

CONTRAVARIANCE, COVARIANCE, DENSITIES, AND ALL THAT: AN INFORMAL DISCUSSION ON TENSOR CALCULUS

CHRIS TIEE

ABSTRACT. This collection of notes attempts to demystify some of the vagaries of tensor analysis, to emphasize the connections between the old way of doing things and the new ways, as well as to hopefully illuminate some of the more mathematically obscure aspects which turn up all the time in physics but never get any formal treatment mathematically. In short it is a summary from what I have learned from the Quest For The Holy Grail of Understanding Tensors.

CONTENTS

0. Introduction	2
0.1. Notational Conventions	3
0.2. A Brief Aside on Useful Visualizations	6
0.3. Acknowledgements	7
1. Terminological Confusions	7
1.1. Tensors vs. Tensor Fields	7
1.2. Do Vectors Have Location?	8
1.3. Invariance Under Coordinate Changes	9
1.4. Contravariance and Covariance	13
1.5. Geometric Quantities?	14
1.6. Infinitesimals vs. Linear Functionals, Volume Forms, and Rethinking Integration	15
2. Transformation Laws and Their Relation to Fancy-Schmancy Tensors	17
2.1. Transformation Laws Again	17
2.2. Fancy-Schmancy Tensors: Densitization	19
2.3. Visualization of Forms as Tensor Densities	23
3. Pseudotensors	25
3.1. The Plan	26
3.2. Recap of Orientation on Vector Spaces	27
3.3. The Pseudoscalar Algebra	28
3.4. Pseudovectors and Their Tensor Spaces	30
3.5. The Orientation Line Bundle	33
3.6. Differential Pseudoforms and Integration	35
3.7. Oriented Manifolds: Identifying Axial and Polar Quantities	39
3.8. Honest Densities and De-densitization	41
3.9. Flux and Integration Over Hypersurfaces	42
3.10. The Cross Product	45
3.11. Orientability of Vector Bundles: Bonus Section!	45
4. Duals, Double-Duals, and Canonical Transformations	46
5. Bilinear Forms, and Their Relation to Metrics and Eigenstuff	47

Appendix A. A Note on Texts	47
Appendix B. On Heartbreak: A Parody on Physics Texts and Tensors	50
B.1. Introduction	50
B.2. Stendhal's Law and Sensitivity	50
B.3. Explanation of the Terms	51
References	52

0. INTRODUCTION

Anyone who has survived learning the calculus of tensors and their meaning must have, in fact, known what a pain in the back end it is. It is no surprise that it took an Einstein to invent General Relativity. Heck, it was a major reason why GR took 10 years to formulate after SR, as Einstein himself had many complaints about his mathematical troubles from this era. There are posters that show a picture of Einstein which have the quote "If you think you have troubles with math, believe me, mine are far greater." This undoubtedly refers to his learning of this monstrosity.

Over the years, of course, as the higher powers of abstraction in mathematics gained maturity, and the field of differential geometry became more invigorated (much thanks to Einstein himself for that), the tensor concept has been clarified, re-clarified, and elegantized¹, that is, distilled into a nicer, more streamlined theory. It nevertheless is still a bit hairy in many places, but perhaps Einstein would not have as much trouble with it if he had today's machinery to start with. Unfortunately much of the physics community has been much slower in catching on (not of course because of any intellectual inferiority, but rather the fact they have been far more entrenched in old conventions due to the fact that they have to use them a lot,²), and so there are a variety of texts that just give them as definitions of quantities which transform in a certain way, and since computations can only be usefully done when a coordinate system has been chosen, everything is done that way to practice, and they of course don't care why things work out mathematically, only that they do work out. On the other hand, mathematicians are snobbish and turn their noses up at the "old" way, regarded as "computational BS," and instead replace it with the "new way" which of course is "formal BS"---a decidedly elegant theory of tensors, but hardly more meaningful than "quantities which transform as bar-this-bar-that is equal to lots and lots and lots of indices in funny places times unbarred-this-unbarred-that."

In short, the whole connection between old and new gets lost, and there are precious few "bridge texts" which help relate the old and new ways. In addition there are a few concepts which get marooned in the transition... concepts like tensor densities, pseudotensors (axial quantities), and the fact that orientability is a major source of identifications in a similar manner of how the metric is used to identify contravariance and covariance. These tensor-related topics (anything related to which, shall henceforth be qualified with the adjective **fancy-schmancy**, John Baez's terminology for these things) get a brief mention in physics texts (with one exception) and are completely swept under the rug in mathematical treatments. Also, all texts seem to make assumptions about some kind of terminology, which may seem obvious at one moment but at odds with what one's understanding the next.

The goal of this informal discussion is to complain about and hopefully clarify some issues, as well as organize stuff that still is to be demystified---collecting the known unknowns,

¹I don't know if that's a word, but well, I'll make it up.

²And the fact I'm writing this for a physicist.

so to speak. As such, it is a work in progress and many sections will be marked with the dreaded “(Under Construction. . .).” Also, in the name of informality, lightheartedness, and general optimism in taming the monster that is Tensor Analysis, I shall of course at times insert strange and sometimes irrelevant commentary, and jokes (some bad ones, too). Don’t take things too seriously. So, *since you have been warned*, we shall henceforth abide by the following principle:

0.1. **Mike Kinally’s Principle of Offense.** *“If you are offended, well, it’s your own damn fault.”*

Now let’s get on to some more important stuff.

0.1. **Notational Conventions.** Now for the obligatory initial conventions section in any math text/paper. I assume that you’ve learned some tensor calculus at some point and know how to work with them on a formal level. Also hopefully you’ve learned at least what the invariant conceptions of these things are (multilinear maps, sections of the vector bundle of tensors over a manifold, and so forth), although I’ll briefly describe them as I go along. I’m not expecting readers to either be satisfied about their tensor-analytic knowledge nor for them to have a deep understanding of the concepts³, and I *am* expecting the reader to have complaints and be frustrated over this. That’s why I’m writing this. . . Also, some things in here might be unfamiliar but are clarified later on, so don’t be too put off if you should be confused here; skip ahead and see what you like.

The notations I use are relatively standard. I use upper indices for contravariant components, lower for covariant components, with the convention that an “upper index in the denominator” such as $\frac{\partial}{\partial x^i}$ is equivalent to a lower index. I use the **Einstein summation convention**, namely, indices appearing as lower indices and upper indices in the same expression get summed over. Some texts allow the convention for all repeated indices. Generally I will reserve it only for upper-lower pairs, but there are a few instances where there is a conflict, as there are, as you shall see, covariant quantities which when viewed a different way, become contravariant (and vice versa). So it does not hurt to use explicit summation signs and they shall appear in here from time to time. The manifolds we work with will be C^∞ and the ring of real-valued functions are denoted $C^\infty(M)$.

Notation

In modern terms, a **tensor** on the space at p is a multilinear map with ℓ slots that accept vectors (**ℓ -covariant**), and k slots that accept covectors (**k -contravariant**), using the canonical double-dual identification $V \cong V^{**}$ (we will devote an entire section to talking about why precisely the double-dual identification is canonical. But for most of the discussion it suffices to know that this identification will be systematically used to describe vectors in terms of covectors). Of course, this formulation is nonsense (circular) if we do not have some “primitive” quantity to start off with. The usual development is to make a primitive realization of vectors on a manifold, and then get happy with multilinear maps later on to concretely realize tensors. The vector space (or linear space, **MVE**⁴ space, or just space) of all k -contravariant, ℓ -covariant tensors (tensors of **valence** $\binom{k}{\ell}$) at the point p in a manifold M will be denoted $T_\ell^k(M)_p$, with TM_p and T^*M_p denoting the special

Modern definition of tensors

³If you, O reader, do have such an understanding, contact me immediately, because I sure as heck don’t. These notes are more complaining, remember?

⁴Eric Michelsen’s terminology for vectors in the abstract mathematical sense. It stands for **mathematical vector element** given in [Mic06].

cases of **vectors** $\binom{1}{0}$ -tensors) and **covectors** or **1-forms** $\binom{0}{1}$ -tensors), respectively. We will mention some good visualizations of these two special cases in the next section.

True vectors vs. abstract vectors

We'll use the term "vector" to just mean $\binom{1}{0}$ -tensor and use MVE for abstract mathematical vectors in any vector space.⁵ A major annoyance for the learner is that there are a zillion different identifications one can make (which results in headaches because sometimes you forget which one is being used), based on representing a function by fixing one variable and considering it as a function of the remaining: T as a mapping that sends (v_1, \dots, v_n) to $T(v_1, \dots, v_n)$ can also be realized as a map $\tilde{T} : v_1 \mapsto T(v_1, \cdot, \dots, \cdot)$, that is, the v_1 fills in a slot and determines a function of the remaining guys. That is to say, to write things funny, $\tilde{T}(v_1)(v_2, \dots, v_n) = T(v_1, \dots, v_n)$, essentially replacing a "(" with a comma. Similar isomorphisms exist for the other slots, and even more than one slot.⁶ We elaborate upon this in the section on duality. The thing is, all of these identifications are "legal" because they are *canonical*; there is an identification of vectors and covectors, too, but it is noncanonical in the absence of a metric structure. Again, more on the precise meaning of canonicity later.

Tensor bundles

Considering the totality of all the $\binom{k}{\ell}$ tensor spaces based at each $p \in M$ gives us a structure called a **tensor bundle** $T_\ell^k(M)$. This can be made into a manifold intuitively by "gluing the spaces together."---more on this in the next paragraph. The corresponding space of **sections** of these bundles, that is, tensor **fields** will be denoted $\mathcal{T}_\ell^k(M)$ (more on this in the next section (pun not intended)). For the space of m -forms, that is, totally antisymmetric covariant tensors of degree m , we write $\Lambda^m(M)_p$, and its associated bundle, by dropping the p . For the corresponding space of sections of the alternating tensor bundles (m -form **fields**) we write $\Omega^m(M)$. Note that $\mathcal{T}_0^0(M) = \Omega^0(M) = C^\infty(M)$. Antisymmetric tensors have an bit of structure, a special product called **wedge product**, written $(\alpha, \beta) \mapsto \alpha \wedge \beta$. The theory of totally antisymmetric tensors is of course intimately related to the study of determinants and we shall use the following two facts which relate wedge products of 1-forms to determinants:

$$(0.1) \quad \omega^1 \wedge \dots \wedge \omega^m(X_1, \dots, X_m) = \det(\omega^j(X_k))$$

where $(\omega^j(X_k))$ is the matrix formed by evaluating the j th 1-form on the k th vector, indices j and k ranging from 1 to m , and also, for forms of top degree n we have, given T a linear transformation of the cotangent space, that the determinant of T comes out when it is applied factorwise:

$$(0.2) \quad (T\omega^1) \wedge \dots \wedge (T\omega^n) = \det(T)\omega^1 \wedge \dots \wedge \omega^n$$

Multivectors

Totally antisymmetric *contravariant* tensors also come up---these represent area or volume elements. They are dual to forms and also have a wedge product. Such a thing is called a **multivector** of degree m , or m -vector. The bundle of such guys will be denoted $\Lambda_m(M)$ (I don't think this is standard notation, but it sure is convenient; note that $\Lambda_1(M)$ are then just vectors) In small cases we use the terms **bivector** for Λ_2 and **trivector** for Λ_3 . These should

⁵Of course the term "vector space" itself uses the term vector, but these two guys appearing together should be noticeable from the context, as the "true vector" space will always be referred to as the **tangent space**. I could use the term "linear space" or "MVE space," but I am an entrenched mathematician, har har.

⁶Computer scientists refer to this as "currying." In set-theoretic terms, this is the linear-category realization of the so-called **exponential law** $(X^Y)^Z \cong X^{Z \times Y}$ which says functions from a set to a set of functions is equivalent to regrouping parentheses which give functions from the product set to the innermost range space. That is, writing $f(x, y) = f(x)(y)$. In other words, $(X \rightarrow (Y \rightarrow Z)) \cong (X \times Y) \rightarrow Z$. There are various analogues of this trick on more structured sets, and is very useful to reduce somewhat intractably large function spaces down to maps between reasonably-sized spaces, since forming function spaces tends to yield infinite-dimensional spaces while products preserve finite-dimensionality.

be visualized as m -dimensional parallelepipeds determined by m vectors.⁷ In particular, the wedge product of two vectors is given by sticking their tails together and looking at the oriented area of the parallelogram they complete (like the cross product in \mathbb{R}^3 ; hence the wedge product actually is a generalization of the cross product⁸), and given the direction “in which it’s facing.”

We’ll mention the general terminology of vector bundles, which perhaps is familiar to gauge-happy physicists. I don’t have time to go into very much detail about these guys but here is the shortest explanation I can give. A vector bundle defines a natural target for fields of funky objects defined on the manifold (for example, a vector field). A **rank- m smooth vector bundle** (henceforth referred to just as a bundle, since we will not have occasion to use some of the more exotic fiber bundles that physicists and topologists like to talk about) over the manifold M is a space E with a smooth map $\pi : E \rightarrow M$, called the **projection** whose **fiber** above each point p , $E_p = \pi^{-1}(p)$ is a vector space of dimension n , which has the property that E locally looks like a product; that is, for each p there is a neighborhood U of p such that $\pi^{-1}(U)$ is diffeomorphic to $U \times \mathbb{R}^m$ in a nice fiber-preserving way, and well-behaved on overlaps (the map that takes vectors in one part of the overlap to the other is called a **transition function** of the bundle). In this local product the projection map looks like the usual projection map onto the first factor. This is mathematical nonsense for “ E looks locally like $M \times \mathbb{R}^m$ except it might be globally ‘twisted’.” Of course it may happen that E isn’t twisted and so it is globally the product---the cartesian product $M \times \mathbb{R}^m$ with projection map being the genuine projection onto the first factor. $M \times \mathbb{R}^m$ is called the **trivial bundle**.

General vector bundles

The canonical trivial example is the bundle \mathbb{R}^n where the fiber is also \mathbb{R}^n , giving the product \mathbb{R}^{2n} . This is just the space where vectors based at that point live, and the usual vector field on \mathbb{R}^n indeed maps \mathbb{R}^n into this. Because \mathbb{R}^n is nice and flat, all these individual spaces are identified in some natural way, so people usually ignore basepoints (more discussion on this in “Do Vectors Have Location?”) and so that’s why you hear of vector fields as being maps from \mathbb{R}^n to \mathbb{R}^n . Anyway here is the translation: given a traditional vector field $\mathbf{F} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ then the associated übersophisticated-mathematically-snobby vector field is the mapping $p \mapsto (p, \mathbf{F}(p))$. For general manifolds, the tangent bundle might have some kind of “twist” so it is harder to separate the so-called “point part” and “vector part” of the equation. The standard counterexample here is the nontrivial **Möbius bundle** over the circle which has fibers going all around, experiencing a literal half-twist along the way---if we tried to separate p and $F(p)$ for some vector field in this guy, we’d get some kind of turnaround. More on this later.

In practice, vector bundles are formed by taking a collection of equal-dimensional vector spaces associated with a point $p \in M$ and “gluing” them together---considering the disjoint union of all these spaces, indexed by the point, and giving them a topology and smooth structure that lends a sense of “coherence” to these vector spaces---if their base-points are close together, then the fibers are somewhat related, i.e. tied together in a . . . bundle (what else?). The visualization is, of course, really only possible in the case of rank-1 bundles except in some really trivial 2-dimensional cases. Hence the terminology “fiber.” Sometimes a rank-1 bundles are called **line bundles**.

⁷Technically this is only an accurate picture of *decomposable* m -vectors, that is, those that can be expressed as m -fold wedge products of ordinary 1-vectors; for the general case we consider their linear combinations, though the picture is not as easy to see!

⁸Textbooks often say “the cross product does not generalize” is that in general it maps into a higher dimensional space, and Λ_1 is *not* closed under this product. Even in \mathbb{R}^3 , of course, Λ_1 is not closed under wedge product, but since Λ_2 is of the same dimension we just identify them.

0.2. A Brief Aside on Useful Visualizations.

“Gradient a 1-form? How so? Hasn’t one always known the gradient as a vector? Yes, indeed, but only because one was not familiar with the more appropriate 1-form concept.”---Charles Misner, Kip Thorne, and John Archibald Wheeler, in [MTW]

Visualizations are great and I probably will have much more to say about this here in the future. However, many texts in fact actually do a good job in this department in effort to be intuitive. The chief purpose I have in mind for these notes is just taming the variety and illuminating some mysterious points. I should mention my very good primary source on these visualization, and the only source that really develops a picture of fancy-schmancy rank-1 tensors (the only shortcoming being that it only talks about ranks 0 and 1): the text *Geometrical Vectors* [We97] by Gabriel Weinreich (more praise to be heaped upon it in Appendix A on texts). Because I shall use these visualizations as models, however, I should definitely introduce them and mention them here.

Vectors as arrows

Traditionally, 3-dimensional MVEs tend to get represented by arrows. In \mathbb{R}^3 , the metric structure is implicitly used to represent all rank-1 tensors (including their densitized varieties) as true vectors ($\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ -tensors). This is in fact an appropriate picture for true vectors, and they should always be visualized in this way, even on a general manifold, where you should visualize them sticking out tangentially (though it seems obvious, this is still an important thing to mention, since the actual definition of a vector on a manifold is often given by some algebraic or functional-analytic object such as a directional derivative).

The Stack

A better picture for a covector ($\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ -tensor) in \mathbb{R}^3 , given by Weinreich, is a *stack*. [Insert picture here!]. A stack is a bunch of planes with evenly spaced sheets in space---the relative density of which determines the magnitude, and the orientation of which determines the direction (i.e. the direction is the direction in which the sheets are facing). Classically this direction would be given by a normal vector to the planes, and for pedagogical purposes we notate a stack by sticking an arrowhead to indicate the direction. Technically the direction is already obvious from looking at the sheets themselves; the only thing left to do is assign a “sense” to the direction i.e. which way the sheets are increasing---if one were going for a completely arrow-less presentation, this could be indicated by distinguishing the “top” sheet. The horizontal extent of the stack (i.e. how big the planes should be) is immaterial, just as, for example, the thickness of the shaft of the arrow representing a vector is irrelevant.

