construct a new derivative operator that reduces to the usual partial derivative in Cartesian coordinates. The simplest modification to the partial derivative $\frac{\partial v^i}{\partial x^j}$ (11.2) is to add a correction term. Since I wish to cancel out the underlined term in (11.5), and since the underlined term contains a factor of v^i, the correction term should contain a factor of v^i. These considerations lead me to define a new derivative operator ∇_j, called the *covariant derivative*, which acts on v^i in the following way:

$$\nabla_j v^i = \frac{\partial v^i}{\partial x^j} + \Gamma^i_{jk}\,v^k \tag{11.6}$$

where Γ^i_{jk} are coefficients, which are sometimes called *connection coefficients*. The trick here is that the coefficients Γ^i_{jk} *do not* transform as tensors. Instead, I demand that the coefficients Γ^i_{jk} satisfy the following transformation law:

$$\Gamma^\alpha_{\beta\gamma} = \left(\frac{\partial y^\alpha}{\partial x^i}\frac{\partial x^j}{\partial y^\beta}\frac{\partial x^k}{\partial y^\gamma}\right)\Gamma^i_{jk} - \left(\frac{\partial x^j}{\partial y^\beta}\frac{\partial^2 y^\alpha}{\partial x^j \partial x^i}\right)\frac{\partial x^i}{\partial y^\gamma} \tag{11.7}$$

Note that the bracketed quantity in the second term appears in the underlined term in (11.5). If we demand that ∇_i reduces to the ordinary partial derivative in Cartesian coordinates, then in Cartesian coordinates, we set $\Gamma^i_{jk} = 0$. It follows that the second term in (11.7) can be used to compute the connection coefficients $\Gamma^\alpha_{\beta\gamma}$ in any other coordinate system.

If the coefficients Γ^i_{jk} transform according to the above transformation law (11.7), then it is not difficult to show (see exercise **XI.1**) that the quantity $\nabla_j v^i$ (11.6) transforms as a tensor:

$$\nabla_\beta v^\alpha = \frac{\partial y^\alpha}{\partial x^i}\frac{\partial x^j}{\partial y^\beta}\nabla_j v^i \tag{11.8}$$

where $\nabla_\beta v^\alpha$ is given by:

$$\nabla_\beta v^\alpha = \frac{\partial v^\alpha}{\partial y^\beta} + \Gamma^\alpha_{\beta\gamma}\,v^\gamma \tag{11.9}$$

Equation (11.8) states that the covariant derivative operator ∇_i yields a rank-2 tensor when acting on a vector—success!

The next thing to do is to construct a covariant derivative for tensors of higher rank. Note that for a rank-2 tensor G^{ij} with raised indices each index will pick up a factor of transformation matrix $\frac{\partial y^\alpha}{\partial x^i}$ under a coordinate transformation. Each factor of $\frac{\partial y^\alpha}{\partial x^i}$ will generate an extra term which needs to be canceled out (see exercise **XI.2**). The covariant derivative for G^{ij} takes the form:

$$\nabla_k G^{ij} = \frac{\partial G^{ij}}{\partial x^k} + \Gamma^i_{km}\,G^{mj} + \Gamma^j_{km}\,G^{im} \tag{11.10}$$

For a rank-3 tensor Q^{ijl}, I pick up yet another term:

$$\nabla_k Q^{ijl} = \frac{\partial Q^{ijl}}{\partial x^k} + \Gamma^i_{km}\,Q^{mjl} + \Gamma^j_{km}\,Q^{iml} + \Gamma^l_{km}\,Q^{ijm} \tag{11.11}$$

I can continue to higher rank, but I think you see the pattern; when taking the covariant derivative, each index requires a term containing a factor of Γ^i_{jk} and the original tensor components.

I'll take the opportunity to quickly introduce the following notation for the gradient of a scalar function:

$$\nabla_k f = \frac{\partial f}{\partial x^k}, \tag{11.12}$$

which follows from running the pattern of equations (11.6), (11.10) and (11.11) in reverse—a scalar function f has zero indices, so no extra correction term is needed. In case you don't recognize the pattern, I'll summarize the covariant derivative for f, v^i, G^{ij} and Q^{ijl}:

$$\begin{aligned}
\nabla_k f &= \frac{\partial f}{\partial x^k}\\
\nabla_k v^i &= \frac{\partial v^i}{\partial x^k} + \Gamma^i_{km}\,v^m\\
\nabla_k G^{ij} &= \frac{\partial G^{ij}}{\partial x^k} + \Gamma^i_{km}\,G^{mj} + \Gamma^j_{km}\,G^{im}\\
\nabla_k Q^{ijl} &= \frac{\partial Q^{ijl}}{\partial x^k} + \Gamma^i_{km}\,Q^{mjl} + \Gamma^j_{km}\,Q^{iml} + \Gamma^l_{km}\,Q^{ijm}
\end{aligned} \tag{11.13}$$

As an exercise (see exercise **XI.3**, try to construct the covariant derivative for a rank-4 tensor.

I now wish to construct covariant derivatives for lower indexed objects. I'll begin by constructing the derivative for a dual vector, which I expect to have the form:

$$\nabla_k w_i = \frac{\partial w_i}{\partial x^k} + C^j_{ki}\,w_j \tag{11.14}$$

Recall that the quantity $v^i\,w_i$ is a scalar. From (11.12), the covariant derivative ∇_k acting on a scalar is just the

[1] If the second derivative of $y^\alpha(x^a)$ vanishes, then the underlined term in (11.5) vanishes, and we recover the tensor transformation law. However, if x^i are Cartesian coordinates, this is also the condition that the coordinate lines are straight; the only admissible coordinate transformations correspond to rigid rotations and translations.

partial derivative, so:

$$\nabla_k(v^i\,w_i) = \frac{\partial(v^i\,w_i)}{\partial x^k} = w_i\,\frac{\partial v^i}{\partial x^k} + v^i\,\frac{\partial w_i}{\partial x^k} \tag{11.15}$$

where the product rule has been used in the second equality. Now a good definition for a derivative operator should respect the product (Leibniz) rule, so I demand that the covariant derivative for dual vectors be consistent with the following property:

$$\nabla_k\left(v^i\,w_i\right) = w_i\,\nabla_k v^i + v^i\,\nabla_k w_i \tag{11.16}$$

I now have two equations, (11.15) and (11.16), for the same quantity $\nabla_k\left(v^i\,w_i\right)$. If I expand (11.16) and subtract (11.15) from (11.16), I obtain the following result:

$$\begin{aligned}\nabla_k\left(v^i\,w_i\right) - \nabla_k\left(v^i\,w_i\right) &= w_i\,\Gamma^i_{kj}\,v^j + v^i\,C^j_{ki}\,w_j \\ &= w_i\,\Gamma^i_{kj}\,v^j + v^j\,C^i_{kj}\,w_i = 0\end{aligned} \tag{11.17}$$

Where I have performed a relabeling of dummy indices in the last term (recall the discussion in section VIII on the relabeling of dummy indices). If I demand that equation (11.17) holds for *all* v^i and w_i, the coefficients C^i_{kj} must satisfy $C^i_{kj} = -\Gamma^i_{kj}$, so that the covariant derivative for dual vectors is given by:

$$\nabla_k w_i = \frac{\partial w_i}{\partial x^k} - \Gamma^j_{ki}\,w_j \tag{11.18}$$

Given the pattern (11.13) and the covariant derivative (11.18) for dual vectors, I may infer that for a rank-2 tensor K_{ij} with lowered indices, the covariant derivative is:

$$\nabla_k K_{ij} = \frac{\partial K_{ij}}{\partial x^k} - \Gamma^a_{ki}\,K_{aj} - \Gamma^a_{kj}\,K_{ia} \tag{11.19}$$

I may also infer that for a rank-2 tensor $T^i{}_j$ with mixed indices, the covariant derivative is:

$$\nabla_k T^i{}_j = \frac{\partial T^i{}_j}{\partial x^k} + \Gamma^i_{ka}\,T^a{}_j - \Gamma^a_{kj}\,T^i{}_a \tag{11.20}$$

Given (11.19), (11.20), and the "tower" of equations in (11.13), you should be able to infer the covariant derivative of a tensor of arbitrary rank. If you want a more careful justification of (11.19) and (11.20), see exercise **XI.4**.