Gradient as a stack

The prototypical model of a stack (actually a stack *field*) is given by the gradient or differential of a scalar function---given a function f , we look at its level surfaces (fixing some scale which describes the increment from one surface to the next), which in a neighborhood of a point, should also be approximately evenly spaced, and approximately flat. Then if we cut out a small portion (Weinreich describes this as cutting the level surfaces with a melon-baller) and consider them as flat, we get the covector which describes the gradient of f at that point (the positive sense of the stack is given by which direction f is *increasing*). Note that this characterization makes the “covector” nature of the gradient immediately apparent in a way that “normals to level surfaces” does not. The informal term “density” for the magnitude of these sheets is closely related to, but not exactly the same as the term density used in *densitization* which we shall describe as a process that can be applied to any tensor; I shall explain the process of densitization in full detail later. The important thing to remember here is that the density is *linear*, and if we scale the space, only the scaling along the direction of the sheets contributes to a change in magnitude. That is, the “units” would be inverse distance, coinciding with the intuition that these guys are dual to arrows. (One would expect the unqualified term “density” to mean volume density, that

is scaling in any direction has an effect). Once again looking at the gradient example, the more closely spaced the level surfaces are, the faster f is increasing, and so the larger the gradient covector. Contrast this to the arrow, which gets *longer* when its magnitude increases. This is why stacks give a better picture for the gradient. Evaluating a covector on a vector (contraction) is then given by a very simple method: put the arrow representing the vector across the sheets of the stack, and count how many intersections you get, including of course fractions. The answer is positive or negative according to whether the arrow traverses the sheets in increasing or decreasing order. If the arrow crosses none at all, it is said to be *contained* in the stack. Classically this means the arrow and the arrow that would be used to represent the stack are orthogonal, but note that the concept of containment is metric-independent concept (in topology, the property of *not* being contained in a stack is the notion of **transversality**).

The n -dimensional generalization of a stack is given by replacing the words “sheet” by “hyperplane” and “level surface” by “level hyperplane” or just “level set” of a scalar-valued function. Since the complement of an $(n - 1)$ -dimensional space is only 1-dimensional (true whether or not a metric is present, although a metric allows us to specify the *orthogonal* complement), things like “spacing between hyperplanes” and “direction” are still well-defined. Once again, the magnitude is dependent on the relative *linear* density of the hyperplanes.

The stack in n dimensions

Weinreich really pulls out the big visualization guns when talking about vector calculus. He gives more pictures that indicate quite naturally how curls and divergences come about via intuitive geometric operations (involving appropriate visualizations of 2-forms), and their integration as well.

0.3. Acknowledgements. Profs, students, various authors, comrades, and da math h0miez. Eric Michelsen, for inspiring me to put my tensorial complaints on paper (or at least in $\text{T}_\text{E}\text{X}$). Ben Chow, for really jump-starting and putting me into high geometric gear. Hans Lindblad, who gives renewed hope to the realization of the dream to understand GR. John Baez for really insisting on the reality of the existence of these things and that there is a truly organized framework in deal with tensor concepts. See the 325-posting argument on sci.physics.research and you’ll see what we mean by dedication to one’s field. The title is a(n) homage to Harry Schey’s classic text *div, grad, curl and All That: An Informal Text on Vector Calculus* [Scy97]⁹.

1. TERMINOLOGICAL CONFUSIONS

There are a lot of tacit terminological conventions when speaking the lingo of tensor analysis. In order the clear the confusion we must explicitly state these conventions and note their pitfalls. There are surprisingly many even in the seemingly most obvious cases. Here we shall fix M to be some n -dimensional smooth manifold.

1.1. Tensors vs. Tensor Fields. First and foremost we have to address the issue of distinguishing tensors from the maps on M which assign to each point p a tensor (in the fiber over p) in a smooth manner. . . i.e. the objects which represent tensors “sprouting” out of each point (hence the term “field”). In übersophisticated bundle lingo, these are called (**cross-**) **sections** of the tensor bundle, which stems from the only real visualizable case: some bundle over a circle whose fiber is drawn vertically at each point (picture here?) so as to make the total space look like a cylinder. A cross section of this bundle is then just a curve

⁹This text of course limits its discussion to \mathbb{R}^3 and does not mention any kind of tensor existing beyond plain old vectors at all! But it is still a great text! (For first learning vector calc, that is!).

on the cylinder, which indeed is like “slicing” the cylinder with some kind of surface to give a selection of points in the fiber above each point. In particular, smooth sections of the trivial line bundle $M \times \mathbb{R}$ are just functions in $C^\infty(M)$. More abstractly, a section is a right inverse of the natural projection map of the bundle. Just keep right on thinking it is what it really is. . . a whole bunch of tensors, each sprouting out at each point in a manifold. The thing is, however, physicists often omit the word “field” when talking about these. Occasionally I am prone to doing the same thing. It is important to take note of this when talking about transformation laws for tensors, because, on a single vector space, we can change the basis to anything we want using change-of-basis matrices.

When speaking of tensor transformation laws, we are usually speaking of converting (tensor products of) *coordinate* basis (co)vectors from one system to another, and this is what yields that $\left(\frac{\partial \bar{x}^i}{\partial x^j}\right)$ business. The thing is, any random field of basis vectors over the manifold that gives us a change-of-basis at each point are not necessarily going to be actual coordinate basis fields (a field of basis vectors which cannot be so represented **nonholonomic**¹⁰). This is described in more detail in the section on transformation laws, and the reader can skip over to there to get precisely what I mean by this. The thing is, it is *always* possible to change a basis at a point and have some change-of-basis matrix, but it is not always possible to do so coherently on a whole coordinate chart. We will see however, nonholonomic fields which are not directly coordinate vectors themselves, but rather somehow *derived* from the coordinate vectors, are in fact useful, and yield funny-looking but important transformation laws (hint: our fancy-schmancy tensors can be looked at in this manner. We’ll get to that). I will try to always say “tensor field” when I mean it.

1.2. Do Vectors Have Location? You were always taught in vector calculus that “vectors have no location; they can be slid around with impunity.” This is blatantly false on a general manifold for two reasons. First, there is no real “canonical” way to do this sliding on a general manifold. In \mathbb{R}^n there is a very natural and obvious identification of tangent spaces, namely sliding things to the origin. This is known to not cause any problems because, the standard structure of \mathbb{R}^n always implicitly assumes the presence of a flat (Euclidean) metric. Using a metric structure (or more precisely, a **connection**) the ability to slide vectors along makes somewhat of a comeback. One place where sliding is implicit is in differentiation of vector fields on \mathbb{R}^n ---in the difference quotient, we have $\partial \mathbf{V} / \partial x^j = \lim_{h \rightarrow 0} (\mathbf{V}(x + h\mathbf{e}_j) - \mathbf{V}(x)) / h$, the two vectors, although they are very close together, still do not technically live in the same tangent spaces, and hence technically cannot be subtracted without taking making their tails meet. This is the whole reason why covariant differentiation was originally invented.

The second reason vectors should be considered to have location is in fact apparent even in \mathbb{R}^n ---the concept of a vector field of course demands each point be assigned some vector. You’d expect to think of this as the range vector being “planted” at its basepoint, and not all of them merrily sitting simultaneously at the origin (although occasionally this viewpoint is useful, e.g. the Gauss map, and the differential-topological concept of the *index* of a vector field). I mean, if you’re going to integrate a vector field over a surface, your vectors better be at those points. So I emphasize this when TAing vector calculus, that just because you *can* slide your vectors in \mathbb{R}^n , it doesn’t mean that you always *should*. One further issue about tensor fields is that there is an equivalent realization of a tensor field as being operators on the associated tensor bundles into $C^\infty(M)$ which is essentially given by filling up all its slots

¹⁰It is well-known that a necessary and sufficient condition for the field of basis vectors $\{E_j\}_j$ to be a coordinate basis is that their Lie brackets $[E_i, E_j]$ all vanish identically.

with appropriate 1-form- and vector fields which I shall clarify in the section on “geometric quantities.”

1.3. Invariance Under Coordinate Changes. Since tensors were invented to describe physical phenomena independent of coordinate system, we absolutely must clarify what we mean by this. That is, we shall demystify the *very* oft-used phrases “**invariant under coordinate changes**” and “**independent of coordinate system**” and other such permutations. Ok so everybody knows what this means, right? You look at it in different coordinates, and the object itself doesn’t change. The potentially confusing point, though, is that some kind of *representation* for the object does change. There are definitely still some subtleties to be pointed out. Here we relate true vectors (and other guys too) with the old picture of vectors being n -tuples of a whole bunch of numbers. The term “invariant object” (which is sometimes confusingly called a “covariant object” in physics, confusing because covariance will actually not be used in this meaning in all that follows) means precisely this. Basically a quantity is invariant if it is defined without reference to a coordinate system. In particular, if it is defined in terms of *other* invariant objects, it is also invariant. It is important to realize, however, that one can define invariant quantities in a non-invariant manner, and this is why it can be confusing.

First off, consider M^n to be a manifold as usual. The most obvious invariant object on our manifold is a smooth function $f \in C^\infty(M)$. Its value at a point is intrinsically defined, although usually we have to use a chart, given by a map ϕ into \mathbb{R}^n , to actually compute it. But there we are computing $(f \circ \phi^{-1})(\phi(p))$ so the contribution of the use of coordinates in dealing with f cancels out. Essentially this canceling out principle is exactly how we characterize invariance for objects which otherwise look like they depend on the choice of chart.

1.1. Example (Constructing smooth functions on a manifold from functions defined on charts). Let us clarify this with a model example. How would we define a global smooth function on a manifold M via charts $(U_\alpha, (x_\alpha^i))$ on M ? Exactly how you think it is done: have a bunch of local smooth functions g_α defined on each U_α and make sure the obvious compatibility condition on the overlaps holds: $g_\alpha = g_\beta$ on the intersection $U_\alpha \cap U_\beta$. Then defining $g(p) = g_\beta(p)$ where $p \in U_\beta$, we have a perfectly good well-defined smooth function on M .

Now let us turn to (tangent) vectors and their fields. Vectors at p are invariantly defined, in many modern treatments, as directional derivatives on $C^\infty(M)$ at the point p . There are other ways to do it, but this way looks directly like invariants being defined in terms of other invariants. The coordinate system on M at p forms the basis $\left. \frac{\partial}{\partial x^i} \right|_p$ for the tangent space TM_p (these are given by taking the ordinary i th partial derivative with respect to the corresponding chart map composed with f) and so given a vector $V_p \in TM_p$ we define the components V^i to be the unique coefficients such that

$$(1.1) \quad V_p = V_p^i \left. \frac{\partial}{\partial x^i} \right|_p.$$

Throughout a chart U about p , of course, the differential operators $\left. \frac{\partial}{\partial x^i} \right|_p$ without the $|_p$ ’s form a *field* of basis vectors. So a vector field V on M defines smooth functions V^i such that V_p^i is given, at p by (1.1) above. It is customary, of course, to rewrite the law without the subscript p ’s. A standard result of the theory that, despite all the abstract hoopla used to define tangent spaces and so forth, a vector field is smooth if and only if the V^i are all smooth

as functions on U , i.e. a vector field is smooth if and only if all its component functions are smooth in any chart.

Now for two subtleties that are related to each other. First, the choice of basis is of course not invariant (it's determined by the coordinate system!). But the *operator* $X = \frac{\partial}{\partial x^i}$ is in fact invariantly defined: it does a certain thing to functions on a manifold. If we change the coordinate system, X does the same thing as it did before. The thing is, it now doesn't have such a nice expression in the new basis. But it does have a perfectly predictable expression when expanded out in the new coordinate system:

$$(1.2) \quad X = \frac{\partial}{\partial x^i} = \frac{\partial \bar{x}^j}{\partial x^i} \frac{\partial}{\partial \bar{x}^j}$$

In other words X_p is an invariant quantity that happens to coincide with a coordinate basis vector but is not obligated to continue being so in different coordinates (in fact, there is a theorem that states that in the neighborhood of any point $p \in M$ where a smooth vector field V does not vanish, there exist coordinates (y^i) on M such that $V = \frac{\partial}{\partial y^1}$).¹¹

Now for the dual idea. The n -tuple V^i determined by (1.1) consist of invariant functions on U which don't change. But they *represent* something, and thus when the coordinate system changes, if we insist on saying that these functions don't change, we have in fact changed what we are representing. This point bears repeating. Even though we *represent* the invariant quantity V by the invariant scalar functions V^i , this just means that there are n functions on M that happen to coincide with the components of V with respect to the coordinate basis. When we change coordinates, those n functions no longer are the components of V in the new basis, although the new functions \bar{V}^i do have a nice expression in terms of the old ones.

Perhaps it is more straightforward to see this from the fact that the transformation of the components is derived from *requiring that the sum (1.1) remain invariant*:

$$(1.3) \quad V = V^i \frac{\partial}{\partial x^i} = V^i \frac{\partial \bar{x}^j}{\partial x^i} \frac{\partial}{\partial \bar{x}^j} = \bar{V}^j \frac{\partial}{\partial \bar{x}^j}$$

means, when collecting terms, that

$$(1.4) \quad \bar{V}^j = V^i \frac{\partial \bar{x}^j}{\partial x^i}$$

Notice how it is backward from the original basis transformation laws. The upshot of all this is that the collection of n functions (V^i) is, in and of itself, insufficient to determine a vector field---we need the information that they do in fact determine a vector field, and so therefore the transformation law applies.

To really drive the point home, we will give two ways to construct a vector field that really illustrates the difference between components and what they represent---the "wrong way" and a "better way."

1.2. Example (Wrong way to construct a vector field). Suppose we are given n invariant functions $g^1, \dots, g^n \in C^\infty(M)$, and a collection of charts (U_α, x_α) on M , that is (x_α^i) are the coordinate functions on the α th chart. Then if we try to naively define a global vector field X on M by setting

$$X = g^i \frac{\partial}{\partial x_\alpha^i}$$

¹¹Even more is true: given n smooth vector fields E_i on M which, at p , form a basis for TM_p , and such that their pairwise Lie brackets vanish, there exist coordinates (y^i) in a neighborhood U such that $E_i = \frac{\partial}{\partial y^i}$ over all of U .

in every chart---that is to say, the g^i 's don't change no matter what, this X is certainly not going to be well-defined in general (the case that α is a singleton and M can be covered with one chart being the exception here).

1.3. Example (Better way to construct a vector field). To fix the situation up, let's consider something analogous to what we do with defining smooth functions from a collection of individuals defined on the charts. That is, we suppose instead we have a whole bunch of functions $g_\alpha^1, \dots, g_\alpha^n$, that is, n smooth functions for each chart U_α . This looks so far like an identical situation in defining a global smooth function on M . But the twist is now in the compatibility condition: g_α^i is related to g_β^i on the overlap $U_\alpha \cap U_\beta$ *not by the silly condition* $g_\alpha^i = g_\beta^i$ *but rather by the more complicated transformation law for components of vector fields given above.* That is, the appropriate compatibility condition is

$$g_\beta^i = g_\alpha^j \frac{\partial x_\beta^i}{\partial x_\alpha^j}$$

which is the same equation as (1.4) except instead of using barred and unbarred, we're using specific labels α and β . Then the naïve way of doing things goes through; just define $X_p = g_\beta^i(p) \frac{\partial}{\partial x_\beta^i}$ when U_β contains the point p . Using the non-silly condition would glue everything together into n smooth global functions on M , giving us the situation above.

In short terms, *n -tuples alone do not a vector make.* The point of the transformation law is to be able to see when defining some indexed quantities on charts will yield a globally-defined invariant object. Actually, there is nothing wrong with defining n -tuples of functions via the silly condition, that all their values on overlaps are equal. But this does not define a vector field (a section of the tangent bundle TM), but rather it defines a mapping into \mathbb{R}^n (a section of the trivial bundle $M \times \mathbb{R}^n$). Those two bundles are *not* the same!

In general any sum over upper and lower indices is supposed to represent an invariant, and the transformation law of one is dual to the other. That's why basis vectors and components transform oppositely. That's also why the upper-lower index summation convention was invented---not to make it a pain for readers (however like that it may seem), but rather as an aid to tell at a glance which expressions are supposed to represent invariant quantities. For another example, consider the transformation law for covectors. This is done in the manner to preserve the invariance of contraction, i.e. for a 1-form ω we have

$$(1.5) \quad \omega(V) = \omega_i V^i = \bar{\omega}_j \bar{V}^j = \bar{\omega}_j V^i \frac{\partial \bar{x}^j}{\partial x^i}$$

where ω_i is determined by $\omega = \omega_i dx^i$, dx^i being the dual to $\frac{\partial}{\partial x^i}$. This means the components of ω *must* satisfy:

$$(1.6) \quad \omega_i = \bar{\omega}_j \frac{\partial \bar{x}^j}{\partial x^i}$$

which is again backward from the transformation of the vector law. Appealing to duality once more with ω_i and dx^i we can also get the transformation law for those too.

1.4. Example. Now we move on to a more interesting example that will be relevant to our tentative development of fancy-schmancy tensors. Recall that, for an n -dimensional manifold M , the vector space $\Lambda^n(M)$ at each point is only 1-dimensional, i.e. it is a rank-1 bundle, due to the alternating nature of the algebra. Λ^n itself is often called the **determinant line bundle**. This means that, even though $\Lambda^n(M)$ sits in the very large n^n -rank bundle of $\binom{0}{n}$ -tensors on M , it is most naturally expressed in terms of one basis

element. For example, $dx^1 \wedge \cdots \wedge dx^n$ is such a basis. So a general n -form, in coordinates, is represented by a function $f \in C^\infty(M)$ times the basis n -form $dx^1 \wedge \cdots \wedge dx^n$. The issue that now appears before us here is, since $\Lambda^n(M)$ is rank-1, as is \mathbb{R} , can we identify their cross sections, namely $\Omega^n(M)$ and $\Omega^0(M) = C^\infty(M)$? The former is represented by a single function times some n -form, while the latter is just a function. The thing is, however, coordinates appear in $f dx^1 \wedge \cdots \wedge dx^n$. Let's look at the transformation law. Using the usual 1-form basis transformation law, we write: $d\bar{x}^i = \frac{\partial \bar{x}^i}{\partial x^j} dx^j$. But note that this represents a linear transformation T that sends dx^i to $d\bar{x}^i$, whose action is given by the Jacobian (matrix) $\left(\frac{\partial \bar{x}}{\partial x}\right)$. So rewriting, we have $d\bar{x}^i = \left(\frac{\partial \bar{x}}{\partial x}\right) dx^i$ and

$$d\bar{x}^1 \wedge \cdots \wedge d\bar{x}^n = \left(\frac{\partial \bar{x}}{\partial x}\right) dx^1 \wedge \cdots \wedge \left(\frac{\partial \bar{x}}{\partial x}\right) dx^n = \det\left(\frac{\partial \bar{x}}{\partial x}\right) dx^1 \wedge \cdots \wedge dx^n$$

by the formula (0.2) involving factorwise application of a linear transformation to an n -form pulling out to become the determinant. Hence we have the following nice transformation law:

$$(1.7) \quad d\bar{x}^1 \wedge \cdots \wedge d\bar{x}^n = \det\left(\frac{\partial \bar{x}}{\partial x}\right) dx^1 \wedge \cdots \wedge dx^n$$

Notice how this transformation law is completely foreign to the ones we have seen so far (which, on $T_n(M)$, would involve the coordinate basis elements $dx^{i_1} \otimes \cdots \otimes dx^{i_n}$. See the discussion on transformation laws in §2.1 below).¹² Anyway, if we insist, for example, that we have an n -form ω on the manifold given in a particular coordinate chart as $\omega = f dx^1 \wedge \cdots \wedge dx^n$ remain invariant, then

$$\omega = f dx^1 \wedge \cdots \wedge dx^n = \bar{f} d\bar{x}^1 \wedge \cdots \wedge d\bar{x}^n = \bar{f} \det\left(\frac{\partial \bar{x}}{\partial x}\right) dx^1 \wedge \cdots \wedge dx^n$$

it follows that, inverting, $\bar{f} = f \det\left(\frac{\partial x}{\partial \bar{x}}\right)$. On the other hand, in $C^\infty(M)$ the transformation law is $f = \bar{f}$ by definition. This means even though an element of $\Omega^n(M)$ is represented by *one element*, its “single coordinate” still changes under transformations (thus defeating the notion that perhaps n -tuples have to change but singletons do not); further, this really distinguishes the two line bundles Λ^0 and Λ^n . Now for our canonical compare and contrast example.

1.5. Example (Wrong way to make an n -form on M). Let's show the analogous non-example to the one we did for the vector transformation laws above: given $g \in C^\infty(M)$, suppose we decide to define a global n -form ω by naively defining in every chart $\omega = g dx_\alpha^1 \wedge \cdots \wedge dx_\alpha^n$ where the x_α 's are charts covering M . This is the wrong thing to do as ω is seriously ill-defined unless x_α is the only chart.

1.6. Example (One right way to do it). What we can do instead is have a bunch of g_α 's that are related in the overlapping charts by the determinant transformation law just mentioned:

$$g_\beta = g_\alpha \det\left(\frac{\partial x_\alpha}{\partial x_\beta}\right)$$

on $U_\alpha \cap U_\beta$, which is very distinct from the silly condition $g_\beta = g_\alpha$ that would give rise to the preceding example.

¹²Well, some physics texts do define odd transformation laws, but they are never explained there in depth, nor developed in detail in math texts. It is our purpose to bring this out.

Elements of Λ^n , with the transformation law that includes a determinant, are called **densities**¹³, and can be used to construct some more general fancy-schmancy densitized objects. The reasoning for this term is that if we take a scaling transformation, say the new coordinates are all double the old ones, then $\det\left(\frac{\partial x}{\partial \bar{x}}\right) = 2^{-n}$ and the new \bar{f} will be 2^{-n} times the original f in order for $\bar{f}d\bar{x}^1 \wedge \cdots \wedge d\bar{x}^n$ to represent the same object. That is it scales inversely as a volume. Hence the term density. We will have a lot more to say about this guy, and transformation laws, later. We must admit, here, that calling an n -form a density is not actually the right thing to do; we only call it this way because it is sort of already standard, i.e. what has been called a tensor density over the years. One should really say that the (natural) *component* of an n -form is the density; this is again a holdover from the days when people thought that tensors *were* their components and nothing else. There will be more of a rant on this when we introduce these guys in greater generality.

1.4. Contravariance and Covariance. We have been usually avoiding the terms *contravariant* and *covariance* in preference to the more hip, modern notations $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ to indicate types. However I have resisted retitling this set of notes as “ $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$, $\Lambda^n \otimes$, and All That” (as Harry Schey still does not retile his book, now in its 4th edition, as “ $\nabla \cdot$, ∇ , $\nabla \times$, and All That.”) The terms contravariant and covariant are sort of relic terms from the bad old days and are confusing. The original model was that covariance means that decreasing the scale of space along the direction of a such a tensor increases the magnitude, whereas contravariant ones get larger along with space. That sounds backward, right? You’d expect covariant to mean that it *varied with* the space, and contravariant to mean it *varies against*. The reason for this is because components were the main thing people thought of back in the day, rather than invariant vectors, so when scaling the space, one would change the size of the *basis vectors*, and so the components, whose transformation properties are opposite to those of bases, in actuality do have the right correspondence in scaling and so forth. A better picture of what is going on is that covariant things have “plane-type” directionality, in which direction is given by the *transverse* dimensions (like the planes of a stack “facing” in its direction) whereas contravariant things have have “line-type” directionality, i.e. direction is given along whatever they are, such as an arrow “pointing.” This is all pretty vague without metric notions, so just think of covariant objects having direction arising from the object’s orthogonal complement (i.e. being oriented by a suitably generalized notion of “normal vector”), and contravariant ones as things whose direction arises from the space spanned by the object itself (“tangency”). Basically it’s much harder to picture in cases apart from 1- and $(n - 1)$ -degree objects; “tangential” vs. “transverse” directions seem to me the best way to think about it. This does go somewhat awry with densitization, although there is a nice explanation for what happens here, at least for densitization of m -vectors (see §2.2).

The modern mathematical terminology “covariant” and “contravariant” pertains to the category-theoretic notion of *functors*, meaning whether some map of manifolds induces a map of some corresponding space going in the same direction oppositely. It so happens that maps of manifolds *do* induce maps on tensor spaces, and the relationship is also backward from the old terminology: vectors “push forward” which means that a map from M to N induces a map of tangent bundles going forward $TM \rightarrow TN$ —a terminology that says T is a *contravariant* functor, but forms (actually form *fields*) “pull back,” that is the map from M to

¹³Some authors (e.g. Lee in [Lee02]) use density to mean a closely related object, which includes an absolute value in the above transformation law. This alternative type of density is really a pseudo-form, used to integrate over nonorientable manifolds. More on the interplay between pseudo-ness and densitization later.

N induces a backwards map $\mathcal{T}^*N \rightarrow \mathcal{T}^*M$ (there is a slight weirdness that only form *fields* can pull back, but the totality of the bundle itself cannot pull back, whereas, vector fields cannot push forward in general, but the whole bundle can).

Finally, physicists have a term “general covariance” which describe that the laws of physics are independent of coordinate systems. They should actually call it “general *invariance*.” The coordinate components of a covariant vector is a great demonstration of what is *not* general covariance!

1.5. Geometric Quantities? There is a major complaint that I have in terminology. It is fairly standard to say that when a whole bunch of indexed quantities transform in a certain way, and it coincides with the tensor transformation law, these indexed quantities define a **geometric quantity**. This is misleading. For example, the Christoffel symbols Γ_{ij}^k are nontensorial but it represents the mother of all this geometric business: the covariant derivative operator. You can’t get more geometric than that. Let’s clarify for a moment what distinguishes tensors. We can talk about tensor fields in yet another way.

Currently we are conceiving of a tensor field as a map $T : M \rightarrow T_\ell^k(M)$ a section of the tensor bundle or in less snobbish terminology: give T a point, and it will sprout a tensor at p . That is, to write things out functionally explicitly, given vector fields X_1, \dots, X_ℓ and 1-forms $\omega^1, \dots, \omega^k$, $T_p = T(p) \in T_\ell^k M_p$ can act on all these guys: $T(p)(X_1(p), \dots, X_\ell(p), \omega^1(p), \dots, \omega^k(p))$. But remembering our usual set-theoretic law, similar to “Replace)(with a comma” to our tensor identifications, we do another trick: move the (p) all the way to the right: We have map

$$\tilde{T} : (\mathcal{T}(M))^\ell \times (\mathcal{T}^*M)^k \rightarrow C^\infty(M)$$

given by

$$\tilde{T}(X_1, \dots, X_\ell, \omega^1, \dots, \omega^k)(p) := T(p)(X_1(p), \dots, X_\ell(p), \omega^1(p), \dots, \omega^k(p))$$

Again, a standard result of the theory is that T is a smooth tensor field if and only if its associated \tilde{T} actually defines a smooth function of p (i.e. it is well-defined as a map into $C^\infty(M)$). Clearly \tilde{T} is an \mathbb{R} -multilinear map. But in fact, it is much more than that: it is multilinear over $C^\infty(M)$: that is for any function f , we have $\tilde{T}(\dots, fX, \dots)(p) = f(p)\tilde{T}(\dots, X, \dots)(p)$. As it turns out, this characterizes tensors completely (given as an exercise in [Lec97]):

1.7. Theorem. *Let $S : (\mathcal{T}(M))^k \times (\mathcal{T}^*(M))^\ell \rightarrow C^\infty(M)$ be a map. Then S arises as \tilde{T} of some tensor field if and only if S is multilinear over $C^\infty(M)$.*

Proof. We have already seen the “only if” direction. So now what we must do is, given S multilinear over $C^\infty(M)$, define a smooth section $T : M \rightarrow T_k^\ell(M)$ such that $\tilde{T} = S$. It suffices to define $T(p)$ on the tensor product of basis (co)vectors, so we simply define $T(p)$ on the basis fields at p to be S of the basis fields. We have to show that this is in fact independent of the choice of field, however. This uses Taylor’s theorem with remainder [details to be worked out here later]. Therefore yields something well-defined pointwise and smooth on all vector fields and 1-forms because the multilinearity over C^∞ which make all the smooth functions pull out; and the remaining evaluation on the basis vectors yield the smooth function of p given by S . Therefore T is *smooth*. By definition, then, $\tilde{T} = S$. \square

Note this says something important about global tensor fields, that the multilinear maps that tensors define only are dependent on the point and not the values of any of its arguments in neighborhoods of a point. So in particular it depends only on values of vector fields themselves and not their derivatives at p .

Incidentally, readers who know module theory will realize that, since $\mathcal{T}(M)$ is a module over $C^\infty(M)$ and $\mathcal{T}(M)^* \cong \mathcal{T}^*(M)$, T being a multilinear map over $C^\infty(M)$ in all those variables means that T is a *single tensor* (as opposed to a tensor *field*) over the module $\mathcal{T}(M)$. So in the most weird, twisted mathematical way, physicists saying tensors when they mean tensor fields is justified. A tensor field is just a larger kind of tensor (übertensor?) defined over some humongous modules now.

Let us apply this now to our observations about how Christoffel symbols define a geometric quantity. Recall that Γ_{ij}^k define the Levi-Civita connection ∇ just as any other 3-indexed quantity would define a tensor, except that this one isn't a tensor. What this means is that,

$$\nabla : \mathcal{T}(M) \times \mathcal{T}(M) \rightarrow \mathcal{T}(M) \cong \mathcal{T}(M) \times \mathcal{T}(M) \times \mathcal{T}^*(M) \rightarrow \mathbb{R}$$

is not in fact multilinear over C^∞ . But we already know that: in $\nabla_X Y$, it's multilinear over C^∞ in X but not in Y ---it satisfies the Leibnitz Rule there instead. As you can see, there are geometric quantities much broader than the multilinear maps over C^∞ . Another thing to observe is that independence of the values of the derivatives at p stresses that tensors only observe first-order effects. This is in accordance with the whole approximation-by-linear maps or by first-order effects philosophy of physics. The covariant derivative does not fit in with this because it is a *derivative* so it by definition depends on neighborhoods of a point, as you'd expect.

1.6. Infinitesimals vs. Linear Functionals, Volume Forms, and Rethinking Integration. It is often noted, for pedagogical reasons, in first-year calculus, that dx and dt and so forth should be thought of as “infinitesimal quantities,” and then using this notion to justify the concept of integration as “summing up” these things and differentiation as “ratios” of these things. Trying to make this notion precise has led to a big overhaul in mathematical thinking, including the invention of the concept of limit and so forth. . . much to the dismay of many physicists and indeed many who would like a more intuitive explanation of things (people have argued passionately over which method is more intuitive). Finally, Weierstraß won out and infinitesimals were completely banned from modern calculus teaching. As a result of this legacy, the notation for differentiation and integration has some weird holdovers, and the problem is especially acute when learning about differential forms. We are told that we should think of dx and so forth as linear functionals. But one thing tends not to get mentioned: why is it better thought of as a linear functional, and what is its relation to the old way of doing things. It so happens that calculus with infinitesimals *has* been made mathematically precise by Abraham Robinson, using some hardcore 20th-century set theory (which is *not* just your garden-variety “naïve” set theory that is the bread and butter of all working mathematicians), so these things have an air of legitimacy that they lacked in Newton's and Leibniz' time, but nevertheless the stigma remains.

So what *should* we think of dx as, really? An infinitesimal? Or a linear functional? Matters are especially confusing here since if dx means infinitesimal displacement, dx should be a small vector. Yet we are told in tensor analysis: *NO, NO, NO, NO, NO!!! dx is a covector, and unless you have a metric, it can't be identified with vectors canonically!! It belongs to the dual space of vectors!* As a 21st century mathematician, I must concur that dx as a linear functional is necessary and in fact, proper. But let me say why. A possible question that arises in learning differential forms is, *Why do n-vectors ($\Lambda_n(M)$) get so little press? Why is integration so covariantly biased?* The truth of the matter is, integration actually puts both n -vectors and n -forms on equal ground: how, after all, does an integral sign manage to get rid of the dx 's density units anyway?

The secret lies in that very important piece of information that we call the *domain of integration*. Recall that in integration, we break up the domain into little pieces and evaluate a function on each piece, and add up all the results. Now here is the concept. In geometrical physics we are told to think that tangent vectors are the fabled infinitesimal displacements we seek. In the one-dimensional integral, we can then think of breaking the domain up into infinitesimal displacements (tangent vectors), head-to-tail, from start to end. However tangent vectors are not *numbers*. In integration we want to get a number at the end. What dx does, as a linear functional, is *assign a real number to the infinitesimal displacement*, which we should call the *oriented length* of the tangent vector. In n -dimensional integration, one can consider breaking the domain up into little cells or “volume elements.” But what are little cells, anyway? One possibility for good definition of “cell” is a *parallelepiped*. But as we know, parallelepipeds are determined by the vectors along their edges. In fact, it is the wedge product of those vectors, i.e. *parallelepipeds are n -vectors*. And what does an n -form do to a parallelepiped? *Assign it a real number*, exactly in accordance to what forms should do. So, for example, $dx^1 \wedge \cdots \wedge dx^n$ is defined as something that gives the coordinate cell $\frac{\partial}{\partial x^1} \wedge \cdots \wedge \frac{\partial}{\partial x^n}$ the number 1. However, there is still a problem that sticks: We’d expect the volume of a parallelepiped to always be positive, and that’s where the pseudoforams come in. We will get to this point later, but for now, bear in mind that the following motivation for the methods we use to define integration on manifolds is not, strictly speaking, correct, but most of the comments still apply.

In the theory of integration of differential (pseudo) forms, the major shift in conceptual thinking occurs when we rethink the integral

$$\int_a^b f(x) dx$$

from being the two “disconnected” pieces $\int_a^b dx$, the integral operator, and the function f , to, instead, the two pieces \int_a^b and the 1-form $f(x) dx$. What is $f(x) dx$ really, then? One can think of $f(x)$ as being a weighting function for a “standardized” object dx .

In multivariable calc, this two-way thinking gets confusing, especially if one has had advanced analysis (the development of integration from the viewpoint of measure theory, called **Lebesgue integration**). One gets ingrained with the idea that *functions* are *the* objects of choice to be integrated, over the whole space (the first kind of reading of $\int f(x) dx$), for example, and all this integration with other kinds of objects such as differential forms seems somewhat silly. How do we reconcile this with geometry? The answer is once again that coordinate charts tend to get in the way. In more than one variable, and on a manifold, the conundrum becomes:

$$\int_M f(x) dx^1 \wedge \cdots \wedge dx^n$$

being read as $\int_M dx^1 \wedge \cdots \wedge dx^n$ and $f(x)$, or \int_M and a single object, the n -form $f(x) dx^1 \wedge \cdots \wedge dx^n$. Indeed, in order to get away from coordinate-thinking, a general n -form is one object ω which, as we have seen, when coordinatized, happens to show up as some function f times the usual $dx^1 \wedge \cdots \wedge dx^n$ and it can always be expressed this way. Here we are saying f is a weighting function for a standardized, coordinate-dependent object $dx^1 \wedge \cdots \wedge dx^n$. Therefore as we integrate over a coordinate patch, we insert standardized infinitesimal cells represented by $\frac{\partial}{\partial x^1} \wedge \cdots \wedge \frac{\partial}{\partial x^n}$ into the standard coordinate n -form, which gives some number, and then multiply by the function evaluated at the point to get a new number, and finally totals everything up. The problem here, of course, is that it only works

for coordinate charts. Which is not bad if you're developing integration on \mathbb{R}^n which is the space of choice even for die-hard analysts. But suppose are given an object called a *volume form* on the manifold M , namely a *globally defined* choice of preferred n -form. Then at each point, we call this object a "standard." We can then declare that the integral of an honest function f over the manifold to be the integral of f against this chosen form. And now it's guaranteed to be coordinate-independent (modulo sign issues---again, we'll explain it in the section on pseudotensors). For orientable Riemannian manifolds (why only *orientable*, of course, is a question we also hope to address, because obviously we can do surface integrals over nonorientable things like the Möbius strip---it has an area, after all), the metric is used to select this preferred n -form on a manifold, which immediately reduces to the usual integration of functions when applied to \mathbb{R}^n and the Euclidean metric. This volume form (or its pseudo-version) is also what is used to define stuff like surface area and the traditional scalar line integral (as opposed to the vector-field or covector line integral) which describes things like energy contained in the tension of a string or soap film.

What is the upshot of the discussion, in terms of infinitesimals? Well, integration of a *form* is now more simply expressed as breaking up the domain into little cells, described by a frame of n -vectors, say, feeding it into the form we're integrating, taking the number it yields, and totalling them all up. Period. No multiplication by anything.

For integrals of *functions*, this occurs when we want to compute some "global" quantity, out of another quantity described as "local" or "microscopic," or "per unit volume," whose precise nature is encoded in this function f , *provided that we have given a standard reference*, we then chop up the domain as usual, feed the cells into the standard reference form, then taking the number it yields and multiplying it by the weight that f assigns at that point, and finally totals everything up.

One other thing one should speak of is calculating the volume (or surface area) of a manifold itself. This is another can of worms, because of some ingrained prejudices again. First off it should be obvious from the get-go that the everyday volume of a manifold is a metric concept, i.e. one should integrate the Riemannian volume form to get what we want (this is expressed, intuitively, as the general principle that n -volumes should scale as the n th power of *distance*). This is going to conflict with the notion of volume element (n -vectors) because when inserting volume elements into the form, we get a scalar. So is the volume of a manifold a scalar, or should it be an n -vector which transforms with $\det\left(\frac{\partial \bar{x}}{\partial x}\right)$?

We will see that volume forms are useful for much more than just allowing us to define integrals of functions later on when we develop tensor densities and their pseudo-version more thoroughly. In particular is the notion of *flux* which takes this local-to-global functional description to another level. The notion of flux brings up even more interesting and weird problems, because it is integration of a form of lower degree, over a submanifold with corresponding dimension. However, the interplay between orientability of submanifolds and the orientability of the ambient manifold is not entirely straightforward (just think about the facts that the Möbius strip, and the fact that it can be embedded in \mathbb{R}^3 , but the Klein bottle, a closed nonorientable surface, cannot).