At this point, I can state the punchline and end this section. However, I wish to present a useful formula for Γ^i_{jk} in terms of the metric tensor:

$$\Gamma^i_{jk} = \frac{1}{2}g^{ia}\left(\frac{\partial g_{ja}}{\partial x^k} + \frac{\partial g_{ak}}{\partial x^j} - \frac{\partial g_{jk}}{\partial x^a}\right) \tag{11.21}$$

In this form, the coefficients Γ^i_{jk} are called *Christoffel symbols*. You can derive the formula (11.21) for Christoffel symbols from the following two conditions[1] on the covariant derivative ∇_i:

1. The Torsion Free Condition: *For a scalar function f, ∇_i satisfies:*

$$\nabla_i\nabla_j f = \nabla_j\nabla_i f \qquad \Rightarrow \qquad \Gamma^k_{ij} = \Gamma^k_{ji} \tag{11.22}$$

2. Metric Compatibility:

$$\nabla_k g_{ij} = 0 \tag{11.23}$$

Since equations (11.22) and (11.23) are constructed from covariant derivatives, they transform as tensors, and have the same form in all coordinates. It is not too hard to verify that both (11.22) and (11.23) are true in Cartesian coordinates, provided that we demand that $\Gamma^i_{jk} = 0$ in Cartesian coordinates. Thus, instead of computing Γ^i_{jk} via the transformation law (11.7) in curvilinear coordinates, we may first work out an expression for the metric in curvilinear coordinates, then use (11.21) to obtain an expression for Γ^i_{jk}.

Now that I have shown you how to take the covariant derivative ∇_i for a tensor, I can deliver the punchline for this section:

> **The covariant derivative of a tensor transforms as a tensor. This is useful because it allows us to write derivatives in a form that looks the same in all coordinate systems.**

The formulas for the covariant derivatives of a tensor are given in (11.19), (11.20), and (11.13). Given a metric g_{ij}, the coefficients Γ^i_{jk} can be obtained from equation (11.21). Alternately, one may demand that $\Gamma^i_{jk} = 0$ and use the transformation law (11.7) to obtain an expression for Γ^i_{jk} in an arbitrary curvilinear coordinate system.

Since covariant derivatives become partial derivatives in Cartesian coordinates, all you need to do to convert a PDE to a tensor equation is to write down the PDE in Cartesian coordinates, and replace partial derivatives with covariant derivatives:[2]

$$\frac{\partial}{\partial x^k} \quad \rightarrow \quad \nabla_k \tag{11.24}$$

---

[1] Some terminology: Sometimes, a covariant derivative ∇_i is called a *connection*, and Γ^i_{jk} are called *connection coefficients*. The specific covariant derivative defined by the properties (11.22) and (11.23) is called the *Levi-Civita connection* (this is not to be confused with the Levi-Civita symbol and Levi-Civita pseudotensor, which are entirely different quantities that I will introduce later on).

[2] This is sometimes called the "comma-to-semicolon rule." This term refers to a commonly used shorthand notation for partial derivatives and covariant derivatives, in which partial derivatives of tensors are denoted by placing the derivative index after a comma: $T^i{}_j{}_{,k} = \frac{\partial T^i{}_j}{\partial x^k}$, and covariant derivatives are similarly expressed, this time with a semicolon: $T^i{}_j{}_{;k} = \nabla_k T^i{}_j$. The comma-to-semicolon rule refers to the replacement of partial derivatives (commas) with covariant derivatives (semicolons). For further discussion of this rule and the subtleties of applying it in non-Euclidean geometries, refer to Chapter 4 of [31] and Box 16.1 of [18].

Curvature

In these notes, I have avoided the discussion of non-Euclidean geometries, lest I be accused of doing General Relativity (the horror!). However, I suspect that you may nonetheless be curious about the applications of this formalism to non-Euclidean geometries. In particular, this formalism may be used to describe curved 2d surfaces, but it does so in a way that can be generalized to n-dimensional manifolds. Specifically, the intrinsic geometric properties of a curved 2d surface is determined by the metric, from which one can define distances (using the line element) and angles (using the inner product). A metric for a curved 2d surface differs from that of a flat metric in that there exists no coordinate system such that the metric reduces to the Kronecker delta δ_{ij} *everywhere* on the surface. The metric is important for studying the intrinsic geometry of curved 2d surfaces and their higher-dimensional generalizations, but the covariant derivative (called the connection) is even more important, because you can use it to construct a *tensor* that corresponds to the curvature of the surface (or a manifold), which is a measure of the degree to which a surface/manifold fails to satisfy the geometric properties of Euclidean space. The tensor that measures the intrinsic curvature of a surface/manifold is called the Riemann curvature tensor $R^i{}_{jkl}$, which is formed from the connection coefficients $\Gamma^i{}_{jk}$ and their partial derivatives—the definition and formula for $R^i{}_{jkl}$ can be found in any textbook on General Relativity (see for instance [5], [31], and [18]).

Exercise XI.1

Show that if the coefficients Γ^i_{jk} satisfy the transformation law (11.7), the covariant derivatives $\nabla_j v^i$ and $\nabla_\beta v^\alpha$ (as defined in (11.6) and (11.9)) satisfy (11.8):

$$\nabla_\beta v^\alpha = \frac{\partial y^\alpha}{\partial x^i}\frac{\partial x^j}{\partial y^\beta}\nabla_j v^i$$

Exercise XI.2

Given the tensor transformation law for G^{ij}:

$$G^{\alpha\beta} = \frac{\partial y^\alpha}{\partial x^i}\frac{\partial y^\beta}{\partial x^j}\,G^{ij}, \tag{11.25}$$

show that the partial derivatives $\frac{\partial G^{ij}}{\partial x^k}$ and $\frac{\partial G^{\alpha\beta}}{\partial y^\gamma}$ satisfies the following expression:

$$\begin{aligned}\frac{\partial G^{\alpha\beta}}{\partial y^\gamma} =& \frac{\partial x^k}{\partial y^\gamma}\frac{\partial y^\alpha}{\partial x^i}\frac{\partial y^\beta}{\partial x^j}\frac{\partial G^{ij}}{\partial x^k} \\ &+ \frac{\partial y^\beta}{\partial x^j}\frac{\partial x^k}{\partial y^\gamma}\frac{\partial^2 y^\alpha}{\partial x^k \partial x^i}\,G^{ij} \\ &+ \frac{\partial y^\alpha}{\partial x^i}\frac{\partial x^k}{\partial y^\gamma}\frac{\partial^2 y^\beta}{\partial x^k \partial x^j}\,G^{ij}.\end{aligned} \tag{11.26}$$

Use (11.26) to show that the covariant derivative $\nabla_k G^{ij}$ as defined in (11.10) transforms as a tensor, or that:

$$\nabla_\gamma G^{\alpha\beta} = \frac{\partial x^k}{\partial y^\gamma}\frac{\partial y^\alpha}{\partial x^i}\frac{\partial y^\beta}{\partial x^j}\,\nabla_k G^{ij}, \tag{11.27}$$

with $\nabla_\gamma G^{\alpha\beta}$ being given by the expression: $\nabla_\gamma G^{\alpha\beta} = \frac{\partial G^{\alpha\beta}}{\partial y^\gamma} + \Gamma^\alpha_{\gamma\sigma}\,G^{\sigma\beta} + \Gamma^\beta_{\gamma\sigma}\,G^{\alpha\sigma}$.

Exercise XI.3

Following the pattern in (11.13), write out the expression for the covariant derivative of the rank-4 tensor R^{ijkl}.

Exercise XI.4

In this exercise, I will walk you through a more careful justification of equations (11.19) and (11.20). Begin by assuming that the covariant derivatives of $T^i{}_j$ and K_{ij} have the form:

$$\begin{aligned}\nabla_k K_{ij} &= \frac{\partial K_{ij}}{\partial x^k} + A^a_{ki}\,K_{aj} + A^a_{kj}\,K_{ia} \\ \nabla_k T^i{}_j &= \frac{\partial T^i{}_j}{\partial x^k} + P^i_{ka}\,T^a{}_j + Q^a_{kj}\,T^i{}_a\end{aligned} \tag{11.28}$$

where A^a_{ki}, P^i_{ka}, and Q^a_{kj} are coefficients to be determined. Now consider the scalars $u^i\,v^j\,K_{ij}$ and $w_i\,v^j\,T^i{}_j$, where u^i, v^i are arbitrary vectors, and w_i is an arbitrary dual vector. Note that:

$$\begin{aligned}\nabla_k\left(u^i\,v^j\,K_{ij}\right) &= \frac{\partial\left(u^i\,v^j\,K_{ij}\right)}{\partial x^k} \\ \nabla_k\left(w_i\,v^j\,T^i{}_j\right) &= \frac{\partial\left(w_i\,v^j\,T^i{}_j\right)}{\partial x^k}\end{aligned} \tag{11.29}$$

Expand the right hand side, using the product rule. Demand that the covariant product rule holds:

$$\begin{aligned}\nabla_k\left(u^i\,v^j\,K_{ij}\right) =& v^j\,K_{ij}\,\nabla_k u^i + u^i\,K_{ij}\,\nabla_k v^j \\ &+ u^i\,v^j\,\nabla_k K_{ij} \\ \nabla_k\left(w_i\,v^j\,T^i{}_j\right) =& v^j\,T^i{}_j\,\nabla_k w_i + w_i\,T^i{}_j\,\nabla_k v^j \\ &+ w_i\,v^j\,\nabla_k T^i{}_j\end{aligned} \tag{11.30}$$

and expand out the covariant derivatives of the vectors u^i, v^i and dual vector w_i using equations (11.6) and (11.18). Subtract equations (11.29) from (11.30), and demand that the result holds for all choices of u^i, v^i, w_i, $T^i{}_j$ and K_{ij}. From this demand, you may infer that $A^a_{ki} = -\Gamma^a_{ki}$, $P^i_{ka} = \Gamma^i_{ka}$, and $Q^a_{kj} = -\Gamma^a_{kj}$.