2. TRANSFORMATION LAWS AND THEIR RELATION TO FANCY-SCHMANCY TENSORS

2.1. Transformation Laws Again. Although we touched on this above in the section on tensors vs. their fields, it is worth repeating that the standard $\binom{k}{\ell}$ -tensor transformation laws talk about changing the standard *coordinate* bases, that is, the coordinate system (x^i)

determines the bases $\left\{ \frac{\partial}{\partial x^i} \right\}$ and $\{dx^j\}$ and the components of a $\binom{k}{\ell}$ tensor are given as the coefficients with respect to the basis of all possible tensor products of k $\frac{\partial}{\partial x}$'s and ℓ dx 's. Just this once, I'll write the whole delightful transformation in coordinates just to say precisely what I mean:

$$(2.1) \quad \bar{T}_{i_1 \dots i_\ell}^{j_1 \dots j_k} = T_{r_1 \dots r_\ell}^{s_1 \dots s_k} \frac{\partial \bar{x}^{j_1}}{\partial x^{s_1}} \cdots \frac{\partial \bar{x}^{j_k}}{\partial x^{s_k}} \frac{\partial x^{r_1}}{\partial \bar{x}^{i_1}} \cdots \frac{\partial x^{r_\ell}}{\partial \bar{x}^{i_\ell}}$$

Gee, I really hope I got that right. For some slightly nonstandard terminology, I shall refer to the whole combination of partial derivative factors above as **junk**. Now, what I am repeatedly emphasizing is what the components $T_{i_1 \dots i_\ell}^{j_1 \dots j_k}$ are actually the coefficients of the tensor with respect to that certain very special basis, the coordinate basis. $T_{i_1 \dots i_\ell}^{j_1 \dots j_k}$ is paired with the basis tensor consisting of all the guys tensored together, namely $dx^{i_1} \otimes \cdots \otimes dx^{i_\ell} \otimes \frac{\partial}{\partial x^{j_1}} \otimes \cdots \otimes \frac{\partial}{\partial x^{j_k}}$ and not some other funky basis derived from the coordinates. That is to say, given a tensor \mathbf{T} defined by all those components, we have that, using the summation notation in full force,

$$(2.2) \quad \mathbf{T} = T_{i_1 \dots i_\ell}^{j_1 \dots j_k} dx^{i_1} \otimes \cdots \otimes dx^{i_\ell} \otimes \frac{\partial}{\partial x^{j_1}} \otimes \cdots \otimes \frac{\partial}{\partial x^{j_k}}.$$

However, as we have seen with differential forms, even though they lie in spaces of tensors of high covariant rank, the actual dimension of the space containing all such forms is reasonably-dimensional, and in fact becomes smaller (recall the **Grassmann algebra** is the graded algebra $\Lambda^*(M)_p$ of all forms on M at p with the wedge product, corresponding to the direct sum of all the individual alternating tensor spaces; above each point, this space is always finite-dimensional, of dimension 2^n ; and the dimension of the space of grade k is given by the combinatorial symbol $\binom{n}{k}$.) In other words, on such a space, there is a more natural basis to work with, namely the basis

$$\mathcal{B}_k = \{dx^{i_1} \wedge \cdots \wedge dx^{i_k} \mid 1 \leq i_1 < \cdots < i_k \leq n\}.$$

We will get to a very concrete example in the next section and you can skip there now if you like, but here I'll sketch the general big picture of what is going on with funny bases. Given some q -dimensional subspace of, say, $T^k(M)$ we could suppose, in theory, we have q basis elements $B^I(x)$, with some exotic multiple-index scheme indicated by the I , which are functions of the coordinates; and an element T of this space can be expanded as $T_I(x)B^I(x)$ (using a suitably generalized summation convention). Then, changing coordinates, we have $\bar{T}_J(\bar{x}) = Q_J^I(x, \bar{x})T_I(x)$ where the Q_J^I 's are some transition functions that juggle lots of indices. The fact is, even though the B^I are derived from coordinates, they will not, in general, actually be coordinate basis elements of the form used in (2.2) above. Therefore it is imperative to realize that *the usual transformation laws do NOT apply!* Let's suppose we have an n -dimensional subspace; and the B^I some exotic linear combination of the dx^i 's which depend on the coordinates in some systematic way. Then the transformation law

$$\bar{T}_i = Q_i^j T_j$$

does not imply, just because the space is n -dimensional and the tensors are covariant, that the Q_i^j have to equal the usual $\frac{\partial x^i}{\partial \bar{x}^j}$. We have already seen one example of this, where the index set is completely empty---i.e. "scalar-like" spaces \mathbb{R} and Λ^n in our $\Omega^0(M)$ vs. $\Omega^n(M)$ example. Here the transition function for Ω^0 is just the identity---functions satisfy $\bar{f} = f$

in any coordinates. But the transition function for Ω^n is $\det\left(\frac{\partial x}{\partial \bar{x}}\right)$. We will exploit this fact, using noncoordinate, but coordinate-related bases, to show some strange things---for example, objects being both covariant and contravariant at the same time, if viewed in different ways (*without* using a metric).

2.2. **Fancy-Schmancy Tensors: Densitization.**

“Hi, I’m George. I am your density. Err, I mean, destiny.”---Crispin Glover
(as George McFly) to Lea Thompson (as Lorraine Baines) in *Back to the Future*.

We finally come down to the real deal and the real purpose for putting these notes together. Physics texts often introduce the concept of tensor density with a strange transformation law and subsequently never give an invariant conception of them (i.e. always talking about them as components), which mathematicians then dutifully ignore. Let us recall the coordinate-messy definition now.

2.1. **Definition.** A $\binom{k}{\ell}$ **tensor density** with **weight** w is a quantity whose components transform as follows:

$$(2.3) \quad \bar{T}_{i_1 \dots i_\ell}^{j_1 \dots j_k} = T_{r_1 \dots r_\ell}^{s_1 \dots s_k} \det\left(\frac{\partial x}{\partial \bar{x}}\right)^w \frac{\partial \bar{x}^{j_1}}{\partial x^{s_1}} \cdots \frac{\partial \bar{x}^{j_k}}{\partial x^{s_k}} \frac{\partial x^{r_1}}{\partial \bar{x}^{i_1}} \cdots \frac{\partial x^{r_\ell}}{\partial \bar{x}^{i_\ell}}$$

that is, exactly like the usual law in (2.1) except for an extra factor of $\det\left(\frac{\partial x}{\partial \bar{x}}\right)^w$ (recall that the notation denotes the Jacobian matrix of the coordinate change) preceding the usual junk. The weight can range over all integers, positive and negative (although the cases of interest usually are just 1, -1, -2, and of course zero, which gives us back our old tensors). Negative-weight tensor densities might also be referred to as **tensor capacities**, although I think I’ve made this one up; the only thing I have to go by is to generalize the term from the one and only place I’ve seen the word capacity used with tensor concepts: Weinreich’s text [We97]. Occasionally it is defined with the *absolute value* of the determinant which is in fact what happens when we examine the interplay between tensor densities and pseudotensors (namely, adding pseudoness can fiddle with the sign).¹⁴

Now that we have given the boring definition, the burning question you may have is, “What invariant objects have components which could transform in such a manner?” What geometric object makes these come about? Math books won’t tell you (or maybe they will tell you in an exercise, as in Spivak’s otherwise great reference, [Sp03]). We have already seen one case in the above example, and this is the most important case: components where $w = 1$ (one determinant factor) and $\binom{k}{\ell} = \binom{0}{0}$ (no junk), namely, a **scalar density**. We saw that the object giving rise to these guys was the line bundle Λ^n . I say this is perhaps the most important case, because, all other density-related objects (except the absolute values) can be obtained from this. One can “force” something to be a density by a process called *densitization* which simply takes the tensor product of Λ^n with everything else (mathematicians try to make it sound really cool by calling it **twisting by the determinant line bundle**). So we shall give a more formal nonsense mathematical definition, which says that all of these guys have a somewhat “easy” invariant formulation:

2.2. **Theorem.** *Consider the tensor bundle $T_\ell^k(M)$. Let $w \geq 0$. Then $T_\ell^k(M) \otimes (\Lambda^n(M))^{\otimes w}$ (where the notation $V^{\otimes w}$ means the w -fold tensor product of V with itself) is a bundle of $\binom{k}{\ell}$ -tensor densities of weight $w \geq 0$. If $w < 0$, we can define $T_\ell^k(M) \otimes (\Lambda_n(M))^{\otimes -w}$, which yields the corresponding densities*

¹⁴With the absolute value, it is possible to even define weights for all real numbers.

of negative weight. More concretely, tensor densities of weight w are precisely multilinear mappings taking k 1-forms, and ℓ vectors and yielding w -fold densitized scalars.

Notice that since Λ^n is rank-1, taking additional tensor products with it does not change the rank. No matter what w is, the multilinear mappings always end up spitting out funky modified scalars, but they really still are souped up 1-dimensional objects. Don't let the w scare you into thinking we have gotten into the realm of vector-valued forms or something like that (those actually are useful, by the way---Cartan developed a whole theory of objects like that, involving the whole moving frames deal. I haven't had much time to study them though).

We motivate our next example as follows. The magnetic field "vector" is described, in more advanced texts as actually being a pseudovector. But in fact this is not what the magnetic field most naturally is. It turns out that the magnetic field is most properly conceived as a 2-form, which, as we shall see, is not a pseudo- anything, but rather a $\binom{1}{0}$ -tensor density. What, you say? Haven't you always known a 2-form as a $\binom{0}{2}$ totally antisymmetric genuine tensor i.e. weight-0 density?¹⁵

2.3. Example (The Magnetic Field). First off, let us see why the pseudovector issue is resolved by a 2-form. Given the magnetic "vector" field, a triple $\mathbf{B} = (B^x, B^y, B^z)$ we consider the 2-form given by

$$\mathcal{B} = B^x dy \wedge dz + B^y dz \wedge dx + B^z dx \wedge dy.$$

This should actually be a vaguely familiar construction, as taking curls and using Stokes's theorem and such involves taking a coordinate and pairing it with the excluded coordinates in the somewhat bizarre xy , yz , and zx order. Recall the naïve definition of a pseudovector, namely, when all the coordinates flip (considering only orthogonal transformations, since nonrelativity physics texts are always so enamored with Cartesian tensors), we get $\bar{B}_i = B_i$ rather than, for regular vectors, $\bar{A}_i = -A_i$. You see, if the magnetic field were really a true vector, then in coordinates it would have the expression $B^x \frac{\partial}{\partial x} + B^y \frac{\partial}{\partial y} + B^z \frac{\partial}{\partial z}$ and we know that each of the $\frac{\partial}{\partial x^i}$ guys morphs into its negative when flipping. Hence the corresponding components must also change sign. However, flipping the coordinates morphs the dx^i guys into their negatives, too. Since the 2-form has them all wedged together in pairs, though, the two negatives cancel each other out. So the corresponding component remains the same. So this looks good. Another perhaps more convincing reason why the 2-form works is by evaluating \mathcal{B} on a single genuine tangent vector $v = v^i \frac{\partial}{\partial x^i}$ (and using its remaining slot to make the result a 1-form). Note that the above definition the B_i are not the genuine $(0, 2)$ -tensor components of \mathcal{B} , because first there are only three of them, while there would be nine (doubly-indexed) quantities that would give us the components of $(0, 2)$ -tensor. We can ask, though, what are the genuine $(0, 2)$ -tensor components of \mathcal{B} . These are simply given by $B_{xy} = -B_{yx} = B^z$, $B_{yz} = -B_{zy} = B^x$, and $B_{zx} = -B_{xz} = B^y$, and finally $B_{xx} = B_{yy} = B_{zz} = 0$. This is just an expression of the fact that the Λ^m 's are spaces of

¹⁵ It turns out that in the case of *orthogonal* transformations, pseudovectors and $\binom{1}{0}$ -tensor densities coincide, because the determinant is reduced to being ± 1 ---and orientation-reversing ones give the -1 that cause everything to be pseudo. That is, the books calling magnetic fields a pseudovector are not only avoiding the use of differential forms, but also restricting attention to *Cartesian tensors*. . . which I really haven't had the time to really research in depth. It turns out all tensors are Cartesian tensors, but not all Cartesian tensors are tensors. One would think it was the other way around.

antisymmetric tensors. More succinctly, we can write the components of B in matrix form:

$$\begin{pmatrix} 0 & B^z & -B^y \\ -B^z & 0 & B^x \\ B^y & -B^x & 0 \end{pmatrix}$$

which might ring a bell. Now let's stick qv into \mathcal{B} 's first slot. We find $\mathcal{F} = \mathcal{B}(qv, \cdot) = B_{ij}qv^j dx^i$, or that $F_j = \mathcal{B}(qv, \cdot)_j = qB_{ij}v^i$. But that just says that the triple $\mathbf{F} = (F_x, F_y, F_z)$ is the matrix product of that antisymmetric matrix and the triple $q\mathbf{v} = (qv^x, qv^y, qv^z)$:

$$\begin{pmatrix} F_x \\ F_y \\ F_z \end{pmatrix} = \begin{pmatrix} 0 & B^z & -B^y \\ -B^z & 0 & B^x \\ B^y & -B^x & 0 \end{pmatrix} \begin{pmatrix} qv^x \\ qv^y \\ qv^z \end{pmatrix} = \begin{pmatrix} qv^y B^z - qv^z B^y \\ qv^z B^x - qv^x B^z \\ qv^x B^y - qv^y B^x \end{pmatrix}$$

that is, the triples are related by $\mathbf{F} = q\mathbf{v} \times \mathbf{B}$, the plain old Lorentz force law with the cross product! In the modern notation this means $\mathcal{F} = \mathcal{B}(qv, \cdot)$ which yields force as a 1-form, actually¹⁶. This actually is natural because in electromagnetism, F is a quantity that always gets (line-)integrated to get the energy.

Note that in the above example we have two ways of characterizing the magnetic field 2-form: its $\binom{0}{2}$ -tensor components B_{ij} , and its natural 2-form basis components, B^i . Why am I writing B^i with the guy upstairs? It turns out that the B^i transform like a $\binom{1}{0}$ -tensor density! It is in this sense B can be both a $\binom{0}{2}$ -tensor and a $\binom{1}{0}$ -tensor density: it all depends on how we define *components* and hence what kind of basis we are choosing. Here, we shall see, we get an exotic transformation law by forgoing the usual definition of a basis element as given by \cdot . In general, $\Lambda^{n-1}(M) \cong T^*(M) \otimes \Lambda^n(M)$ that is all $(n-1)$ -forms can be realized as n -dimensional $\binom{1}{0}$ -tensor densities (with weight 1).

We can prove this two ways: the abstract nonsense way (efficient---it actually will give us information on the ways we can identify *all* the forms on manifolds with certain densities) and the messy way (which actually does give some information on what's going on). First, the messy way, for $(n-1)$ -forms. In Λ^n we take as a basis

$$\mathcal{B}'_k = \{(-1)^{i+1} dx^1 \wedge \cdots \wedge \widehat{dx^i} \wedge \cdots \wedge dx^n \mid i = 1, \dots, n\}$$

where the notation $\widehat{dx^i}$ means to omit that term in the product. This is almost the usual \mathcal{B}_k as defined before, except for a factor of $(-1)^{i+1}$ preceding each. This is analogous to choosing the cyclic permutations $dx \wedge dy$, $dy \wedge dz$, and $dz \wedge dx$ in n dimensions. For brevity, we write ξ_i to be the i th basis vector in this set (note the lower index).

2.4. Example (Messy Computation of $\Lambda^{n-1} \cong \Lambda_1 \otimes \Lambda^n$). We would like to show that given some $B \in \Lambda^{n-1}$ and expanding it in terms of this natural basis, $B = B^i \xi_i$ (and *not* the usual $(n-1)$ -rank tensor basis elements), that these B^i satisfy the transformation law

$$\bar{B}^i = \det \begin{pmatrix} \frac{\partial x}{\partial \bar{x}} \end{pmatrix} \left(\frac{\partial \bar{x}^i}{\partial x^j} \right) B^j$$

The proof is given by using the classical adjoint formula for the inverse of a matrix. Recall that for an invertible matrix $A = (A_{ij})$, we have

$$(A^{-1})_{ij} = \frac{1}{\det A} (-1)^{i+j} \det(M_{ji})$$

¹⁶This operation of inserting into the first slot and leaving the remaining blank is done quite often; the standard notation for this is $\iota_{qv}(\mathcal{B})$ or $qv \lrcorner \mathcal{B}$ (a notation I'm not too fond of, grr), the **interior product** of qv and \mathcal{B} .

where M_{ij} is the $(n-1) \times (n-1)$ minor matrix with the i th column and j th row deleted. Now, we try to write out the basis element $\bar{\xi}_i = d\bar{x}^1 \wedge \cdots \wedge \widehat{d\bar{x}^i} \wedge \cdots \wedge d\bar{x}^n$ in terms of the old basis. The j th component in terms of the old basis is given by evaluation on the vectors $\frac{\partial}{\partial x^\ell}$ but excluding the j th component, and multiplying by $(-1)^{j+1}$ because of that darn sign flip; plus we multiply the whole thing by $(-1)^{i+1}$ when expanding ξ_i using formula (0.1):

$$\begin{aligned} \bar{\xi}_i \left(\frac{\partial}{\partial x^\ell} \right)_{\ell \neq j} &= (-1)^{i+1} \det \left(d\bar{x}^k \left((-1)^{j+1} \frac{\partial}{\partial x^\ell} \right) \right)_{\substack{k \neq i \\ \ell \neq j}} \\ &= (-1)^{i+j} \det \left(\frac{\partial \bar{x}^p}{\partial x^\ell} d\bar{x}^k \left(\frac{\partial}{\partial \bar{x}^p} \right) \right)_{\substack{k \neq i \\ \ell \neq j}} = (-1)^{i+j} \det \left(\frac{\partial \bar{x}^k}{\partial x^\ell} \right)_{\substack{k \neq i \\ \ell \neq j}} \end{aligned}$$

where the first equality is given by formula (0.1). But the rightmost term is just the (i, j) th minor of the Jacobian $\left(\frac{\partial \bar{x}}{\partial x} \right)$ with the factor $(-1)^{i+j}$. In other words, this is the entry in the cofactor matrix of the inverse. This is to say that the (j, i) component of the inverse times the determinant $\det \left(\frac{\partial \bar{x}}{\partial x} \right)$. In other words,

$$\bar{\xi}_i \left(\frac{\partial}{\partial x^\ell} \right)_{\ell \neq j} = \det \left(\frac{\partial \bar{x}}{\partial x} \right) \left(\frac{\partial x^j}{\partial \bar{x}^i} \right)$$

It therefore follows that if $B = B^i \xi_i = \bar{B}^i \bar{\xi}_i$ the transformation law we want for the B^i , by inverting, is satisfied. Therefore an $n-1$ form is a $\binom{1}{0}$ tensor density. Despite the apparent ugliness the derivation, it suggests something about all of this: because we have a term with a determinant (which scales in all independent directions) and a single partial of the inverse (therefore scaling in only one direction, but in an opposite manner), the effect is to cancel off one degree. This explains why an $(n-1)$ -fold covariant object can somehow be contorted into a contravariant object: the *densitization* adds n -fold covariance upon an object; hence a singly contravariant quantity, when densitized, has n degrees of covariance added on---namely making it an $(n-1)$ -fold covariant object.