Exercise XI.5

In order for the operator ∇_i to be properly regarded as a derivative operator for the vector components u^i and v^i, it should satisfy the following two properties for the vector components u^i and v^i:

1. Additivity:

$$\nabla_i \left(u^j + v^j\right) = \nabla_i u^j + \nabla_i v^j \quad (11.31)$$

2. Product Rule (Leibniz Rule):

$$\nabla_i \left(u^j\, v^k\right) = v^k\, \nabla_i u^j + u^j\, \nabla_i v^k \quad (11.32)$$

Show that the covariant derivative satisfies these properties (for the product rule, apply the definition (11.10) to the product $G^{jk} = u^j\, v^k$). Write down the corresponding properties for a pair of dual vectors w_i and p_i, and show that ∇_i satisfies those properties for dual vectors. Finally, check that the Leibniz rule holds for a vector v^i and a dual vector w_i:

$$\nabla_i \left(v^j\, w_k\right) = w_k\, \nabla_i v^j + v^j\, \nabla_i w_k \quad (11.33)$$

Exercise XI.6

Derive (11.21) from equations (11.22) and (11.23). To do this, use (11.23) to obtain the expression:

$$\nabla_i g_{kj} + \nabla_j g_{ik} - \nabla_k g_{ij} = 0 \quad (11.34)$$

Use (11.22) to simplify the result.

Exercise XI.7

Use equation (11.21) and with the metric you obtained in exercise **X.4** to derive the coefficients Γ^i_{jk} in cylindrical coordinates. You should only have two unique[1] nonzero coefficients.

Exercise XI.8

Use equation (11.21) and with equation (10.3) to derive the coefficients Γ^i_{jk} in spherical coordinates.

Exercise XI.9

Extend the properties (11.31) and (11.32) to tensors of arbitrary rank, and show that the covariant derivative operator ∇_i satisfies these properties. For instance, given the tensors $T^i{}_j$, $K^i{}_j$ and $Q_i{}^{jk}$, $P_i{}^{jk}$, show that the covariant derivative satisfies the following properties:

1. Additivity:

$$\nabla_i \left(T^j{}_k + K^j{}_k\right) = \nabla_i T^j{}_k + \nabla_i K^j{}_k \quad (11.35)$$

$$\nabla_i \left(Q_j{}^{kl} + P_j{}^{kl}\right) = \nabla_i Q_j{}^{kl} + \nabla_i P_j{}^{kl} \quad (11.36)$$

2. Product Rule (Leibniz Rule):

$$\nabla_i \left(T^j{}_k\, Q_l{}^{mn}\right) = Q_l{}^{mn}\, \nabla_i T^j{}_k + T^j{}_k\, \nabla_i Q_l{}^{mn} \quad (11.37)$$

and use this result to convince yourself that ∇_i satisfies similar properties for tensors of arbitrary rank.

Exercise XI.10

Show or argue[2] that:

$$\nabla_k g^{ij} = 0. \quad (11.38)$$

Note that the above expression differs from (11.23) in that it involves the covariant derivative of the *inverse* metric instead of the metric.

[1] There are a total of three, but two of the coefficients are equivalent due to the symmetry in the lowered indices of Γ^i_{jk}.

[2] Hint: Note that $\nabla_k g^{ij}$ is a tensor. What is $\nabla_k g^{ij}$ in Cartesian coordinates?

XII. DIVERGENCES, LAPLACIANS AND MORE

PDEs are often expressed in terms of divergences and Laplacians. In this short section, I describe their construction in terms of covariant derivatives ∇_i. In the standard Gibbs-Heaviside notation for vector analysis, the divergence of a vector field v is written as the following:

$$\vec{\nabla} \cdot \vec{v} \tag{12.1}$$

If I take equation (12.1) literally, $\vec{\nabla}$ is a vector operator, with "components" ∇^i given by:

$$\nabla^i := g^{ik}\, \nabla_k \tag{12.2}$$

Now if I interpret the dot product as an inner product, I may write:

$$\vec{\nabla} \cdot \vec{v} = g_{ij}\, \nabla^i v^j = g_{ij}\, g^{ik}\, \nabla_k v^j = \delta^k_j\, \nabla_k v^j \tag{12.3}$$

which leads to the following expression for the Divergence of a vector field:

$$\vec{\nabla} \cdot \vec{v} = \nabla_i\, v^i = \frac{\partial v^i}{\partial x^i} + \Gamma^i_{ik}\, v^k \tag{12.4}$$

Suppose that you wish to take the divergence of a matrix field; in other words you wish to write something to the effect: $\vec{\nabla} \cdot \mathbf{M}$ for some 3×3 matrix field $\mathbf{M}$. An expression of the form $\vec{\nabla} \cdot \mathbf{M}$ requires a good deal of explanation, but if the matrix $\mathbf{M}$ is a rank-2 tensor with components $M^i{}_j$, it has a straightforward tensor expression:

$$\vec{\nabla} \cdot \mathbf{M} = \nabla_i\, M^i{}_j = \frac{\partial M^i{}_j}{\partial x^i} + \Gamma^i_{ik}\, M^k{}_j - \Gamma^k_{ij}\, M^i{}_k \tag{12.5}$$

This example illustrates the power of the tensor formalism; expressions that are difficult to interpret in the standard Gibbs-Heaviside notation become straightforward when written in tensor notation. Conversely, it is not too difficult to imagine operations that are straightforward in tensor notation, but become extremely cumbersome to express in Gibbs-Heaviside notation—consider, for instance, the following operator, which I call the "matrix Laplacian":

$$M^i{}_j \nabla_i\, \nabla^j \tag{12.6}$$

Another example is the following quantity:

$$(\vec{v} \cdot \vec{\nabla})\vec{u} \tag{12.7}$$

which can be thought of as the "directional derivative" of a vector $\vec{u}$. This quantity caused me great confusion when I first saw it, because it was written as $\vec{v} \cdot \vec{\nabla} \vec{u}$; in that form, I didn't know whether to apply the dot product to $\vec{\nabla}$ or $\vec{u}$, so in my mind, there were two meanings for $\vec{v} \cdot \vec{\nabla} \vec{u}$:

$$g_{kj}\, v^k\, \nabla^i u^j \tag{12.8}$$

$$g_{kj}\, v^k\, \nabla^j u^i \tag{12.9}$$

When brackets are placed around $\vec{v} \cdot \vec{\nabla}$, as in (12.7), then it becomes clear that one is referring to second expression (12.9). I don't know of any straightforward way to express the quantity (12.8) in Gibbs-Heaviside notation—this is another limitation of that formalism.

Now I turn to the Laplacian acting on a scalar function $\phi(x^a)$ is given by the following expression in Gibbs-Heaviside notation (the "2" on the left hand side is an *exponent*, not an index):

$$\nabla^2 \phi = \vec{\nabla} \cdot \vec{\nabla} \phi \tag{12.10}$$

Using (12.2), I may write the Laplacian operator as:

$$\nabla^2 := \vec{\nabla} \cdot \vec{\nabla} = g_{ij}\, \nabla^i \nabla^j \tag{12.11}$$

If the covariant derivative ∇_i satisfies the condition of metric compatibility (11.23,11.38) (it should, if we demand that $\Gamma^i_{jk} = 0$ in Cartesian coordinates), then one can use the product rule to obtain the result for the operator:

$$\nabla^2 = g^{ij}\, \nabla_i \nabla_j \tag{12.12}$$

For a scalar function ϕ, (12.10) becomes:

$$\nabla^2 \phi = g^{ij}\, \nabla_i \nabla_j \phi = g^{ij} \left(\frac{\partial^2 \phi}{\partial x^i \partial x^j} - \Gamma^k_{ij}\, \frac{\partial \phi}{\partial x^j} \right) \tag{12.13}$$

The Laplacian operator (12.12) constructed from covariant derivatives ∇_i may be used to obtain the action of the Laplacian on vectors:

$$\begin{aligned} \nabla^2 v^k = g^{ij}\, \nabla_i \nabla_j v^k &= g^{ij}\, \nabla_i \left(\frac{\partial v^k}{\partial x^j} + \Gamma^k_{ja}\, v^a \right) \\ &= g^{ij} \left[\frac{\partial^2 v^k}{\partial x^i \partial x^j} + \frac{\partial \Gamma^k_{ja}}{\partial x^i}\, v^a + \Gamma^k_{ja}\, \frac{\partial v^a}{\partial x^i} \right. \\ &\qquad \left. + \Gamma^k_{ib} \left(\frac{\partial v^b}{\partial x^j} + \Gamma^b_{ja}\, v^a \right) - \Gamma^b_{ij} \left(\frac{\partial v^k}{\partial x^b} + \Gamma^k_{ba}\, v^a \right) \right] \end{aligned} \tag{12.14}$$

Since we know how covariant derivatives act on tensors of arbitrary rank, you could in principle construct the explicit formula for the Laplacian for a tensor of arbitrary rank.

Exercise XII.1

Obtain the components of $(\vec{v} \cdot \vec{\nabla})\vec{u}$ (12.7) (in tensor form, $v^k\, \nabla_k u^i$ (12.9)) in cylindrical coordinates (recall exercise **X.4** and **XI.7**).

Exercise XII.2

The Schrödinger equation for a single particle of mass m in nonrelativistic Quantum Mechanics has the form:

$$i\,\hbar\,\frac{\partial\Psi}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\Psi + V(x^a)\,\Psi$$

where t is time, Ψ and $V(x^a)$ are scalar functions, $\hbar$ the reduced Planck constant, and $i = \sqrt{-1}$. Rewrite the Schrödinger equation in tensor form, then write down the explicit expression for the Schrödinger equation in spherical coordinates (recall exercises **X.2** and **XI.8**). We're not doing relativity, so don't transform $\partial/\partial t$; time t is an absolute parameter here.

Exercise XII.3

The nonrelativistic Euler equation for an inviscid fluid takes the form:

$$\rho\,\frac{\partial\vec{v}}{\partial t} + \rho\,(\vec{v}\cdot\vec{\nabla})\vec{v} = -\vec{\nabla}P + \frac{1}{\rho}\,\vec{f}$$

where t is time, $\vec{v}$ is the velocity of a fluid, ρ is the fluid density, P is the fluid pressure, and $\vec{f}$ is an external force density (force per unit mass). Rewrite the Euler equation in tensor form, then write down the explicit expression for the Euler equation in cylindrical coordinates (again recall exercises **X.4** and **XI.7**). Assume that the coordinate transformations do *not* depend on t, so that $\partial\vec{v}/\partial t$ transforms just like the vector $\vec{v}$.

Exercise XII.4

Write out the components for the Laplacian of a *vector* $\nabla^2\vec{v}$ in spherical coordinates.

Exercise XII.5

Write out an explicit expression for the "matrix Laplacian" of (12.6) acting on a scalar function:

$$M^i{}_j\nabla_i\,\nabla^j\phi \tag{12.15}$$

in terms of ϕ, g_{ij}, g^{ij} and Γ^i_{jk}.

Exercise XII.6

Given the expression (12.11), work out an explicit expression for the Laplacian of a rank-2 tensor field:

$$\nabla^2 T^i{}_j \tag{12.16}$$

in terms of $T^i{}_j$, g_{ij}, g^{ij} and Γ^i_{jk}.

XIII. THE LEVI-CIVITA TENSOR: CROSS PRODUCTS, CURLS, AND VOLUME INTEGRALS

Many PDEs contain expressions that involve cross products and curls. In the standard Gibbs-Heaviside notation, the cross product and curl respectively take the form:

$$\vec{A}\times\vec{B} \tag{13.1}$$

$$\vec{\nabla}\times\vec{A} \tag{13.2}$$

for vectors $\vec{A}$ and $\vec{B}$. In both cases, the result is a vector.

To express (13.1) and (13.2) in component form, I need to define the following quantity[1] $\underline{\epsilon}_{ijk}$, called the *permutation symbol*:[2]

$$\underline{\epsilon}_{ijk} = \underline{\epsilon}^{ijk} = \begin{cases} 1 & \text{if } \{i,j,k\} \text{ is an even permutation of } \{1,2,3\} \\ -1 & \text{if } \{i,j,k\} \text{ is an odd permutation of } \{1,2,3\} \\ 0 & \text{if any two indices are equal} \end{cases} \tag{13.3}$$

Explicitly, this means that the nonzero components of $\underline{\epsilon}_{ijk}$ are:

$$\begin{aligned} \underline{\epsilon}_{123} = \underline{\epsilon}_{312} = \underline{\epsilon}_{231} = 1 \\ \underline{\epsilon}_{132} = \underline{\epsilon}_{321} = \underline{\epsilon}_{213} = -1 \end{aligned} \tag{13.4}$$

I must point out something important:

> **The permutation symbol $\underline{\epsilon}_{ijk}$ does NOT form the components of a tensor!**

The reason for this is simple: the values of $\underline{\epsilon}_{ijk}$ do not depend on the coordinate system. Alternately, $\underline{\epsilon}_{ijk}$ has the same set of values in *all* coordinate systems. More pointedly, $\underline{\epsilon}_{\alpha\beta\gamma}$ has values defined by the same equation (13.4) as $\underline{\epsilon}_{ijk}$. I'll say more about this in a bit.

In Cartesian coordinates, the components of $\vec{A}\times\vec{B}$ and $\vec{\nabla}\times\vec{A}$, which I respectively write as $[\vec{A}\times\vec{B}]^i$ and $[\vec{\nabla}\times\vec{A}]^i$, are given by the following formulas:

$$[\vec{A}\times\vec{B}]^i = g^{ia}\,\underline{\epsilon}_{ajk}\,A^j\,B^k = \delta^{ia}\,\underline{\epsilon}_{ajk}\,A^j\,B^k \tag{13.5}$$

$$[\vec{\nabla}\times\vec{A}]^i = g^{ia}\,\underline{\epsilon}_{ajk}\,\nabla^j A^k = \delta^{ia}\,\delta^{jb}\,\underline{\epsilon}_{ajk}\,\frac{\partial A^k}{\partial x^b} \tag{13.6}$$

Where I have set $g^{ij} = \delta^{ij}$ and $\nabla_i = \frac{\partial}{\partial x^i}$. To be clear: **Equations (13.5) and (13.6) are NOT tensor equations!** Equations (13.5) and (13.6) are ONLY valid in

---

[1] Note the underline!

[2] The permutation symbol $\underline{\epsilon}_{ijk}$ is also called the *Levi-Civita symbol*, not to be confused with the Levi-Civita connection (see footnote 1 on page 24), which is an entirely different concept.

Cartesian coordinates. This is because, as stated earlier in bold, $\underline{\epsilon}_{ijk}$ is not a tensor.

To see that $\underline{\epsilon}_{ijk}$ is not a tensor, I'll feed it the components of three vectors, $\vec{A}$, $\vec{B}$ and $\vec{C}$, to obtain the following expression:

$$\vec{A}\cdot(\vec{B}\times\vec{C}) = \underline{\epsilon}_{ijk}\, A^i\, B^j\, C^k \tag{13.7}$$

Under a coordinate transformation, I obtain the left hand side of the following expression (assuming $\frac{\partial x^i}{\partial y^\alpha} \neq \delta^i_\alpha$):

$$\underline{\epsilon}_{ijk}\frac{\partial x^i}{\partial y^\alpha}\frac{\partial x^j}{\partial y^\beta}\frac{\partial x^k}{\partial y^\gamma}\, A^\alpha\, B^\beta\, C^\gamma \neq \underline{\epsilon}_{\alpha\beta\gamma}\, A^\alpha\, B^\beta\, C^\gamma \tag{13.8}$$

The left hand side is not in general equal to the right hand side because (again, assuming $\frac{\partial x^i}{\partial y^\alpha} \neq \delta^i_\alpha$):

$$\underline{\epsilon}_{ijk}\frac{\partial x^i}{\partial y^\alpha}\frac{\partial x^j}{\partial y^\beta}\frac{\partial x^k}{\partial y^\gamma} \neq \underline{\epsilon}_{\alpha\beta\gamma} \tag{13.9}$$

If one insists that both sides (13.9) be equal for an arbitrary coordinate transformation, then one obtains a contradiction: either $\underline{\epsilon}_{ijk}$ or $\underline{\epsilon}_{\alpha\beta\gamma}$ must fail to satisfy (13.4), but *both* $\underline{\epsilon}_{ijk}$ and $\underline{\epsilon}_{\alpha\beta\gamma}$ are *defined* by equation (13.4). This demonstrates that the permutation symbol $\underline{\epsilon}_{ijk}$ cannot form the components of a tensor.