Now let's get onto the more elegant way of proving this. It actually gives isomorphisms of all m -forms to densitized $(n-m)$ -plane elements, and is the starting point of an important duality principle in the study of differential forms, called **Hodge duality**, which is used extensively in electromagnetics (we will mention this briefly in our big duality section).

2.5. Theorem. *There exist canonical isomorphisms $\Lambda^m \cong \Lambda_{n-m} \otimes \Lambda^n$, for $0 \leq m \leq n$. Hence all m -forms can be viewed as suitably densitized $(n-m)$ -plane elements.*

Proof. We have a natural bilinear mapping m -forms and $(n-m)$ -forms into n -forms: the wedge product itself! That is, we write

$$\Lambda^m \otimes \Lambda^{n-m} \xrightarrow{\wedge} \Lambda^n$$

We would like to “move the Λ^{n-m} over to the other side,” in a standard identification that causes the moved space to go to its dual. That is, our mapping should be $\omega \in \Lambda^m$ mapping to $\omega \wedge (\cdot)$ where the (\cdot) indicates a “slot” for evaluation on an $(n-m)$ -form. The funny object $\omega \wedge$ is then something that takes $(n-m)$ -forms to yield n -forms, that is, an element of the dual space of $(n-m)$ -plane elements, but densitized (to be an element of the true dual, it would have to take $(n-m)$ -forms to genuine real numbers).

The only thing that remains here is to show that the mapping is one-to-one, that is, the wedge product is **nondegenerate**. That is, we should check that $\omega \wedge \eta = 0$

for all η implies $\omega = 0$. But if $\omega \neq 0$, then there exists J such that $B^J \neq 0$, where $\omega = B^J \xi_J$, and $\xi_J = dx^{j_1} \wedge \cdots \wedge dx^{j_m}$ is some basis element of Λ^m . Then we just have that dx^J wedged with the complementary, remaining wedged basis 1-forms dx^L . Then $\omega \wedge dx^L = B^J dx^1 \wedge \cdots \wedge dx^n \neq 0$. This means the map $\omega \mapsto \omega \wedge$ is injective, and hence an isomorphism into $\Lambda_{n-m} \otimes \Lambda^n$. \square

The moral of the preceding story is that the phenomena of contravariance and covariance are not mutually exclusive, even when there is no metric or other structure present, but rather, dependent on how one views components. The thing is, the terms contravariance, covariance, and even density are somewhat inaccurate terms for the modern view of tensors (perhaps this is why nobody hears of tensor densities in the math books). These terms more accurately describe the situation when physicists think of vectors and tensors as *being* their components rather than objects which can be represented in some kind of component form. In other words, contravariance/covariance/densitization has more to do with what kind of components one chooses for representations, and not the actual invariant thing it represents. When the components and respective bases are paired up together (in a perfect upper-lower index harmony), as we have noted, we have created something that does not change at all under transformations.

2.3. Visualization of Forms as Tensor Densities. Allright, all the preceding really is nonsense unless we can give some kind of way to visualize these things. First off, n -forms, in the special case $m = 0$ are densitized scalars---exactly what we computed in our main example above. Next, regular scalars are “densitized volumes” which means what you think it means: multiplying a density by a volume cancels everything exactly.

There is no simpler $\binom{n-1}{0}$ -tensor density than the *gradient*, “ df ,” of a function f . Gradient a $\binom{n-1}{0}$ -tensor density? Hasn't one always known the gradient as a 1-form? Yes, indeed, but only because one was not familiar with the more appropriate $\binom{n-1}{0}$ -tensor density concept. This sounds exotic, to be sure, but the $\binom{n-1}{0}$ business just means that df should be represented by hyperplanes ($(n - 1)$ -dimensional spaces), and further, it is their *density* that determines their magnitude. This, in fact, explains Weinreich's picture of a stack---why there seems to be a only a single-directional density---the area of the sheets is immaterial but the spacing between the sheets counts. . . it would, *a priori*, look kind of weird for something purporting to represent an area element, if you ask me. The reason why is that $(n - 1)$ -area elements are in fact almost a full volume. But the true density of Λ^n cancels off all of the “area-like” scaling properties of such a form, and then some.

The canonical map in the above should be called the **Weinreichian dual**. There is also a very nice visualization of an $(n - 1)$ -form which we arrive at via the identification of Λ^{n-1} with $\Lambda_1 \otimes \Lambda^n$. . . which is essentially Faraday's notion of field lines¹⁷. It is termed a *sheaf* by Weinreich (for the mathematically seasoned, this has nothing to do with the abstract sheaves which describe things like germs of functions, other than that they vaguely have the same kind of picture. Weinreich's term is much more direct). In \mathbb{R}^3 it looks like a bundle of arrows all pointing in the same direction; the *areal* density of the arrows in the transverse direction indicating the magnitude [PICTURE] (i.e. it has no measure of magnitude along the direction of its arrows which can be extended as far as one likes, unlike the single arrow case). Weinreich describes a nice way to do scalar multiplication or addition of these guys.

Gradient of a function as a $\binom{n-1}{0}$ -tensor density

¹⁷Field lines are a macroscopic version of what we're describing; whereas a $\binom{1}{0}$ -tensor density would be the linearized or infinitesimal version of field lines.

Once again, note the “density” property going on here. Since sheaves have an extended “line” direction that determines their orientation, scaling along this dimension leaves it invariant, while scaling along all transverse directions has an effect, in sort of an opposite manner that the hyperplane elements are subject to for 1-forms. This reflects the fact that it has a single contravariant rank, and it cancels out one dimension of the densitization.

The $(n - 1)$ -fold product of 1-forms can be thought of as taking $(n - 1)$ 1-forms and intersecting them to form “tubes.” In \mathbb{R}^3 this is taking two stacks and considering the tubes formed by the intersection (MTW’s “egg crates”). In n dimensions, of course, one needs to intersect more forms in order to get tubes, spaces that extend in only 1 dimension. The relation of this to the sheaf is that if we thread lines through the tubes, giving them direction according to the right-hand rule as usual. These form the requisite bundle of lines. However, not all sheaves natively represented this way can be written to be a product like this; it may only be a linear combination.

One final note: one of the first brushes with fancy-schmancy tensors which physicists may encounter is the totally antisymmetric **Levi-Civita permutation symbol**

$$(2.4) \quad \varepsilon_{i_1 \dots i_n} = \begin{cases} 1 & \text{if } i_1, \dots, i_n \text{ are an even permutation of } 1, \dots, n \\ -1 & \text{if } i_1, \dots, i_n \text{ are an odd permutation of } 1, \dots, n \\ 0 & \text{otherwise} \end{cases}$$

which is defined to keep its identity in all coordinate systems. However this restriction causes the components $\varepsilon_{i_1 \dots i_n}$ to not actually transform as a true tensor of rank n . Various authors seem to be confused about this issue and call $\varepsilon_{i_1 \dots i_n}$ by various different names as a densitized or pseudo object, and they can’t seem to decide which. We should see exactly what goes wrong; what happens if we try to do it anyway? Consider the “tensor of rank n ” defined by trying to use $\varepsilon_{i_1 \dots i_n}$ as components:

$$\varepsilon = \varepsilon_{i_1 \dots i_n} dx^{i_1} \otimes \dots \otimes dx^{i_n}.$$

where the summation convention is in full force. If it really defines a tensor, then it has to have the same expression in different coordinates. But all terms with repeated indices are clobbered, and permutations of $1, \dots, n$ contribute a sign. In other words,

$$\varepsilon_{i_1 \dots i_n} dx^{i_1} \otimes \dots \otimes dx^{i_n} = \sum_{\sigma \in \mathcal{S}_n} \text{sgn}(\sigma) dx^{\sigma(1)} \otimes \dots \otimes dx^{\sigma(n)} = dx^1 \wedge \dots \wedge dx^n,$$

which is not, in fact invariant under coordinate transformations: it gains a Jacobian determinant on transformation. How do we kill off a Jacobian determinant? Introduce something of the opposite kind. Namely, we consider the *invariant* object

$$\epsilon = \varepsilon_{i_1 \dots i_n} dx^{i_1} \otimes \dots \otimes dx^{i_n} \otimes \left(\frac{\partial}{\partial x^1} \wedge \dots \wedge \frac{\partial}{\partial x^n} \right) = dx^1 \wedge \dots \wedge dx^n \otimes \left(\frac{\partial}{\partial x^1} \wedge \dots \wedge \frac{\partial}{\partial x^n} \right)$$

It really is coordinate-invariant, because the effects of the coordinate 1-forms are canceled off exactly by those of the coordinate vectors. So we have the following: *If we that $\varepsilon_{i_1 \dots i_n}$ be something invariant, and yet have the same components in every coordinate system, then it must transform as a capacitized covariant tensor (densitized tensor of weight -1) of rank n . In other words, $\varepsilon_{i_1 \dots i_n}$ satisfies*

$$\bar{\varepsilon}_{i_1 \dots i_n} = \varepsilon_{i_1 \dots i_n} = \varepsilon_{j_1 \dots j_n} \det \left(\frac{\partial \bar{x}}{\partial x} \right) \frac{\partial x^{j_1}}{\partial \bar{x}^{i_1}} \dots \frac{\partial x^{j_n}}{\partial \bar{x}^{i_n}}.$$

Similarly the “dual” Levi-Civita tensor, which numerically is identical to the usual one (and hence sometimes no distinction is made) transforms as a contravariant density, and it goes

with the following invariant object:

$$\epsilon^* = \epsilon^{i_1 \dots i_n} \frac{\partial}{\partial x^{i_1}} \otimes \dots \otimes \frac{\partial}{\partial x^{i_n}} \otimes dx^1 \wedge \dots \wedge dx^n.$$

More work to be done:

- Even a degenerate case works out: $\Lambda^1(\mathbb{R}^2)$, where it is both $\binom{0}{1}$ and $\binom{1}{0}$ weight-1 density.
- The metric and its involvement in all this, de-densitization, the determinant of g_{ij} as 2 weight density or 2-weight capacity---next section, actually
- LOTS OF PICTURES.
- clean up

3. PSEUDOTENSORS

Minus times minus is plus;

The reason for which we need not discuss.

--- A classic rhyme detailing one's frustration with rote learning in mathematics.

As we have seen, one case where we might have benefitted in considering pseudotensors is the so-called “magnetic field pseudovector.” But it is resolved completely by considering it as a densitized object instead. And for a long time, I thought that *was* the way to solve the problem: I thought pseudo-ness does not actually exist, it is only a relic of the (bad) old days when everything was a vector field, everyone thought vectors *were* their coordinates, and so forth. But there's a crack in that assumption. It is evidenced by the fact that scalar integrals over submanifolds require orientability, because the basic tool in integration (of functions) that we have is given by the Riemannian volume form, which is only definable for orientable manifolds, because we have to choose a sign for a square root. However, nonorientable objects such as the Möbius strip obviously have area. Note that the volume form also includes a factor of $\sqrt{\det(g_{ij})}$ so these two problems really are one and the same.

What we need to do is somehow force pseudoness onto scalars, and then form a line bundle of pseudoscalars. Then, twisting all our tensor bundles by the pseudoscalar line bundle, we can get all the usual floræ and faunæ of modified tensors, including combining it with densitized guys. So we can get some really exotic objects, and applying it to forms we can get densitized pseudo twisted forms (see [BBF]) and other monstrosities ☺. Once again, I look at the one resource I have, [We97], on pseudoscalars (which is in fact covered, as I shall say in my text reviews, he covers all rank-0 and rank-1 flora and fauna). Intuitively, pseudoscalars are like counterclockwise and clockwise---this is called the “axial sense” as opposed to usual scalars which have “polar sense.” (Polar) scalars split up neatly into positive and negative, while axial scalars can only be either called positive or negative using a convention, such as the right-hand rule (this explains why the cross product is associated with all these handedness things). With pseudoscalars in hand, we can now define a pseudotensor as an honest multilinear mapping that eats vectors and forms and spit out pseudoscalars. “Honest” multilinear mapping, because many authors fix it all up by taking absolute values of forms, which cannot be linear when scaling by negative constants.

Even more weirdly, both Weinreich and Frankel (in [Fr04]) hint at the possibility of mixing and matching these flora and fauna, inserting certain kinds into the slots of a tensor of the other kind: e.g. making pseudotensors eat other pseudotensors, and undoing the pseudoness to spit out genuine scalars, or inserting pseudostuff into slots of genuine tensors, to make them spit out other pseudostuff, and so on. Weinreich describes this when speaking

of taking cross products and dot products with various varieties of pseudo- and nonpseudo-objects. We will resolve this issue also; interestingly the resolution of this, and in fact the whole pseudo-shebang, came from my trying to make sense of *complexifying* bundles, when I was learning about complex manifolds---which are always orientable!

Despite the fact that pseudothings can be constructed via multilinear maps into pseudoscalar; it is nice to actually try to picture what pseudovectors and pseudocovectors look like in \mathbb{R}^3 i.e. what does it mean in terms of arrows, stacks, etc. Ok, enough talk, let's get to action.

3.1. The Plan. What we need to do is take Weinreich's hint (i.e. that axial scalars are like quantities with really funny signs that cannot be identified with usual $+$ and $-$ without making an arbitrary choice---we will call arbitrary choices **non-canonical**), and run with it. There is a bit of mathematical formalism that is cooked up to solve this very issue. Basically, if you want to create objects that cannot be "canonically" identified with $+$ and $-$, create new objects that are distinct from, yet behave exactly like them. What should be these objects be? They should somehow be based in orientations of vector spaces, of course. If a vector space V is oriented, that means we have chosen a certain kind of basis, and declared it "positive." It means, for example, we can give a choice of sign for $\sqrt{\det g_{ij}}$ when making a volume form, at so forth. Obviously, reversing orientation should be akin to multiplying the existing one by -1 (surprise!). In cooking up new objects which behave like $+$ and $-$ and yet are not, we can still allow them to interact with (the genuine) $+$ and $-$. In other words, we want to somehow formalize Weireich's observations:

- Polar times Polar is Polar,
- Axial times Polar is Axial,
- Polar times Axial is Axial,
- Axial times Axial is Polar.

And unlike the old rhyme above, we will *not* abide by

*Axial times axial is polar;
The reason for which merely makes you grow colder.*

One is extremely tempted to say that axial behaves like $-$ and polar like $+$, but do not yield to that temptation! In fact we shall come up with a situation where the opposite is true!

Recall that, one way of construction \mathbb{C} (one could argue, not a very good way) is to "declare by fiat" that $i^2 = -1$, and everything else is just linear combinations of a real number plus i times another real number, and multiplication is then just defined to distribute over addition, including over that factor of i , and then to otherwise be real multiplication (collecting any two factors of i into -1).

We do the same thing to define pseudoscalars (but we'll motivate the "declaration by fiat", so it doesn't look like a rabbit pulled out of a hat---that's why I find that particular definition of \mathbb{C} , fashionable in the style of lean-and-spare popularized by Rudin and the rest during that era, somewhat annoying!). In fact our algebra will have a "unit" which behaves very much like i , except that its square is 1 , instead of -1 . This simple change gives rise to a very different algebra!

Our definition of pseudoscalar (which we shall hereafter abbreviate ψ -scalar) will depend on a particular vector space, although in theory one could just define an abstract space called "the one and only true space of ψ -scalars," sort of like trying to claim that \mathbb{R} is "the one and only true space of real numbers." But the reason why we want to make things depend on a vector space is that we eventually want to have spaces of ψ -scalars varying

over a manifold by using orientations of the underlying tangent spaces; it is this varying that allows us to relate to orientability of the whole manifold.

3.2. Recap of Orientation on Vector Spaces. Let V be a vector space (over \mathbb{R}). Recall that an *orientation* of V is an equivalence class of (ordered!) bases \mathcal{B} . Namely, two bases are declared to have the same orientation if the change-of-basis matrix has positive determinant; otherwise the orientations are different. Since every vector space has a basis, we know that there is at least one such orientation; and by negation of a single coordinate (or swapping of two coordinates), there are at least two. That there are no more than two is a simple consequence of the fact that there are only two ways to move away from the origin in \mathbb{R} , forward and back.¹⁸ We should also give one particular clarification: colloquially when one speaks of orientations in the context of geometry, one is often speaking about how things are positioned in space. For example, if you're facing north, you're "oriented differently" from facing east. This is a different use of the word, although it is vaguely related to our usage, because both of them deal with what direction things are facing in. Our more "fundamental" orientation has more to do with deciding how to define and distinguish north from south, than to talk about what particular objects happen to be facing in those directions. There is an other colloquial use of the word *orientation* which actually corresponds much more directly to what we are discussing here, namely that of the sexual kind. We will refrain from making jokes about this; mathematicians whose minds are in the gutter will need to seek their entertainment elsewhere.

For arbitrary vector spaces, that is, ones that aren't \mathbb{R}^n , there is *no standard choice of orientation*. This cannot be overemphasized. In \mathbb{R}^n we always have the right-hand vs. left-hand choice, namely, the declaration that $(1, 0, \dots, 0)$, $(0, 1, \dots, 0)$, etc. is right-handed. Given an arbitrary V , we can always map it isomorphically to \mathbb{R}^n , and "pull the orientation of \mathbb{R}^n back," that is, we declare the basis that defines the transformation of V to \mathbb{R}^n positive. But a different isomorphism of V with \mathbb{R}^n may very well have the reverse orientation. Namely for a general vector space we know that there are two orientations, and we can always tell when orientations are the same, or when they are different, but we don't know exactly which one is going to be called positive. Without extra information about V , there is no basis (HA HA) for choosing (except maybe a coin toss?).

It may be instructive to show an nontrivial example of when there *is* a natural choice of orientation. Let V be an n -dimensional *complex* vector space. Choose a complex basis $\mathcal{B} = (v_1, \dots, v_n)$, and then consider V as a real vector space of twice the dimension, by considering the real basis $\mathcal{B} \cup i\mathcal{B} = (v_1, \dots, v_n, iv_1, \dots, iv_n)$. It turns out that *no matter what* complex basis one chooses, *all the corresponding real bases are related by a transformation with positive determinant*. The reason is actually very simple: if \mathcal{B} and \mathcal{C} are two bases, with complex change-of-basis matrix A , then the corresponding real change-of-basis matrix B satisfies $\det B = |\det A|^2 > 0$. In other words, any complex basis determines one and only one orientation as a real vector space of twice the dimension. So every complex vector space has a natural ("canonical") choice of orientation as a real vector space.