I can't proceed without introducing an expression for the determinant of a matrix. The permutation symbol $\underline{\epsilon}_{ijk}$ can be used to define the determinant of a matrix $\mathbf{M}$ with components $M^i{}_j$:

$$\det(\mathbf{M}) = \underline{\epsilon}_{ijk}\, M^i{}_1\, M^j{}_2\, M^k{}_3 = \tfrac{1}{3!}\,\underline{\epsilon}_{ijk}\,\underline{\epsilon}^{abc}\, M^i{}_a\, M^j{}_b\, M^k{}_c \tag{13.10}$$

where $3! = 3\times 2\times 1 = 6$ is the factorial of 3. From the determinant formula above, one can deduce the following identity:

$$\underline{\epsilon}_{ijk}\det(\mathbf{M}) = \underline{\epsilon}_{abc}\, M^a{}_i\, M^b{}_j\, M^c{}_k \tag{13.11}$$

The justification for this identity is a bit involved, so I will put it in a footnote.[1] The determinant can also be defined for rank-2 tensors with raised or lowered indices. Particularly useful is the determinant of the metric tensor (which I assume to be positive[2]):

$$|g| := \det(g_{mn}) = \underline{\epsilon}^{ijk}\, g_{i1}\, g_{j2}\, g_{k3} = \frac{1}{3!}\,\underline{\epsilon}^{ijk}\,\underline{\epsilon}^{abc}\, g_{ia}\, g_{jb}\, g_{kc} \tag{13.14}$$

Since I require that there exists an inverse metric g^{ij}, the determinant is nonvanishing. What makes $|g|$ useful is the fact that under a coordinate transformation, it yields the square of the Jacobian determinant:

$$|g'| = J^2\,|g| \tag{13.15}$$

where the Jacobian determinant J is defined as:

$$J := \det\left(\frac{\partial x^i}{\partial y^\alpha}\right) \tag{13.16}$$

I assume that the coordinate transformations do not involve reflections (such as $x\to -x$), so that $J>0$. The transformation property (13.15) may be inferred from the following property of determinants for the 3×3 matrices $\mathbf{M}$ and $\mathbf{N}$:

$$\det(\mathbf{M}\,\mathbf{N}) = \det(\mathbf{M})\,\det(\mathbf{N}) \tag{13.17}$$

As you might imagine, the transformation property (13.15) will make $|g|$ particularly useful for constructing volume integrals—I will briefly discuss this later on.

Now, I go back and compare the left hand side of equation (13.9) with the identity (13.11). Since $\frac{\partial x^i}{\partial y^\alpha}$ form the components of the transformation matrix, I write:

$$\begin{aligned} \underline{\epsilon}_{ijk}\frac{\partial x^i}{\partial y^\alpha}\frac{\partial x^j}{\partial y^\beta}\frac{\partial x^k}{\partial y^\gamma} &= \det\left(\frac{\partial x^i}{\partial y^\alpha}\right)\underline{\epsilon}_{\alpha\beta\gamma} \\ \Rightarrow \qquad \underline{\epsilon}_{ijk}\frac{\partial x^i}{\partial y^\alpha}\frac{\partial x^j}{\partial y^\beta}\frac{\partial x^k}{\partial y^\gamma} &= \underline{\epsilon}_{\alpha\beta\gamma}\, J \end{aligned} \tag{13.18}$$

Equation (13.18) can in some sense be regarded as a transformation law for the permutation symbol $\underline{\epsilon}_{ijk}$. In fact, I can divide both sides of (13.18) by $\det\left(\frac{\partial x^i}{\partial y^\alpha}\right)$ to obtain the following result:

$$\underline{\epsilon}_{\alpha\beta\gamma} = \frac{1}{J}\frac{\partial x^i}{\partial y^\alpha}\frac{\partial x^j}{\partial y^\beta}\frac{\partial x^k}{\partial y^\gamma}\;\underline{\epsilon}_{ijk} \tag{13.19}$$

---

[1] To justify (13.11), I need to establish some properties of a completely antisymmetric rank-3 quantity A_{ijk}, which is defined to be a quantity that satisfies the following property:

$$A_{ijk} = A_{jki} = A_{kij} = -A_{jik} = -A_{ikj} = -A_{kji} \tag{13.12}$$

In 3-dimensions, the antisymmetry property constrains the value of A_{ijk} so that it effectively has one independent component. To see this, first note that if any two indices of A_{ijk} have the same value, then $A_{ijk}=0$. The only nonvanishing components are those for which the indices of A_{ijk} are even or odd permutations of 123, of which there are $3!=6$. Equation (13.12) amounts to six constraints on the nonvanishing components of A_{ijk}, which implies that the six nonvanishing components must all be equal to the same variable α or its negative $-\alpha$. One can infer that in 3 dimensions, any antisymmetric tensor is proportional to $\underline{\epsilon}_{ijk}$, since $\underline{\epsilon}_{ijk} = \pm 1$:

$$A_{ijk} = \alpha\,\underline{\epsilon}_{ijk} \tag{13.13}$$

Now note that both sides of equation (13.11) is an antisymmetric rank-3 quantity (it has three indices that aren't summed over). The indices $\{i,j,k\}$ are antisymmetric by virtue of $\underline{\epsilon}_{ijk}$. This observation (that the indices $\{i,j,k\}$ are antisymmetric) tells us two important facts: equation (13.11) is only nonvanishing if the indices ijk are an even or odd permutation of 123, and the right hand side of (13.11), being antisymmetric, is proportional to $\underline{\epsilon}_{ijk}$. That the constant of proportionality is the determinant of the matrices $\alpha = \det(\mathbf{M})$ comes from contracting the right hand side of (13.11) with $\underline{\epsilon}^{ijk}$, applying the identity $\underline{\epsilon}^{ijk}\,\underline{\epsilon}_{ijk} = 3!$, and recognizing the expression for the determinant in (13.10).

[2] If you plan to study Special and General Relativity, be warned: While the metric tensor in Euclidean space has positive determinant, the metric tensor for *spacetime* has negative determinant.

In this form, the permutation symbol $\underline{\epsilon}_{ijk}$ *almost* transforms like a tensor; only the factor of $1/J$ prevents (13.19) from being a tensor transformation law.

Fortunately, there is a simple fix. Recalling that the determinant of the metric $|g| := \det(g_{mn})$ acquires a factor of J^2 under a coordinate transformation (13.15), I can construct from $\underline{\epsilon}_{ijk}$ a quantity that transforms like a tensor by multiplying $\underline{\epsilon}_{ijk}$ by $\sqrt{|g|}$:

$$\epsilon_{ijk} = \sqrt{|g|}\,\underline{\epsilon}_{ijk} \tag{13.20}$$

The quantity ϵ_{ijk} (no underline!) is called the *Levi-Civita tensor*[1], and it satisfies the following transformation law:

$$\epsilon_{\alpha\beta\gamma} = \frac{\partial x^i}{\partial y^\alpha}\frac{\partial x^j}{\partial y^\beta}\frac{\partial x^k}{\partial y^\gamma}\ \epsilon_{ijk} \tag{13.21}$$

With the definition (13.20) for the Levi-Civita tensor ϵ_{ijk}, I can now write the cross product (13.5) and curl (13.6) as tensor equations:

$$[\vec{A}\times\vec{B}]^i = g^{ia}\,\epsilon_{ajk}\,A^j\,B^k = \sqrt{|g|}\,g^{ia}\,\underline{\epsilon}_{ajk}\,A^j\,B^k \tag{13.22}$$

$$[\vec{\nabla}\times\vec{A}]^i = g^{ia}\,g^{jb}\,\epsilon_{ajk}\,\nabla_b A^k = \sqrt{|g|}\,g^{ia}\,g^{jb}\,\underline{\epsilon}_{ajk}\,\nabla_b A^k \tag{13.23}$$

I conclude this section with a brief discussion of the volume integral. It is well-known that the volume element $dx^1\,dx^2\,dx^3$ acquires a factor of the Jacobian determinant J under a coordinate transformation.[2] From equation (13.15), it follows that the square root of $|g| := \det(g_{mn})$ acquires a factor of the Jacobian determinant:

$$\sqrt{|g'|} = J\,\sqrt{|g|} \tag{13.24}$$

In Cartesian coordinates on Euclidean space, the metric g_{ij} is just the Kronecker delta δ_{ij}, which has a determinant of 1. In curvilinear coordinates on Euclidean space, the determinant of the metric is just the square of the Jacobian determinant. This motivates the following definition for the volume element d^3V in curvilinear coordinates:

$$d^3V := \sqrt{|g|}\,dx^1\,dx^2\,dx^3 \tag{13.25}$$

from which we obtain the following expression for the volume integral of a scalar function $f(x^a)$ in curvilinear coordinates:

$$\int f(x^a)\,d^3V = \int f(x^a)\,\sqrt{|g|}\,dx^1\,dx^2\,dx^3. \tag{13.26}$$

[1] Strictly speaking, it is not exactly a tensor since it acquires an extra negative sign under coordinate transformations involving reflections (parity transformations), which I ignore in this section. For this reason, the Levi-Civita tensor is sometimes called a *pseudotensor*.

[2] The full proof of this statement is beyond the scope of these notes; one can find a proof in [1] (see theorem 15.11), or one may provide a justification through the formalism of differential forms, which is also beyond the scope of these notes.

Exercise XIII.1

If you have taken a course in Electromagnetism, you should be familiar with Maxwell's equations. Write down the vector form for Maxwell's equations in the standard vector notation—if you have a background in physics, do it from memory! If not, look them up.[3] Rewrite all of Maxwell's equations in tensor form. Assume that $\partial\vec{E}/\partial t$ and $\partial\vec{B}/\partial t$ transform as vectors.

Exercise XIII.2

Show (or convince yourself) of the following identities for the permutation symbol $\underline{\epsilon}_{ijk}$ (assuming 3 dimensions):

$$\underline{\epsilon}_{ijk}\,\underline{\epsilon}^{ijk} = 3! = 6 \tag{13.27}$$

$$\underline{\epsilon}_{jab}\,\underline{\epsilon}^{iab} = 2\,\delta^i_j \tag{13.28}$$

$$\underline{\epsilon}_{kij}\,\underline{\epsilon}^{kab} = \delta^a_i\,\delta^b_j - \delta^b_i\,\delta^a_j \tag{13.29}$$

Exercise XIII.3

Show that the second equality in (13.10) holds. In particular, show that (expand the sums):

$$\underline{\epsilon}_{ijk}\,M^i{}_1\,M^j{}_2\,M^k{}_3 = \tfrac{1}{3!}\,\underline{\epsilon}_{ijk}\,\underline{\epsilon}^{abc}\,M^i{}_a\,M^j{}_b\,M^k{}_c \tag{13.30}$$

Exercise XIII.4

Consider an antisymmetric rank-2 tensor A_{ij}. Show that in 3-dimensions, the antisymmetry property $A_{ij} = -A_{ji}$ reduces the number of independent components to 3.

Exercise XIII.5

Prove the following expressions by writing them out in tensor form:[4]

$$\vec{A}\cdot(\vec{B}\times\vec{C}) = \vec{B}\cdot(\vec{C}\times\vec{A}) = \vec{C}\cdot(\vec{A}\times\vec{B}) \tag{13.31}$$

$$\vec{A}\times(\vec{B}\times\vec{C}) = \vec{B}\,(\vec{A}\cdot\vec{C}) - \vec{C}\,(\vec{A}\cdot\vec{B}) \tag{13.32}$$

$$\vec{\nabla}\times(\vec{\nabla}\times\vec{A}) = \vec{\nabla}\,(\vec{\nabla}\cdot\vec{A}) - \vec{\nabla}\cdot\vec{\nabla}\vec{A} \tag{13.33}$$

[3] You may find an example of Maxwell's equations in [12, 13].

[4] Hint: The identities (13.27), (13.28), and (13.29) in exercise **XIII.2** may be useful here.

Exercise XIII.6

Recall that partial derivatives commute: $\frac{\partial}{\partial x^i \partial x^j} = \frac{\partial}{\partial x^j \partial x^i}$. Show that the following expressions hold in all coordinate systems:[1]

$$\vec{\nabla} \cdot (\vec{\nabla} \times \vec{A}) = 0 \tag{13.34}$$

$$\vec{\nabla} \times (\vec{\nabla} f) = 0 \tag{13.35}$$

Exercise XIII.7

Write out the expression for the curl $\vec{\nabla} \times \vec{v}$ in cylindrical and spherical coordinates.

Exercise XIII.8

In two dimensions, the permutation symbol $\underline{\epsilon}_{ij}$ only has *two* indices: $\underline{\epsilon}_{12} = -\underline{\epsilon}_{21} = 1$. Write down the two-dimensional version of the integral (13.26). Compute the Jacobian determinant J (13.16) for the polar coordinate functions (10.9) in exercise **X.1**, and also the metric determinant $|g|$; in doing so, show that $J = \sqrt{|g|}$, and write down the explicit expression for the volume element in polar coordinates.

Exercise XIII.9

Work out the expression for the Jacobian determinant J (13.16) for the spherical coordinate functions (10.2), then do the same for $\sqrt{|g|}$ (recall exercise **X.2**). Check that $J = \sqrt{|g|}$, and write down the volume element. You should get the familiar result:

$$d^3V = r^2 \, \sin\theta \, dr \, d\theta \, d\phi \tag{13.36}$$

[1] Hint: Write the equations in tensor form, then show that they hold in Cartesian coordinates. Think about the tensor transformation law (compare your result with equation (9.16))

XIV. SURFACE INTEGRALS, THE DIVERGENCE THEOREM AND STOKES' THEOREM

In this last section, I will review the formalism for surface integrals, and briefly describe how the tensor formalism may be extended to the divergence theorem and Stokes' theorem.

There are two ways of defining a 2d surface in Euclidean space. The first is to define a 2d surface, which I call σ as the level surfaces of some scalar function $\Phi(x^a)$. In particular, points on the surface must have coordinate values such that the following constraint is satisfied:

$$\Phi(x^a) = C \tag{14.1}$$

where C is some constant. This definition for the surface is useful, because the gradient of the function $\Phi(x^a)$ can be used to obtain the components for the unit normal vector to the surface σ (see exercise **XIV.1**):

$$n^i = \frac{1}{\sqrt{g^{ab} \nabla_a \Phi \nabla_b \Phi}} \, g^{ik} \nabla_k \Phi \tag{14.2}$$

where the right hand side of equation (14.2) is evaluated at points that lie on the surface σ.

The other definition is parametric—in particular, I parameterize the surface σ with two parameters, z^1 and z^2, which you can imagine to be coordinates for the surface σ. In fact, I will simply write z^1 and z^2 as z^A, with the convention that capital Latin indices correspond to coordinate indices for the surface σ. The coordinates for points on the surface σ are defined by the coordinate functions $x^i(z^A)$, which are explicitly given by:

$$\begin{aligned} x^1 &= x^1(z^1, z^2) \\ x^2 &= x^2(z^1, z^2) \\ x^3 &= x^3(z^1, z^2) \end{aligned} \tag{14.3}$$

with x^1, x^2, and x^3 being coordinates in 3-dimensional Euclidean space. The parametric definition is useful because it can be used to define the components for a metric tensor[2] γ_{AB} on the surface σ:

$$\gamma_{AB} := \frac{\partial x^i}{\partial z^A} \frac{\partial x^j}{\partial z^B} \, g_{ij} \tag{14.4}$$

The metric tensor γ_{AB} is called the *induced metric* It is not too difficult to see that γ_{AB} is equivalent to the metric tensor g_{ij} for vectors tangent to the surface σ; given the components $\mathcal{T}^A$ for a tangent vector on σ in

[2] The surface σ may in general be curved; in general, there exists no coordinate system on σ such that $\gamma_{AB} = \delta_{AB}$ everywhere on the surface σ. Incidentally, if the surface σ is curved, then it is an example of a non-Euclidean space.

the basis $\frac{\partial}{\partial z^A}$, I can construct the components T^i for a tangent vector to σ in the basis $\frac{\partial}{\partial x^i}$:

$$T^i = \frac{\partial x^i}{\partial z^A}\, \mathcal{T}^A \tag{14.5}$$

It follows from (14.5) that for vector components $\mathcal{T}^A$ and $\mathcal{S}^A$ in the basis $\frac{\partial}{\partial z^A}$, the induced metric yields the same result as the vector components T^i and S^i in the basis $\frac{\partial}{\partial x^i}$:

$$\gamma_{AB}\, \mathcal{T}^A\, \mathcal{S}^B = g_{ij}\, \mathcal{T}^i\, \mathcal{S}^j \tag{14.6}$$

Since the induced metric γ_{AB} is equivalent to g_{ij} for tangent vectors, then the following line element:

$$ds^2 = \gamma_{AB}\, dz^A\, dz^B \tag{14.7}$$

is equivalent to the line element (6.9) constructed from g_{ij} for distances along curves on the surface σ.