Equivalence classes are a little bit cumbersome; it is, after all, trying to avoid equivalence classes that gets us here in the first place. We can't get away from the concept completely, but the way we use it is a lot nicer than just creating a vector bundle by "brute force." Here's another way of thinking about orientation (and it is often the way it is taught). Let V be as before, with dimension n . Consider the 1-dimensional space of n -vectors (oriented parallelepipeds) on V , that is, $\Lambda_n(V)$. The definition of orientation is equivalent to

¹⁸This is not an entirely trivial observation. Modules over various commutative rings can also be endowed with a concept of orientation, which yield many more than just two possible orientations.

choosing, instead of an equivalence class of bases for V , a *half-line* of $\Lambda_n(V) \setminus 0$, given simply by taking the connected component containing the wedge product of all our basis vectors (which cannot be zero, because the vectors are linearly independent). This is the same as our formulation of orientations as equivalence classes of bases, because any two bases for V determining the same equivalence class are related by a transformation T of positive determinant, and we have $Tv_1 \wedge \cdots \wedge Tv_n = (\det T)v_1 \wedge \cdots \wedge v_n$ so that T of this basis determines the same half-line (the two half-lines are closed under multiplication by positive real numbers, and swap places under negation). So the set of two possible orientations of a vector space V can simply be summed up as being the two connected components of the space of all n -vectors on V .¹⁹

3.3. The Pseudoscalar Algebra. Let V be a vector space and o_V be 2-element set of orientations on V , or if you want to get sophisticated about it, $o_V = \pi_0(\Lambda_n(V) \setminus 0)$. Label the two elements of o_V by a and b . You could also call them Green and Purple, Yin and Yang, Male and Female (in fact Weinreich refers to axiality and polarity as the “gender” of a vector!), and speaking of which, Bicycle and Fish, or two otherwise unrelated objects, as long as they don’t coincide with $+$ and $-$, or indeed, not even counterclockwise and clockwise, or left and right, because the identification of those with $-$ and $+$ are too psychologically strong and will confuse the heck out of you as you try to grapple with what is to follow. However, Weinreich suggests \cup and \circ for them, which does in fact suggest counterclockwise or clockwise. This is actually nice, and I would recommend it, but only after you’ve gotten used to the pseudo concept in general and ironed out the wrinkles in your understanding.²⁰

We would like to define an *algebra* of a , b , $+1$, and -1 . We do it by “brute force” as follows: define, formally, $a^2 = b^2 = 1$, $ab = ba = -1$, and 1 to be the identity. The term “formally” just means that the actual contents of a and b do not actually play a role in defining this multiplication; they are rather just tags. However, we shall see it is useful elaborate on what is really happening, and we’ll point things out as they develop. The four units $\{1, -1, a, b\}$ form a group under this multiplication, which is called the Klein 4-group. We should notice right away that $o_V = \{a, b\}$ as a subset is *not* a subgroup. Although we can’t tell which of a and b is negative or positive, we do know that, for $x, y \in o(V)$, if they are “like,” then $xy = +1$ (namely they multiply to a genuine $+$), and if they are “unlike,” then $xy = -1$, the genuine -1 , and applying the genuine -1 to x changes it to the other object. We have revisited the law of signs and given it a twist! But one may wonder why we insisted that the product of these funky sign-like objects will yield the true signed objects $+1$ and -1 . This is actually somewhat of a strange restatement of the fact that a and b cannot just be identified with $+1$ and -1 , and we can see this as follows: if we had made $ab = a$ and $a^2 = b^2 = b$ instead, we have essentially re-created true signs: this would say that b must be the positive sign (because it represents the product of like things), and thus a must correspond to the negative sign. There is an arbitrary choice involved, because what could have prevented us from defining $ab = b$ instead? The point is that products of sign-like things are always “canonically defined,” even if their individual factors are not. This suggests that the right thing to do is to assign the true signs to the products (the other thing that “suggests” this is that we want multiplication by the genuine -1 to always mean

¹⁹In REALLY fancy-schmancy notation, the set of orientations of V is $\pi_0(\Lambda_n(V) \setminus 0)$. π_0 denotes the set of path components of a topological space; it is the logical start point of the homotopy groups $\pi_n(X)$. There is no group structure on π_0 in general, but in the case that it is 2 points, like $\pi_0(\Lambda^n(V) \setminus 0)$, there are exactly two choices of group structure--and that, in fact, is enough information to specify the orientation of V .

²⁰I was considering calling them Red and Black, but any financially savvy person will automatically associate Red with being in debt, i.e. negative net worth, and black with positive!

reversal, so that we have $a = (-1)b = -b$, but this is now, by multiplying both sides by b , the same as the rule for unlike signs). Now, one can't just go back and compare, and then claim that one has to be positive. One might get worried that this means we still haven't gotten away from true signs, but actually this turns out to be exactly what we want. The funky signs are built to interact with the existing ones, not run away from them, and we'll see there is nothing to worry about.

Now we consider vector space $\psi^1(V) = \psi V = \mathbb{R}a$, the span of $\{a\}$, called the **pure pseudo-** or **ψ -** or **axial scalars**, and the direct sum $\psi^*(V) = \psi^*V = \mathbb{R} \oplus \psi^1(V)$ (sometimes we shall write $\psi^0V = \mathbb{R}$), which we shall call the **mixed ψ -scalars** (since $a = -b$, $\mathbb{R}b = \psi^1(V)$ also, so we haven't actually made a dreaded arbitrary choice). The fact that we want ψV to be a vector space already means we have to make true signs interact with a and b . Now, the direct sum ψ^*V is a true associative algebra, that is, closed under multiplication and addition, with the multiplication distributing over the addition (multiplication for the \mathbb{R} part is just the ordinary multiplication). As a vector space, it 2-dimensional, so in this sense, ψ^*V is a lot like \mathbb{C} , where a is analogous to i and b analogous to $-i$, but satisfying different relations. However, unlike the case in \mathbb{C} , we shall mainly consider only pure ψ -scalars (and we shall drop the word "pure") or purely real numbers (called **polar scalars** for emphasis) and not mixed ones.²¹ Pure scalars of either kind are said to have **gender** (and we say the quality of axiality or polarity *is* the gender of such a quantity)²². This is why, if we omit the 1 in $\psi^1(V)$, we still mean $\psi^1(V)$ rather than $\psi^*(V)$ (the $*$ represents "wildcard" here, a notation frequently employed by topologists and, well, computer scientists---don't confuse it with the pullback by a map!). This is a lot like how we defined the algebra $\Lambda^*V = \Lambda^0V \oplus \cdots \oplus \Lambda^nV$ of covectors (the Grassmannian algebra) in which we technically are allowed to work with and add forms of unlike degree, although we never actually do so.²³ $\psi^*(V)$ is called the **pseudo-** or **ψ -scalar algebra**. As the word "scalar" implies, these guys, pseudo or no, have rank 0, and thus are *not* to be confused with the situation for $\Lambda^p(V)$, where the degree p is in fact its tensor rank.

However, just like $\Lambda^*(V)$, we do have a product on $\psi^*(V)$, which we shall just denote by a dot or juxtaposition, because it is commutative, given distributing the above relations over addition. Namely, given real numbers c_1 and c_2 , we have c_1c_2 be the usual product, $c_1a \cdot c_2a = c_1c_2$, $c_1a \cdot c_2b = -c_1c_2$, etc.²⁴ If you like, you can think of $\psi^1(V)$ as " ψ -scalars of degree 1" and real numbers as " ψ -scalars of degree 0," and note that the degree of a product obeys the $\mathbb{Z}/2\mathbb{Z}$ addition law instead of, as in forms, plain old integer addition. Algebraists even have a name for this kind of algebra: a $(\mathbb{Z}/2\mathbb{Z})$ -graded algebra.²⁵ These things come up in the theory of Clifford algebras, with which many physicists may be familiar. Anyway, the upshot is axial and polar scalars as defined above, with the multiplication also defined above, obey Weinreich's rules exactly.

The whole algebra $\psi^*(V)$ is, once again, somewhat like the complex numbers, namely a 2-dimensional vector space with a multiplication, with the basis $\{1, a\}$ (or $\{1, b\}$; recall that $b = -a$ so they aren't linearly independent!). But that's where the similarity ends, and we definitely do not encourage a identifying the spaces. For one thing, given a mixed ψ -scalars

²¹But they are all 100% pseudo!

²²Obviously, mixed ψ -scalars are genderless, or perhaps hermaphroditic. We'll try to keep the jokes to a minimum.

²³Actually, [BBF] does in fact mention offhand some physical theory that makes genuinely good use of forms consisting of sums of terms of unlike degree, but I haven't checked this out yet.

²⁴Had our units $\{1, -1, a, b\}$ formed the other group of order 4, namely the cyclic group, we would instead have the complex numbers \mathbb{C} .

²⁵Or, more facetiously, a Pass/No Pass algebra.

$c+da$ (for some real numbers c, d), we have $(c+da)(c+db) = c^2 + cdb + cda + d^2 ab = c^2 - d^2$, so that if $c = d \neq 0$, we have found a zero divisor! (contrast that with $a = i$ and $b = -i$, where you would get $c^2 + d^2$ which is never zero unless both are!). So as an algebra, $\psi^*(V)$ is not very well-behaved, although as we said, we won't have much occasion to worry about this, since we will only concern ourselves with pure (gendered) scalars. This is akin to only working with real numbers or purely imaginary numbers, but not with fully complex numbers. Note that $1, -1, a$, and b all have unique inverses---themselves (they form a group, remember?), and moreover, all nonzero elements of the individual factors $\psi^1(V)$ and of course \mathbb{R} have inverses. We can define **absolute values** on ψ^*V , which is defined exactly as if it was \mathbb{C} , namely, $|x + ay| = \sqrt{x^2 + y^2}$. For pure scalars this simply is the usual absolute value in \mathbb{R} , and defining $|a| = |b| = 1$ in ψV .

Anyway enough fun with the pseudoscalar algebra, let's actually do something about it.

3.4. Pseudovectors and Their Tensor Spaces. Actually, we've done most of the hard work already. Defining a pseudovector is remarkably easy once we have the concept of pseudoscalar. $\psi^1(V) = \psi V$, just to recap, is the set of pseudoscalars, and is defined formally to be the span of the two units a and b which comprise the set of orientations on V (subject, of course, to the relation $a = -b$). Let V be a vector space as before, and write $\Psi^1(V) = \Psi(V) = V \otimes \psi^1(V)$ and $\Psi^*(V) = V \otimes \psi^*(V)$ (recall: $\psi^*(V)$ contains scalar, rank-0 quantities, not vectors!). (We'll start losing a lot of extra parentheses very quickly, like in the tensor and alternating bundle notations.) We can consider Ψ^*V to be the set of all vectors which can be multiplied by "scalars" in ψ^*V , namely, Ψ^*V is actually a *module* over ψ^*V . This will be an important observation a little bit ahead. Elements of $\Psi^1V \subseteq \Psi^*V$ are called **ψ -vectors** or **axial vectors**, while elements in V will be called just vectors, or for emphasis, **polar vectors**. Again, pure elements of Ψ^*V are said to have **gender**.

The visualization of an element of ΨV , again, given by Weinreich, is that of an arrow missing its head, but instead with a little "curlicue" surrounding the shaft. The general direction of an axial vector, namely how the shaft itself is positioned in space, is in fact the same as it would be for a polar vector; axiality is only deciding the question of what to do with the head. The answer, of course, is to put the head on the curlicue. The sense of the vector is then indicated by which way the curlicue is going around. It is essentially the way a rotating axis is notated, hence the term *axial*. For example, if we draw a rotating planet; we usually indicate its axis of rotation as a line going through the center, and its actual direction of rotation as a curlicue surrounding the axis. This picture also makes it very plainly obvious that the "vector" which is called the angular velocity in most physics textbooks is actually an axial vector, with its curlicue. The fact is, everyone gets too excited about ordinary (polar) vectors that they have to adopt the *right-hand rule* that declares that one must curl the fingers of their right hand around the vector to match the direction of the curlicue, and the direction in which the thumb points is assigned as the polar direction of the vector. This is all well and good, until one tries to reflect the axial vector in a mirror. Let's see what happens with an example.

For concreteness, let's suppose that we're in \mathbb{R}^3 and we have an axial vector parallel to the z-axis with its tail on the xy-plane, for some $y > 0$, with its curlicue running around counterclockwise. Suppose we also have a polar vector, also parallel in the same manner, but pointing up, and also somewhere on the xy-plane where $y > 0$. Put a mirror in the xz-plane, or equivalently, reflect their y-positions. The mirror image of the polar vector of course still points up, as it should. If I'm looking in the mirror and I point my pencil toward the ceiling, both me and my mirror image have pencils pointing toward the ceiling. On the other hand, the mirror image of the axial vector has its curlicue now going the

opposite way, clockwise. Again, that is completely reasonable; the reflection of an upright, spinning top, in a vertical mirror, will look like it's spinning the other way. *But*, if we insist on using the right-hand rule to force polarity on our axial vector, it will have to point up. But the corresponding polar direction of its mirror image will be pointing *down*. So under the reflection, the axial vector behaves as it should, but the “polarized” version of that axial vector does not behave, under reflections, like truly polar vectors should.

If we instead do our reflections on the xy -plane, namely the plane that they're sitting on, we find that the reverse things happen, but the result is the same: polarizing an axial vector creates problems. An up-pointing vector, when reflected in the xy -plane, will point down. Anyone who's ever seen reflections of scenery in a still lake will agree with me on this one. On the other hand, the reflection of the axial vector will have its curlicue also pointing the same way. Spin a top on a hand mirror laid flat on your desk, and you will see that its mirror image rotates in the same direction. But that means their “polarized” versions will both be pointing up, contradicting, again, what would happen with a “true” vector.

Defining the higher tensor bundles is very easy to write down---it is exactly the same process as densitization, tacking an additional factor in the tensor products---but we do have to ferret out some subtleties after all (mostly in the notation). In this process, however, we also solve the problem of defining objects that can accept both axial and polar vectors, changing the result from axial to polar or the reverse, accordingly.

Here's the simple, straight definition of the higher-rank tensors: we define the space of all rank $\binom{k}{\ell}$ - ψ -**tensors** to be

$$\Psi T_{\ell}^k V = T_{\ell}^k V \otimes \psi V = \underbrace{V \otimes \cdots \otimes V}_{\ell \text{ times}} \otimes \underbrace{V^* \otimes \cdots \otimes V^*}_{k \text{ times}} \otimes \psi V.$$

In words, it is the space of all multilinear maps on all of V , with values in ψV . Similarly, the higher ψ -**covectors** $\Lambda^p V \otimes \psi V$ ---all multilinear, alternating, ψV -valued maps (for the special case $p = 1$ this is just $\Psi(V^*)$). We also define higher **mixed ψ -tensors** and **mixed ψ -covectors** to be the above tensor product with $\psi^* V$ rather than just ψV . As things stand, however, these are all maps only on V (and V^*) and *not* ΨV and $\Psi(V^*)$, and hence we *cannot* (yet) insert a ψ -vector into any of the slots of an element of $T_{\ell}^k V$; only the *values* of T lie in $\psi^* V$! We will rectify this situation in short order. It is important, also, to note, for the moment, that the factor ψV appears only *once* in the tensor products for the tensor bundles of higher rank. We should recall the fact that the notions of duality, multilinearity, and tensor product all actually are very dependent on the (algebraic) field or ring we are considering (namely operations on vector spaces also are dependent upon the field over which they are defined). For example the tensor product, for example, over \mathbb{C} , is not the same as the one over \mathbb{R} , if the two vector spaces support those two concepts. We can similarly “pseudo-ize” any vector space E by taking the tensor product with ψV .

However, we *can* actually take tensor products over $\psi^* V$ (this of course technically requires us to use the word “module” when referring to the objects we can do this on, because ψV is not an (algebraic) field, only a ring), and this is the key to allowing ψ -vectors be inserted in to the tensor's slots. Let's motivate this with a simple example, because it is enough to understand these definitions on tensors of rank 2 or less: bilinear maps on $\Psi^* V$. The generic covariant 2- ψ -tensor T is an element of $V^* \otimes V^* \otimes \psi V$, namely, it is a map $T : V \times V \rightarrow \psi V$, taking two vectors and yielding a (pure) ψ -scalar (we should note here that ψV should be considered its own dual---they behave like scalars, remember?). But T can be considered to have values in $\psi^* V$, and we can extend T to a map $T^e : \Psi^* V \otimes \Psi^* V \rightarrow \psi^* V$, by ψV -multilinearity, namely, $T^e(cv, w) = cT(v, w) = T^e(v, cw)$ for all $c \in \psi V$, and $v, w \in V$. For

any ψ^*V -valued bilinear maps on V , which includes both \mathbb{R} -valued and ψV -valued bilinear maps, this extension exists, and is easily seen to be uniquely determined. We've made T able to accept ψ -vectors (actually not just pure, but mixed ones) as well as yield them.

To summarize: the mapping $T \mapsto T^e$ gives a (ψ^*V -module) isomorphism between $V^* \otimes V^* \otimes \psi^*V$ with $\Psi^*(V^*) \otimes_{\psi^*V} \Psi^*(V^*)$ where \otimes_{ψ^*V} means the “module” tensor product over the algebra ψ^*V . Basically \otimes_{ψ^*V} means that we declare things to be multilinear over ψ^*V rather than just \mathbb{R} . Similarly we can use $\otimes_{\mathbb{C}}$ to form spaces of complex-multilinear maps. Anyway, we can generalize the above to all tensors, that is, we can extend any $T \in \Psi^*T_\ell^k V$ to be defined on Ψ^*V and Ψ^*V^* in each of the slots by simply enforcing ψ^*V -multilinearity²⁶.