I now describe a strategy for establishing the relationship between the constraint definition (14.1) and the parametric definition (14.3). One way to bridge the gap is to use an *induced parameterization*, in which two of the Euclidean coordinates (x^1, x^2, x^3) are used parameters, so that the parameterization includes the following coordinate functions:

$$\begin{aligned} x^1 &= z^1 \\ x^2 &= z^2 \end{aligned} \tag{14.8}$$

The last coordinate function $x^3(z^1, z^2)$ is then obtained by solving the constraint equation $\Phi(x^a) = C$ for x^3. Of course, the induced parameterization (14.8) assumes that σ is not a surface of constant x^1 or x^2. Also, the induced parameterization may only work on a portion of the surface, as the surface σ may contain more than one point that has the *same* values for x^1 and x^2.

Given a metric tensor γ_{AB} for the surface σ and equation (13.26) for the volume integral, it is natural to construct the surface integral in the following way:[1]

$$\int_\sigma f(z^A)\, d^2a = \int_\sigma f(z^A)\, \sqrt{|\gamma|}\, dz^1\, dz^2 \tag{14.9}$$

where the 2d metric determinant $|\gamma|$ is given by the following formula:

$$|\gamma| := \det(\gamma_{AB}) = \gamma_{11}\, \gamma_{22} - \gamma_{12}\, \gamma_{21} \tag{14.10}$$

As argued earlier, the metric γ_{AB} provides a measure of distance (in the form of the line element (14.7) on the surface σ that is equivalent to the measure of distance (in the form of the line element (6.9) provided by the metric g_{ij}. From this, one can infer that the area element in (14.9), being constructed from the determinant of γ_{AB} is consistent with the way we usually measure lengths in Euclidean space.

Given the components of the unit normal vector n^i (14.2) and the area element in (14.9) I can define the *directed surface element*:

$$d^2\Sigma_i := n_i\, \sqrt{|\gamma|}\, dz^1\, dz^2 \tag{14.11}$$

For some region U in Euclidean space with a 2d boundary surface ∂U, the *divergence theorem* for a vector field v^i can then be written as:

$$\int_U \nabla_i v^i\, \sqrt{|g|}\, dx^1\, dx^2\, dx^3 = \int_{\partial U} v^i\, d^2\Sigma_i \tag{14.12}$$

Now consider a 2d surface σ bounded by the closed 1-d path $\partial\sigma$. The closed 1-d path $\partial\sigma$ may be described by a parameterized curve $x^i(s)$, where s is some parameter. We may write the differential dx^i for the 1-d path $\partial\sigma$ as:

$$dx^i = \frac{dx^i}{ds}\, ds \tag{14.13}$$

The classical[2] Stokes' theorem for a vector field v^i can then be written as:

$$\int_\sigma g^{ia}\, g^{jb}\, \epsilon_{ajk}\, \nabla_b v^k\, d^2\Sigma_i = \int_{\partial\sigma} v_i\, dx^i \tag{14.14}$$

where ϵ_{ajk} is the Levi-Civita tensor (13.20), and the line integral on the right hand side is performed in the right-handed direction with respect to the directed surface element Σ_i.

Exercise XIV.1

Convert the gradient of a function $\Phi(x^a)$ to a vector (with raised indices) $\mathcal{G}^i$. Take the inner product of the gradient vector $\mathcal{G}^i$ with another vector $\mathcal{T}^i$, and simplify it to obtain a familiar result. Do you recognize the result? Use this result to argue that if the vector $\mathcal{T}^i$ is tangent to the level surfaces (the surfaces of constant $\Phi(x^a)$) of the function $\Phi(x^a)$, the inner product $g_{ij}\, \mathcal{G}^i\, \mathcal{T}^j$ vanishes. Since the inner product between the gradient vector $\mathcal{G}^i$ and a tangent vector $\mathcal{T}$ vanishes, this demonstrates that the gradient vector $\mathcal{G}^i$ is normal to the level surfaces of $\Phi(x^a)$.

---

[1] The 2d surface element $\sqrt{|\gamma|}\, dz^1\, dz^2$ can be derived from the 3d surface element (13.25) by way of differential forms, but differential forms are beyond the scope of these notes.

[2] I say *classical* Stokes' theorem to distinguish (14.14) from the generalized Stokes' theorem, which is expressed with differential forms. Further discussion of the generalized Stokes' theorem may be found in many textbooks on General Relativity such as [5, 8, 9, 18, 21, 31, 32], and also in the books [2, 3, 11, 16, 19, 20, 27].

Exercise XIV.2

Briefly explain why I need to divide by $\sqrt{g^{ab}\nabla_a\Phi\nabla_b\Phi}$ on the right hand side of the formula (14.2) for the components of the unit normal vector n^i.

Exercise XIV.3

Rewrite equations (14.12) and the (classical) Stokes' theorem (14.14) in the (Gibbs-Heaviside) form used in elementary vector calculus courses.[1]

Exercise XIV.4

Rewrite the integral version of Maxwell's equations in "tensor form".

Exercise XIV.5

Consider the usual Cartesian coordinates x, y, z on Euclidean space. The function Φ for a sphere is given by:

$$\Phi = x^2 + y^2 + z^2 \tag{14.15}$$

and the constraint is:

$$\Phi = r^2 \qquad \Rightarrow \qquad x^2 + y^2 + z^2 = r^2 \tag{14.16}$$

for some constant r. Construct an induced parameterization for the sphere. Specifically, parameterize the sphere with the parameters p and q, and write down the two coordinate functions:

$$\begin{aligned} x(p,q) &= p \\ y(p,q) &= q \end{aligned} \tag{14.17}$$

Solve equation (14.16) for z to obtain the function $z(x,y)$, and use (14.17) to obtain the function $z(p,q)$. Now take the following derivatives:

$$\frac{\partial x}{\partial p}\,, \quad \frac{\partial x}{\partial q}\,, \quad \frac{\partial y}{\partial p}\,, \quad \frac{\partial y}{\partial q}\,, \quad \frac{\partial z}{\partial p}\,, \quad \frac{\partial z}{\partial q} \tag{14.18}$$

Finally, use your result for (14.18) with (14.4) to obtain the induced metric γ_{AB} for the sphere.

Exercise XIV.6

Again, find the induced metric γ_{AB} for the sphere, but this time, start in spherical coordinates (as defined in (10.2)) on Euclidean space. What is the constraint function $\Phi(r,\theta,\phi)$?[2] It is appropriate to parameterize the surface using θ and ϕ. You should obtain the following result for the metric components:

$$\begin{aligned} \gamma_{\theta\theta} &= 1 \\ \gamma_{\phi\phi} &= r^2\,(\sin\theta)^2 \end{aligned} \tag{14.19}$$

Now take the determinant of γ_{AB} to obtain the surface element, then perform the following integral over the entire sphere to obtain the surface area:

$$A = \int_\sigma \sqrt{|\gamma|}\, d\theta\, d\phi \tag{14.20}$$

and check that your result is consistent with the surface area of the sphere.

Exercise XIV.7

Given a matrix $\mathbf{M}$ dependent on some variable s, the Jacobi formula for the derivative of the determinant is the following:

$$\frac{\partial(\det(\mathbf{M}))}{\partial s} = \det(\mathbf{M})\, \tilde{M}_j{}^i\, \frac{\partial M^i{}_j}{\partial s}. \tag{14.21}$$

where $\tilde{M}_j{}^i$ form the components of the inverse matrix $\mathbf{M}^{-1}$. Use this to infer a similar Jacobi determinant formula for the derivative of $\sqrt{|g|}$ with respect to some variable s. Use the result to show the following:

$$\frac{\partial\sqrt{|g|}}{\partial x^j} = \sqrt{|g|}\,\Gamma^a_{ja} \tag{14.22}$$

Use (14.23) to show that:

$$\sqrt{|g|}\,\nabla_i v^i = \frac{\partial}{\partial x^j}\left(\sqrt{|g|}\,v^i\right) \tag{14.23}$$

[1] See for instance, equations 10.17 and 11.9 in [4] or equations II-30 and III-13 in [26].

[2] Hint: It's really trivial.

XV. FURTHER READING

As I mentioned in my introductory remarks, these notes are meant to be self-contained, and I have done my best to build the subject from first principles. Though I have written these notes for a broad audience, these notes may be unsatisfactory to some—they may not be pedagogical or rigorous enough to satisfy everyone. Those familiar with early versions of these notes can attest to the significant changes I have made—it turns out that over time, even *I* can become dissatisfied with my notes! I must also mention that these notes do not form a complete discussion of tensor analysis; I have omitted many important topics, such as differential forms (and the associated formalism), Lie differentiation, and of course, Riemannian geometry.