What is the upshot of all this? Getting back to the ψV -valued bilinear case (pure covariant 2- ψ -tensor case), this means that for all $v, w \in V$, and $c \in \psi V$, $T^e(v, w) = T(v, w) \in \psi V$, $T^e(cv, w) = T^e(v, cw) = cT(v, w) \in \mathbb{R}$, and $T^e(cv, cw) = T(v, w) \in \psi V$, that is, T^e extends the tensor's ability to eat both vectors and ψ -vectors. T^e actually has a very strange feature that actually (almost) obliterates the first line of the old rhyme: let $c = a$, which, recall, is one of the units in our ring (one of the orientations of V). Then for $v, w \in V$ (i.e. v, w polar), we have av and aw are axial. No surprises so far. Remember, the very concept of multiplication is a bilinear thing (“bilinear” is a fancy name for “distributive law” --- to really convince you, ask yourself how you would calculate $T(u + v, w + z)$ for any bilinear tensor T ? It's nothing other than FOIL!). So $T^e(v, w)$ is in some sense a “product” of v and w , and, if T is ψ -valued, we have $T^e(v, w) \in \psi V$, or *polar times polar is axial*; $T^e(av, w) = T^e(v, aw) = aT(v, w) \in \mathbb{R}$, or *axial times polar is axial* and *polar times axial is axial*, and $T^e(av, aw) = T(v, w) \in \psi V$, that is, *axial times axial is axial*. What we've found is a seeming contradiction to Weinreich's observations! Actually no, it really isn't. The error is twofold: first, it is only a contradiction if we claim that $a = -1$, so that it becomes the claim “minus times minus is minus.” But we'll see, however, we can make $a = -1$. . . *provided we have chosen an orientation*, i.e. we have said that b is the “positive” orientation. So there's still a contradiction. But this just brings us to the second error: you're already familiar with objects that can be multiplied by -1 , and yet are not themselves negative or positive (or can be even said to have any kind of “sign”): vectors themselves. There's no sense calling the vector $(1, -2, 5)$ is “positive” but $(-1, 2, -5)$ is negative (although the right-hand rule could theoretically make that distinction for us: we'd read the sign off from the last coordinate only. But let's not make things harder than they need to be, ok?).

When referring to the extended tensors T^e , we shall just drop the e from now on and understand that if one has an argument of the wrong “type,” we mean that we have to use the extended operator.²⁷

Now that we have described the algebra of what happens---after the dust settles one can see such a thing really is simple (yes, I know that is hard to believe right now, but you'll see eventually). . . how about the *geometry* of the situation? Let's look at bivectors in \mathbb{R}^3 . Given two (polar) vectors v and w , their wedge product $v \wedge w$ is an oriented parallelogram, a polar 2-vector. The “direction” of a polar 2-vector is actually given by a curlicue! Namely, v denotes one side of the parallelogram, and w denotes the other; the curlicue is chosen so that the increasing direction takes v into w . With the right-hand rule, of course, this amounts to sticking an arrowhead directly onto the parallelogram itself (the construction of what Weinreich calls a “thumbtack” --- but actually our explanation of the gender here is turns out to be opposite of what he's talking about. We will get to that when we describe Hodge

²⁶I apologize for the excess of stars. The fact is, we *must* use ψ^*V in the definition of tensor product for modules, because in module theory, tensor products have to be taken over rings.

²⁷I.e. we “overload” the function T , in the language of computer programming.

duality). The wedge product of two axial vectors also yields a polar 2-vector, and it also rotates from the first to the second---provided that you rearrange them so that the curlicues both look like they're moving in the same direction when you look from the common point. On the other hand, an axial 2-vector results from wedging two forms of unlike gender. The way to do it is arrange it so either end of the axial vector meets the tail of the polar one, and look at the parallelogram they make. The direction in which the curlicue *pierces* the parallelogram determines the sense.

Weinreich identifies such parallelograms with covariant capacities (we saw an exact analogue of this earlier for forms, by looking at the transformation law for the three components of Λ^2 in the natural form-basis, not the general 9 components of the tensor basis).

3.5. The Orientation Line Bundle. Ok, now we are ready for the really good stuff: doing this on a manifold, orientable or not! We should note, once again, that ψV is well-defined regardless of whether or not we have specifically chosen an orientation of V . This is because the set of orientations on V , o_V , always has two elements, the two connected components of $\Lambda^n(V)$, which we gave the arbitrary names a and b , and they are declared to be linearly dependent---negatives of one another. ψV is then just the span. Also, ψV obviously depends on V ; after all, it's the orientations of V that we're talking about.

So, on a manifold, we always have the 1-dimensional vector space (i.e. line) $\psi(T_p M)$ for every tangent space $T_p M$. We can bundle everything up, to define $\psi M = \psi^1 M = \coprod_{p \in M} \psi^1(T_p M)$, and π be the usual projection map that sends everything in the fiber, $\psi_p M$, to p . This is why we wanted ψV to depend on the vector space: we can assign it to each tangent space separately. We call this the **orientation line bundle**. Anyway, we have to check, of course, that this really defines a smooth vector bundle, and that we're not just spouting nonsense (well, we are spouting nonsense, but not *complete* nonsense). In order to do that, we have to specify the local trivializations. Given $p \in M$ and (U, x) a coordinate chart at p , we have that $\frac{\partial}{\partial x^i}$ gives a coordinate *frame*, i.e., a basis for $T_p M$ throughout U . But every coordinate frame gives us a parallelepiped $P = \frac{\partial}{\partial x^1} \wedge \cdots \wedge \frac{\partial}{\partial x^n} \in (\Lambda_n)_p M$. Since the $\frac{\partial}{\partial x^i}$ are a basis, $P \neq 0$ must lie in one of the half-lines of $(\Lambda_n)_p M \setminus 0$ (that is, it is an element of $o_{T_p M}$). Write $o(x)_p$ for it, and $o(x)$ for the function of p .²⁸ Of course, we don't know whether $o(x)_p$ is a or b ; all we know is that it is one of them, and we have no idea which one it is. But we don't really care, for the moment; all we care is that we can multiply it by elements in $\psi_p M$. With this in mind, we define a local trivialization $\Phi : \pi^{-1}(U) \rightarrow U \times \mathbb{R}$, for $(p, s) \in \psi_p U = \pi^{-1}(U)$ as follows:

$$\Phi(p, s) = (p, o(x)_p \cdot s).$$

Note that $o(x)_p \cdot s \in \mathbb{R}$, because both $o(x)_p$ and $s \in \psi_p M$. But what do elements of $\psi_p M$ multiply to? Elements of \mathbb{R} of course. The inverse of Φ is then given by $\Phi(p, r) = (p, o(x)_p \cdot r)$ where now $r \in \mathbb{R}$ so that $o(x)_p \cdot r \in \psi_p M$. It all works out because $o(x)_p$ is its own multiplicative inverse.

Composing with the chart map $p \mapsto x(p) = (x^1(p), \dots, x^n(p))$, we get a coordinate chart on ψM given by $(p, s) \mapsto (x^1, \dots, x^n, o(x)_p \cdot s)$. It remains to show that overlapping

²⁸This is actually Frankel's notation to denote an orientation of a frame, but his definition is confusing because, he specifically says it is function into $\{\pm 1\}$ and not $\{a, b\}$. On the other hand, *the way he uses it* is actually our way of doing things. For example, he even actually casually refers to $o(x, y, z)$ as an orientation of \mathbb{R}^3 . What apparently happened is that in his definition, he makes the major no-no of conflating a and b with ± 1 when we don't have a preferred orientation chosen.

coordinate charts are smoothly compatible. Let (V, \bar{x}) be another coordinate system. Then as before, we have

$$\frac{\partial}{\partial \bar{x}^1} \wedge \cdots \wedge \frac{\partial}{\partial \bar{x}^n} = \det \left(\frac{\partial x}{\partial \bar{x}} \right) \frac{\partial}{\partial x^1} \wedge \cdots \wedge \frac{\partial}{\partial x^n}.$$

This means that

$$o(\bar{x}) = \operatorname{sgn} \left(\det \left(\frac{\partial x}{\partial \bar{x}} \right) \right) o(x)$$

because the half-lines in $o_{T_p M}$ are closed under multiplication by positive numbers, and swap places under multiplication by negatives. All in all, the coordinate transition map is then

$$(x^1, \dots, x^p, r) \mapsto (\bar{x}^1, \dots, \bar{x}^p, \operatorname{sgn} \left(\det \left(\frac{\partial x}{\partial \bar{x}} \right) \right) r).$$

But because $\det \left(\frac{\partial x}{\partial \bar{x}} \right)$ never vanishes over all of $U \cap V$, its sign is therefore a smooth function on $U \cap V$ (in fact, locally constant--it can still jump from -1 to 1 , but only between different connected components of $U \cap V$, which does *not* affect smoothness!). So smoothly compatible charts on M yield smoothly compatible charts on ψM and thus ψM is a smooth vector bundle. Eliminating the step of actually converting coordinates of the point itself, this yields ψM to be a vector bundle with the transition maps

$$(p, r) \mapsto (p, g_{VU}(p)r) = \left(p, \operatorname{sgn} \left(\det \left(\frac{\partial x}{\partial \bar{x}} \right) \right) r \right)$$

or $g_{VU}(p) = \operatorname{sgn} \left(\det \left(\frac{\partial \bar{x}}{\partial x} \right) \right)$ which is, as we remarked, a smooth function into $\mathbb{R} \setminus 0 = GL_1(\mathbb{R})$. Because $x \operatorname{sgn}(x) = |x|$, and $|x| \operatorname{sgn}(x) = x$, these factors are exactly thing needed to transform determinants of Jacobians into their absolute values and vice versa. Since $\operatorname{sgn} \left(\det \left(\frac{\partial x}{\partial \bar{x}} \right) \right) = \operatorname{sgn} \left(\det \left(\frac{\partial \bar{x}}{\partial x} \right) \right)$, things work equally well for contravariant and covariant transformation laws, or densitization and capacitization.

For a comparison, the trivial bundle $M \times \mathbb{R}$ could be constructed in the same way, except the transition functions would all identically be 1 , because the vector element is already in \mathbb{R} to begin with.

So far, we have said absolutely nothing about whether or not M is orientable. The bundle ψM was well-defined precisely because we took both possible orientations of each tangent space into consideration (the span of $\{a, b\}$ with $b = -a$). But M being orientable probably has *some* effect on ψM . What is it? It turns out to be this:

3.1. Theorem. M is orientable if and only if ψM is trivial.

We'll give the proof in the next section, as we will recap the idea of orientations of manifolds in general there. We can also bundle $\psi^* T_p M$ at every point to form $\psi^* M$; this is just the direct sum bundle consisting of the trivial bundle in one factor and ψM in the other. It is the bundle of mixed ψ -scalars. Note that even though all the fibers are distinct, there is a global metric on $\psi^* M$ given by the ordinary \mathbb{R}^2 dot product on the components. It is well-defined and independent of coordinates, because $o(y)$ differs from $o(x)$ only by a factor of ± 1 regardless of what coordinates we choose. The norm defined by this metric is simply the square of the absolute value we defined on ψV earlier, and in this norm, we always have $|o(y)| = |o(x)|$ for any coordinate chart. This norm will be very useful in defining the isomorphisms in the next section. For now it should be noted that if f is a smooth nonvanishing section of $\psi^* M$, then $|f|$ is a smooth function into \mathbb{R}^+ .

Now everything follows as it should: **ψ -tensor bundles** are then defined by tensoring $T_k^\ell M$ with ψM pointwise. In fact, any vector bundle E over M may be pseudo-ified by

this process. Or, to use what is rapidly becoming a favorite phrase of mine, we **twist the vector bundles by the orientation line bundle**. If we additionally densitize by also twisting by the determinant line bundle $\Lambda^p M$ we derive the **densitized pseudo twisted fancy-schmancy bundles**. Actually to make it sound less frightening (even though it also sounds way cool!) we should just define one very important bundle (it consists of the objects we can integrate on non-orientable manifolds!).

3.2. Definition. $\Lambda^n M \otimes \psi M$ is called the **honest density bundle** of M , or the bundle of **ψ -scalar densities** (this is as opposed to just densities in general which only include the $\Lambda^n M$, and dishonest density bundles, which is sometimes defined in place of the honest one as the space of absolute values of n -forms on M). For $0 \leq p \leq n$ we call $\Lambda^p M \otimes \psi M$ the **ψ -forms**. ψ -scalars, of course, are just the 0 - ψ -forms.

3.6. Differential Pseudoforms and Integration. Ok, we now have pseudotensors, and moreover we can mix and match our ways of twisting them. Isn't that nice? We can also have sections of these bundles, namely, entire fields of these arbitrarily exotic objects. But let's focus into our most important variety: the differential ψ -forms $\Omega^p(M) \otimes \mathcal{P}M = \Omega^p(M, \psi M)$, where $\mathcal{P}(M)$ are sections of ψM . We call sections of ψM **pseudoscalar fields**, or just ψ -functions. For you electromagnetism buffs, such a section might give a magnetic charge distribution (if you believe in magnetic monopoles, like I do---they make Maxwell's Equations look even more symmetric!).

One of the classical reasons for developing a theory of integration was to compute volumes, or quantities dependent on volumes---so-called *weighted* volumes where some portion of space somehow "counts more" than other portions. In fact it may be that the phrase "counts more" originates from integration, where instead of plain counting to total up the number of things, somehow, some things might actually have greater value, and, well, count more. So it is of course natural to return to this notion with the souped up theory of integration of differential forms. A pseudo- n -form (i.e. scalar density) on a manifold is a suitable object for calculating the volume of a manifold, because, since it is a density, and integrals total up a density over an entire manifold.

Let's reflect on what goes wrong when trying to integrate true differential forms. Since what we end up really doing when integrating forms is doing the integral in separate patches and adding the results up, the issue is how to do such a sum in a consistent way. We saw that if we rely on just differential forms to do things, different coordinate patches can give different results differing by a sign, and chaos reigns when adding the results on each patch. So long as we are in a single patch, the problem is only a sign, but since we add everything up from different patches, each with a sign ambiguity, the answer could be *anything*. For orientable manifolds, this problem is solved, because now we have something to compare to in each patch, and we can make the correct choice of sign every time. But for nonorientable manifolds, such a process doesn't exist. We need some way of encoding orientation information into the patch itself. This is the job of the little $o(x)$ in the coordinate n - ψ -form $o(x)dx^1 \wedge \cdots \wedge dx^n$, and why pseudoforms are the true integrands. In analysis one believes that functions are the objects of choice for integration, and in most ordinary treatments of differential geometry, one gets ingrained with the idea that differential forms are it. But it really is differential pseudo-forms which are the ideal integrands.

We must give one stern warning, however, because one can get so excited about ψ -forms that they are tempted to think that they are panacea. Not everything works on ψ -forms. One major sacrifice is the ability to pull such forms back at will, by any map. ψ -forms can

only be pulled back via diffeomorphisms.²⁹ Of course, this isn't so bad in general, since the same is true for vector fields and other contravariant guys. It therefore follows that we still will not be able to integrate ψ -forms with impunity over any nonorientable submanifold we can come up with. With forms and an orientable submanifold, recall, the integral of a true form over requires pulling the form back by the inclusion map. There's a way of getting around some of these issues, however (as there must be, since, once again, an embedded Möbius strip has area, and its Riemannian density must somehow be related to the ambient one on \mathbb{R}^3).

First off, we can declare the wedge product to be multilinear over ψ^*M so that $\psi^*M \otimes \Lambda^*M$ is an algebra which combines the two operations. It essentially doubles the number of grades---forms of each degree, and then within each degree, a real and ψ version (or in more consistent terminology, an axial grade and polar grade). Multiplying by axial scalars exchanges grades within a degree, by polar scalars preserves the degree, and by other forms increases the degree and adjusts the gender accordingly. Sometimes ψ - p -forms, i.e. sections of $\Omega^p(M, \psi M)$ are also called **forms of the odd kind**, as opposed to the usual non- ψ -forms, **forms of the even kind**, which is terminology first given by Georges de Rham. Also note that this extended wedge product always obeys Weinreich's observations, that is, the wedge product of forms of like gender is polar, and unlike gender is axial. De Rham's "parity" classification, namely odd and even, is based on the same observation except now by analogizing with the addition laws for odd and even numbers: like parity sums to even, unlike parity sums to odd.

Now given an **honest density** φ , that is, $\varphi \in \Omega^n(M, \psi M)$, we have that φ looks like $sdx^1 \wedge \cdots \wedge dx^n = o(x)fdx^1 \wedge \cdots \wedge dx^n$ where s is a ψ -scalar field and f is a plain ol' scalar field equal to $o(x) \cdot s$, in a coordinate chart. If in another coordinate chart, $\varphi = \bar{s}d\bar{x}^1 \wedge \cdots \wedge d\bar{x}^n$ then $\bar{s} = s \det(\partial x / \partial \bar{x})$ and

$$\bar{f} = o(\bar{x})\bar{s} = o(\bar{x})s \det\left(\frac{\partial x}{\partial \bar{x}}\right) = o(x)s \left| \det\left(\frac{\partial x}{\partial \bar{x}}\right) \right| = f \left| \det\left(\frac{\partial x}{\partial \bar{x}}\right) \right|.$$

In other words there are two ways of conceiving of $\Omega^n(M, \psi M)$ in terms of coordinates: either the basis n -form times a ψ -function s , or f times the basis ψ -form $o(x)dx^1 \wedge \cdots \wedge dx^n$. Sometimes people will write $|dx^1 \wedge \cdots \wedge dx^n| = o(x)dx^1 \wedge \cdots \wedge dx^n$, but we will not do this, because other people write that to mean, literally, the absolute value of $dx^1 \wedge \cdots \wedge dx^n$, which would take n vectors and always yield a positive quantity---something we're trying to avoid having to talk about. Also, some people will write $dx^1 \dots dx^n = o(x)dx^1 \wedge \cdots \wedge dx^n$, i.e. without the wedges, in analogy to the ordinary integrand of multivariable calculus. Again we'll refrain from that, and reserve $dx^1 \dots dx^n$ for times when we are *really* doing the plain ol' integral in \mathbb{R}^n .

We wish to define $\int_M \varphi$. First, suppose φ is supported (i.e. nonzero) in just one coordinate chart (U, τ) . We define:

$$\int_M \varphi = \int_M s dx^1 \wedge \cdots \wedge dx^n = \int_M o(x)fdx^1 \wedge \cdots \wedge dx^n := \int_{\mathbb{R}^n} f \circ \tau^{-1} dx^1 \dots dx^n.$$

where, again, $dx^1 \dots dx^n$ is just the ordinary integrand of Euclidean space (it could also denote the standard measure, if we take our integrals in the sense of Lebesgue). This looks

²⁹The way to do it is as follows: If $F : M \rightarrow N$ is a diffeomorphism, it induces a bundle isomorphism $F^* : \Lambda_n(N) \rightarrow \Lambda_n(M)$ (contravariant tensors always can be pulled back by diffeomorphisms). As F^* is linear on the fibers, it is a continuous map on the fibers, so it maps connected components of $(\Lambda_n)_p(N) \setminus 0$ into those of $(\Lambda_n)_p M$ for each p . This tells where the $a, b \in \psi N$ map to: the set-map images of a and b under F^* as a map of n -parallelepipeds.

like clever BS, which it is in some sense. The whole notation surrounding differential forms was specifically tailored to look like an integrand, because that's what we usually ultimately care about when it comes to dealing with differential forms.