If you wish to learn more about tensors and their applications (or if you find these notes lacking), I have compiled a list of books, many of which I have cited throughout these notes. This list is by no means comprehensive, and mainly reflects my background in General Relativity—I am certain that I have only sampled a small fraction of the literature which discusses tensors and their applications. The books contained in this list, which form a (rather large) subset of the references on the following page, are chosen because they either contain a more complete discussion of the material contained in these notes, contain an extensive discussion of the tensor formalism, or discuss applications of the tensor formalism in geometry and physics.

Books on Tensors

Synge, J. and Schild, A. (1949). *Tensor Calculus: by J.L. Synge and A. Schild.* University Press

Jeevanjee, N. (2015). *An Introduction to Tensors and Group Theory for Physicists.* Springer

Lawden, D. (2012). *An Introduction to Tensor Calculus: Relativity and Cosmology.* Dover

Lovelock, D. and Rund, H. (1989). *Tensors, Differential Forms, and Variational Principles.* Dover

Bishop, R. and Goldberg, S. (1968). *Tensor Analysis on Manifolds.* Dover

Books on Physics and Mathematics

Matzner, R. A. and Shepley, L. C. (1991). *Classical Mechanics.* Prentice Hall

Schutz, B. (1980). *Geometrical Methods of Mathematical Physics.* Cambridge University Press

Baez, J. C. and Muniain, J. (1994). *Gauge Fields, Knots, and Gravity.* World Scientific

Eisenhart, L. (1925). *Riemannian Geometry.* Princeton University Press

Frankel, T. (2011). *The Geometry of Physics: An Introduction.* Cambridge University Press

Nash, C. and Sen, S. (2011). *Topology and Geometry for Physicists.* Dover

Nakahara, M. (2003). *Geometry, Topology and Physics, Second Edition.* Taylor & Francis

Penrose, R. (2004). *The Road to Reality: A Complete Guide to the Laws of the Universe.* Random House

General Relativity Textbooks

Schutz, B. (1985). *A First Course in General Relativity.* Cambridge University Press

Choquet-Bruhat, Y. (2015). *Introduction to General Relativity, Black Holes, and Cosmology.* Oxford University Press

Dray, T. (2014). *Differential Forms and the Geometry of General Relativity.* Taylor & Francis

Zee, A. (2013). *Einstein Gravity in a Nutshell.* Princeton University Press

Carroll, S. (2004). *Spacetime and Geometry: An Introduction to General Relativity.* Addison Wesley

Raychaudhuri, A., Banerji, S., and Banerjee, A. (1992). *General Relativity, Astrophysics, and Cosmology.* Springer

Padmanabhan, T. (2010). *Gravitation: Foundations and Frontiers.* Cambridge University Press

Weinberg, S. (1972). *Gravitation and cosmology: principles and applications of the general theory of relativity.* Wiley

Wald, R. (1984). *General Relativity.* University of Chicago Press

Poisson, E. (2004). *A Relativist's Toolkit: The Mathematics of Black-Hole Mechanics.* Cambridge University Press

Ciufolini, I. and Wheeler, J. (1995). *Gravitation and Inertia.* Princeton University Press

Misner, C. W., Thorne, K. S., and Wheeler, J. A. (1973). *Gravitation.* W. H. Freeman

ACKNOWLEDGMENTS

My advisor Richard Matzner is responsible for most of my knowledge of the tensor formalism, and I thank him for his feedback on earlier versions of these notes. Austin Gleeson, Richard Hazeltine and Philip Morrison taught me some useful identities and introduced me to some interesting applications for the tensor formalism beyond that of General Relativity. Early in my graduate career, Lawrence Shepley and Philip Morrison organized a seminar in differential geometry, which greatly enriched my knowledge and understanding of tensors. I thank Luis Suazo for his visualization of dual vectors/one-forms, and for helping me to think more carefully about the meaning of the formalism. I also thank Brandon Furey for alerting me to some definitions that students may not be aware of. I wish to thank all those that attended the lectures I gave for a crash course in General Relativity, and to all who provided suggestions, typo fixes and feedback (whether in person or by correspondence). Among them are Mark Baumann, Sophia Bogat, Alex Buchanan, Bryce Burchak, Joel Doss, Jean-Jacq du Plessis, Blake Duschatko, Justin Kang, Dave Klein, Baruch Garcia, Bryton Hall, Steven Lobo, Alan Myers, Murilo Moreira, Avery Pawelek, Mark Selover, Zachariah Rollins, Lucas Spencer, Romin Stuart-Rasi, Mehrdad Toofan, Paul Walter, Sam Wang and Shiv Akshar Yadavalli. Special thanks to Sophia Bogat, Alex Buchanan, Jean-Jacq du Plessis, and Murilo Moreira, whose diligence in reading through these notes and working out the homework problems helped me to discover and fix many typos and ambiguities that were present in earlier drafts.

REFERENCES

[1] Apostol, T. (1974). *Mathematical Analysis*. Addison-Wesley.
[2] Baez, J. C. and Muniain, J. (1994). *Gauge Fields, Knots, and Gravity*. World Scientific.
[3] Bishop, R. and Goldberg, S. (1968). *Tensor Analysis on Manifolds*. Dover.
[4] Boas, M. (2005). *Mathematical Methods in the Physical Sciences*. Wiley.
[5] Carroll, S. (2004). *Spacetime and Geometry: An Introduction to General Relativity*. Addison Wesley.
[6] Chappell, J. M., Iqbal, A., Hartnett, J. G., and Abbott, D. (2015). *The vector algebra war: a historical perspective*. (arXiv:1509.00501).
[7] Choquet-Bruhat, Y. (2015). *Introduction to General Relativity, Black Holes, and Cosmology*. Oxford University Press.
[8] Ciufolini, I. and Wheeler, J. (1995). *Gravitation and Inertia*. Princeton University Press.
[9] Dray, T. (2014). *Differential Forms and the Geometry of General Relativity*. Taylor & Francis.
[10] Eisenhart, L. (1925). *Riemannian Geometry*. Princeton University Press.
[11] Frankel, T. (2011). *The Geometry of Physics: An Introduction*. Cambridge University Press.
[12] Griffiths, D. (1999). *Introduction to Electrodynamics*. Prentice Hall.
[13] Jackson, J. (1998). *Classical Electrodynamics*. Wiley.
[14] Jeevanjee, N. (2015). *An Introduction to Tensors and Group Theory for Physicists*. Springer.
[15] Lawden, D. (2012). *An Introduction to Tensor Calculus: Relativity and Cosmology*. Dover.
[16] Lovelock, D. and Rund, H. (1989). *Tensors, Differential Forms, and Variational Principles*. Dover.
[17] Matzner, R. A. and Shepley, L. C. (1991). *Classical Mechanics*. Prentice Hall.
[18] Misner, C. W., Thorne, K. S., and Wheeler, J. A. (1973). *Gravitation*. W. H. Freeman.
[19] Nakahara, M. (2003). *Geometry, Topology and Physics, Second Edition*. Taylor & Francis.
[20] Nash, C. and Sen, S. (2011). *Topology and Geometry for Physicists*. Dover.
[21] Padmanabhan, T. (2010). *Gravitation: Foundations and Frontiers*. Cambridge University Press.
[22] Penrose, R. (2004). *The Road to Reality: A Complete Guide to the Laws of the Universe*. Random House.
[23] Poisson, E. (2004). *A Relativist's Toolkit: The Mathematics of Black-Hole Mechanics*. Cambridge University Press.
[24] Raychaudhuri, A., Banerji, S., and Banerjee, A. (1992). *General Relativity, Astrophysics, and Cosmology*. Springer.
[25] Savov, I. (2017). Linear algebra explained in four pages. Available at: `https://minireference.com/static/tutorials/linear_algebra_in_4_pages.pdf`.
[26] Schey, H. (2005). *Div, Grad, Curl, and All that: An Informal Text on Vector Calculus*. W.W. Norton.
[27] Schutz, B. (1980). *Geometrical Methods of Mathematical Physics*. Cambridge University Press.
[28] Schutz, B. (1985). *A First Course in General Relativity*. Cambridge University Press.
[29] Singh, K. (2013). *Linear Algebra: Step by Step*. OUP Oxford.
[30] Synge, J. and Schild, A. (1949). *Tensor Calculus: by J.L. Synge and A. Schild*. University Press.
[31] Wald, R. (1984). *General Relativity*. University of Chicago Press.
[32] Weinberg, S. (1972). *Gravitation and cosmology: principles and applications of the general theory of relativity*. Wiley.
[33] Zee, A. (2013). *Einstein Gravity in a Nutshell*. Princeton University Press.

INDEX