We do have to make sure that the integral is independent of choice of charts. Let τ, ρ be different charts containing the support of φ . Then

$$\begin{aligned} \int_M \rho dx^1 \wedge \cdots \wedge dx^n &= \int_M o(x) f dx^1 \wedge \cdots \wedge dx^n = \int_{\mathbb{R}^n} (f \circ \tau^{-1}) dx^1 \cdots dx^n \\ &= \int_{\mathbb{R}^n} (f \circ \rho^{-1}) \left| \det \left(\frac{\partial x}{\partial \bar{x}} \right) \right| d\bar{x}^1 \cdots d\bar{x}^n = \int_M o(\bar{x}) \bar{f} \left| \det \left(\frac{\partial x}{\partial \bar{x}} \right) \right| d\bar{x}^1 \wedge \cdots \wedge d\bar{x}^n \\ &= \int_M o(\bar{x}) \bar{f} d\bar{x}^1 \wedge \cdots \wedge d\bar{x}^n = \int_M \bar{\rho} d\bar{x}^1 \wedge \cdots \wedge d\bar{x}^n \end{aligned}$$

Because now f has the right transformation behavior and matches up with the Change of Variables formula in plain old multivariable calculus (which was applied in the equality involving the wedge-less integrals), there is no need to suppose that the charts are “positively” oriented (this is the point where orientation came in when defining the integral of an ordinary form on an oriented manifold). Now we can carry it into forms which are not necessarily only supported in a single coordinate chart by using a partition of unity. The preceding just shows our calculations to be coordinate-invariant. In other words, what we have done is *reduce integration on manifolds to integration in \mathbb{R}^n , something we presumably already know how to do*. Also, it is essential for some theoretical purposes to restrict to forms with compact support (for example, in the de Rham cohomology theory which relates the structure of forms to the topology on the manifold). Of course with Lebesgue integration in a chart we can hope to define a more general integral, but if we get ∞ in a chart, there will be some very bad news when having to add things up from different patches. This is why orientation vs. pseudo-ness is even more essential: you may just end up having to do $\infty - \infty$ if you have the wrong charts. Without the pseudo-business, you can't tell whether it's just your charts that are bad (coordinate system failure), or whether your function is genuinely bad.

We should also define the exterior derivative of a ψ -form φ . The definition is very easy: $d\varphi = o(x)d(o(x) \cdot \varphi)$ in any coordinate chart. Since $o(x) \cdot \varphi$ is an ordinary form, we can use ordinary d on it. Then another $o(x)$ changes it back to a ψ -form. This is independent of choice of coordinate charts simply because two factors of $o(x)$ appear (the fact that the transition functions of the bundle ψM are locally constant functions means the extra sign factor will pass right through the d when comparing what happens in different charts). We have the usual rules $d^2 = 0$, $d(fd) = df \wedge dx$ and all the rest (except pullback). A version of Stokes's theorem holds for the integration of ψ -forms, but this will have to wait until we talk about the tricky task of integrating over submanifolds.

Finally, we can, once again, talk about inserting axial vectors into the slots of a differential form by simply extending them to be ψ^*V -valued, defined on tensor products of Ψ^*V and their duals. Also, with ψ -vector fields, both axial and polar vector fields can be applied to axial and polar scalars to give directional derivatives: we define $Xs = ds(X)$ using this extension if necessary. It is the usual directional derivative if both are polar; otherwise we use the funky d that we just defined. Again, combining scalar and vector fields of like gender yields a directional derivative of polar gender, and of unlike gender yields one of axial gender.

Now, given a metric g on M , let's define the honest Riemannian density on M , which *always* exists regardless of the orientability of M . This allows us to integrate true functions

(not ψ -functions) on manifolds. Define in a coordinate chart,

$$|d\mu_g| = o(x) \sqrt{\det g_{ij}} dx^1 \wedge \cdots \wedge dx^n.$$

Note that positive-definiteness of g is required in this definition to avoid getting an imaginary result (i.e. make the term under the square root positive). It is of course possible to define things for general nondegeneracy of the metric, which is why relativity theory has an even more aggravating number of signs to keep track of.

Anyway, let us see that this is well-defined. Recall that $\det g_{ij}$ transforms as follows:

$$\det(\bar{g}_{ij}) = \det\left(\left(\frac{\partial x^k}{\partial \bar{x}^i}\right) g_{kl} \left(\frac{\partial x^\ell}{\partial \bar{x}^j}\right)\right) = \det\left(\frac{\partial x}{\partial \bar{x}}\right)^2 \det(g_{kl})$$

This shows the quantity $\det(g_{ij})$ to be the single component of a scalar density of weight 2, rather than a function.

Taking its square root means we have to use absolute value signs (remember $\sqrt{x^2} = |x|$, not x):

$$\sqrt{\det(\bar{g}_{ij})} = \left| \det\left(\frac{\partial x}{\partial \bar{x}}\right) \right| \sqrt{\det(g_{kl})}.$$

We will use the briefer (and actually older) notation $\sqrt{\det(g_{ij})} = \sqrt{g}$. There is little ambiguity because you can't take the square root of a tensor anyway. We see that this fellow transforms as the component of a ψ -scalar density (of weight 1). This is exactly what we want! Writing out the full coordinate change gives

$$\begin{aligned} |d\mu_g| &= o(\bar{x}) \sqrt{\bar{g}} d\bar{x}^1 \wedge \cdots \wedge d\bar{x}^n = o(\bar{x}) \left| \det\left(\frac{\partial x}{\partial \bar{x}}\right) \right| \sqrt{g} d\bar{x}^1 \wedge \cdots \wedge d\bar{x}^n \\ &= o(x) \sqrt{g} \det\left(\frac{\partial x}{\partial \bar{x}}\right) d\bar{x}^1 \wedge \cdots \wedge d\bar{x}^n = o(x) \sqrt{g} dx^1 \wedge \cdots \wedge dx^n, \end{aligned}$$

so that the coordinate representations are the same in different charts. Once again, we see that the extra $o(x)$ does its usual job of absorbing factors of $\text{sgn} \det(\partial x / \partial \bar{x})$. We can thus integrate functions even on nonorientable Riemannian manifolds, by simply taking $f \mapsto \int_M f |d\mu_g|$. Since the Riemannian ψ -density is defined solely in terms of the metric, and metrics pull back to any submanifold, it is well-defined concept on submanifolds, and thus, we can integrate functions defined on M on submanifolds, by using the density induced by the pullback metric. This finally puts to rest the question of the area of an embedded Möbius strip. We talk more about what happens on submanifolds later.

The Riemannian honest density is clearly never-vanishing, so it actually shows that $\Psi\Lambda^n(M) = \Lambda^n(M) \otimes \psi M$ is a trivial bundle. Note that $\Lambda^n(M)$ is trivial if and only if M is orientable. Notice how $\Lambda^n(M)$ behaves, when attempting to assess its triviality, exactly like ψM . What we just showed was that $\Psi\Lambda^n(M)$ behaves exactly like the trivial product bundle $M \times \mathbb{R}$. Note the role-reversal of the pseudo-ness. To summarize: *with a metric*, ψ -scalars and true n -forms are alike, and true scalars and ψ -forms are alike. The isomorphism is given, *in both cases*, by multiplying by $|d\mu_g|$.

Actually, since every manifold (noncanonically) admits a Riemannian metric, it shows that those bundles are always (noncanonically) isomorphic. In fact, the honest density bundle allows one to construct the *fanciest-schmanciest tensor bundle of them all*: densities of *fractional* weight! An absolute value may be raised to any real power. There is a nice theorem in representation theory that shows that the classification of the various exotic properties of tensors is given by the classification of the irreducible finite-dimensional representations of $GL_n(\mathbb{R})$ --and it says that there aren't any more varieties of tensors than which have

been given (basically our total list of exotic subspecies are contra- and co, symmetric and antisymmetric, axial and polar, and densitized and capacitized).

3.7. Oriented Manifolds: Identifying Axial and Polar Quantities. Ok, now let us speak of oriented manifolds. As noted before, we should view the distinction between axial and polar as a distinction that can be made to disappear if we have an orientation, in exactly the same way that the distinction between contra- and covariance can be made to disappear if we use a Riemannian metric. A Riemannian metric canonically identifies the tangent and cotangent bundles and all of their products, and we shall show an orientation canonically identifies the ψ - and ordinary tangent bundles and all of their products. But as any physicist and mathematician knows by coming this far, just because we *can* get rid of contra- and covariance with a metric, doesn't mean we *should*, and we have seen situations where one is a more natural way of doing things than the other.

First off, let's recall what we mean by a manifold being orientable. As we have seen, individual tangent spaces always have exactly two orientations. Orientability of a whole manifold then corresponds to whether or not one can make a "consistent, smoothly-varying choice" of orientations of T_pM for all $p \in M$. What this amounts to, in most traditional ways of defining it, is whether or not M admits a covering with coordinate patches such that the Jacobian has positive determinant on all the overlaps. Since the sign of this determinant also precisely describes the transition functions for ψM , this implies ψM is trivial (we'll see this in more detail in a moment). In a single coordinate chart (U, x) , a "smooth choice" of orientation on the tangent bundles is simply given by the coordinate frames $\frac{\partial}{\partial x^i}$, i.e., $o(x)$.

In coordinate-free language, this means that an orientation on M can be specified by a *global*, smooth nonvanishing section s of ψM , i.e. a never-vanishing ψ -function. The orientation for a particular tangent space T_pM is then given by $s(p)/|s(p)|$. Because s is never-vanishing, we see that an orientation is also equivalent to saying that ψM admits a global smooth section s of unit norm, i.e. $|s| = 1$ at every point (recall that the norm defined on ψ^*M is coordinate-independent, i.e. it comes with a canonical metric. TM , by contrast, admits many metrics, and from M alone we could use partitions of unity to construct a global metric; but this is not a *canonical* choice!). A section of unit norm simply means that at every point p , we have that $s(p)$ is the orientation of T_pM (this is why it is so convenient to define the generators in pseudoscalar algebra as actually being the orientations themselves). In particular, the coordinate frame orientation $o(x)$ is simply a section of unit norm over the coordinate patch (U, x) . This shows, in particular, local orientations always exist. That's why we have to make sure to emphasize that word *global* in the above.

An orientation of M can *also* be specified by a smooth nonvanishing section ω of $\Lambda^n M$. One simply takes the orientation in each T_pM to be the half-line in $\Lambda_n M \setminus 0$ which contains the dual of ω_p (recall that $\Lambda^n M = (\Lambda_n M)^*$). Let's show that orientability by charts with positive determinants is equivalent to triviality of ψM .

Proof. If M is orientable, we can cover M with charts all of whose overlaps have positive Jacobian determinant, i.e. for which all the $\text{sgn}\left(\det\left(\frac{\partial \tilde{x}}{\partial x}\right)\right)$'s are equal to $+1$. Then the transition functions for ψM are, in fact, exactly the same as those of $M \times \mathbb{R}$ with the same charts, showing the triviality of ψM .

On the other hand, if ψM is trivial, we can find a global trivialization $\Phi : \psi M \rightarrow M \times \mathbb{R}$ (which, of course, need not come from a single coordinate chart!). Let M be covered with coordinate charts. Then $\text{sgn}\left(\det\left(\frac{\partial \tilde{x}}{\partial x}\right)\right)$ can be compared against the corresponding bundle transition function from each given chart to the global trivialization chart. If the sign matches up, leave the chart alone; otherwise replace the chart with a new one that

swaps the first two coordinates. This new collection of charts, then, must have overlaps with positive Jacobian determinant, because each can be made to go through the global trivialization, and by the adjustment, must have positive sign. \square

There is a similar proof showing that the triviality of $\Lambda^n M$ is equivalent to this. But for now what we do is construct a nonvanishing section by multiplying some Riemannian (honest) density (which always exists, but, you guessed it, just not canonically. . .) with the nonvanishing section of ψM .

Suppose M is orientable. Then a specific choice of any one of those things above (nonvanishing sections of ψM or $\Lambda^n M$, or charts all of whose transition functions have positive Jacobian determinant) is called, you guessed it, an **orientation** of M , and M is said to be **oriented**. Let's get the terminology straight, because it is important to realize this: M is said to be *orientable* if it is possible to find one of those guys mentioned above. *However*, even if M is *orientable*, M is not *oriented* until you have actually chosen a specific *orientation*. The orientation is just an extra structure that one must specify, but orientability merely asserts whether it is possible to put that extra structure on it. It is similar to considering manifolds vs. Riemannian manifolds: a Riemannian manifold has an extra metric structure specified, while a general manifold does not. One can say a manifold is "Riemann-metrizable" if it is possible to put a Riemannian metric on it, but it turns out that this is a silly notion, because it is always possible to do it using a partition of unity. But don't laugh it off so easily, one could ask if manifolds are "Lorentz-metrizable," and this in fact is a highly nontrivial question---some manifolds really cannot have such a metric. But just because we *can* choose an orientation or Lorentz metric on M , doesn't mean we *should*. We should orient manifolds only if we have a good reason to do so.

So let's talk about oriented manifolds. Our most efficient method of specifying an orientation will be, from now on, a global nonvanishing section s of ψM , which we may assume to be of unit length (after all, why would we bother calling it an orientation bundle if it wasn't such a great choice?). Such sections always have coordinate representations that are *locally constant*. In particular, if M is connected, then there are only 2 possible orientations, namely s and $-s$. On the other hand, if M falls into components (the set of which we denote $\pi_0(M)$), then given an orientation s of M , we can multiply any restriction of s to any component by -1 without affecting values of s anywhere else (and, without affecting the smoothness of s), and get another orientation of M . The total number of possible orientations for M , *if M is orientable*, is therefore $2^{|\pi_0(M)|}$.

Now let us (finally!) get to the title of this subsection: identification of axial and polar quantities for oriented manifolds. Given an orientation of M , that is, $s : M \rightarrow \psi(M)$ a smooth section of unit norm, and any vector bundle E over M , we define a bundle isomorphism of $E \otimes \psi M$ with just E by multiplying elements by $s(p)$. Since an element of $E_p \otimes \psi_p M$ is a pure ψ -scalar times some $e \in E$, multiplying it by the ψ -scalar $s(p)$ gives a genuine real number times e . Because s is smooth, the bundle isomorphism is in fact smooth. Again, this is reliant on M being orientable, because we cannot otherwise choose (smooth) sections with unit norm. For example, if M is Riemannian with metric g and honest density $|d\mu_g| \in \Omega^n(M, \psi M)$, then we define the Riemannian volume form to be $d\mu_g = s|d\mu_g| \in \Omega(M)$ for some orientation s . It is important to note that if we use a different orientation, s' , we will get a *different* bundle isomorphism of $E \otimes \psi M$ with E . In the case that M is connected, this isomorphism is just the negative of the other. But for disconnected M , it's a lot more complicated than just a sign difference. This is why it is crucial to distinguish between *orientable* and *oriented* manifolds. An orientable manifold M still only admits noncanonical isomorphisms of $E \otimes \psi M$ with E (this is, however, an

improvement over M being nonorientable, in which case there are no (smooth!) bundle isomorphisms at all!).

3.8. Honest Densities and De-densitization. Now that we have pseudotensors, we can describe another thing that a metric (and the honest density defined by it) can do. Recall that any 1-form can be “made” into a vector by “raising indices,” that is, if $\omega = \omega_i dx^i$ in coordinates, then $\tilde{\omega} = g^{ij} \omega_i \frac{\partial}{\partial x^j}$ defines a vector field based on that 1-form. Pictorially this means we take the “normal vector” to the “sheets” defined by the 1-form. The metric is needed for two things: one to define what “normal” means, and also to scale length accordingly (recall that 1-forms deal with units of inverse length, while tangent vectors deal with units of length). Thus the metric is used to “de-covariantize” a covariant quantity. Without fixing a choice of metric we cannot de-covariantize anything.

Similarly, an orientation (unit section s of ψM) may be used to “de-pseudoize” a quantity, because we can then consistently and smoothly multiply such things by s . Without fixing a choice of orientation, we cannot de-pseudoize anything (and for nonorientable manifolds, it can't be done at all).

Through its honest density, the metric also be used to identify ψ -densitized quantities with un-densitized ones, in exactly the same spirit. Pictorially, we recall that the metric is also used to define what unit volume means. So it is no surprise that it can be used to convert densities. Combined with a metric *and* orientation, we can de-densitize any true (polar) quantity. Let's give an example of how it is done, say with a densitized vector (the idea is exactly the same for other kinds of tensors, so we'll save some notational headaches by doing it this way). Namely, given a quantity with components \mathcal{B}_i which transform like

$$\bar{\mathcal{B}}^i = \det\left(\frac{\partial x}{\partial \bar{x}}\right) \left(\frac{\partial \bar{x}^i}{\partial x^j}\right) \mathcal{B}^j$$

(look familiar? it's just the magnetic field example as before, but with a scripty \mathcal{B} ---there is a reason for choosing just this example!). We could suppose, for example, that the \mathcal{B}^i are the components in the natural $(n-1)$ -form basis. Now consider $B^i = \frac{\mathcal{B}^i}{\sqrt{g}}$. Since \sqrt{g} transforms with a factor of $|\det(\partial x/\partial \bar{x})|$, its inverse just flips the determinant over, and thus the B^i transform as

$$\bar{B}^i = \text{sgn}\left(\det\left(\frac{\partial x}{\partial \bar{x}}\right)\right) \left(\frac{\partial \bar{x}^i}{\partial x^j}\right) B^j,$$

that is, the B^j are the components of the pseudovector $\mathbf{B} = o(x) B^i \frac{\partial}{\partial x^i}$. This means we have de-densitized \mathcal{B} and made it into a pseudovector. This is, implicitly, what beginning physics texts do when they say that the magnetic field is a pseudovector: they are taking what the magnetic field most naturally is, i.e. a 2-form, and converting it to a vector by this trick. It is just that in Euclidean space, $\sqrt{g} = 1$, so that, just as with contravariant vs. covariant ($g_{ij} = \delta_{ij}$), the components of the “modified” vector are exactly the same. That's why claiming that vectors *are* their components is a problem, because in special cases like \mathbb{R}^3 , it fails to distinguish these cases. Plus, since \mathbb{R}^3 is orientable, we can also even get rid of the

“pseudo,” which is why even more crass³⁰ elementary texts don’t mention the true nature of the “vector” \mathbf{B} at all.

As another example which we gave above, we can turn the Levi-Civita tensor capacity (the invariant object defined by the permutation symbol) by multiplying the components with \sqrt{g} instead (because it is a capacity rather than a density). Namely, if

$$\epsilon'_{i_1 \dots i_n} = \epsilon_{i_1 \dots i_n} \sqrt{g},$$

then

$$\epsilon' = \epsilon'_{i_1 \dots i_n} o(x) dx^{i_1} \otimes \dots \otimes dx^{i_n} = \epsilon_{i_1 \dots i_n} o(x) \sqrt{g} dx^{i_1} \otimes \dots \otimes dx^{i_n}.$$

defines a pseudotensor. But recall that from the definition of $\epsilon_{i_1 \dots i_n}$, we have $\epsilon_{i_1 \dots i_n} dx^{i_1} \otimes \dots \otimes dx^{i_n} = dx^1 \wedge \dots \wedge dx^n$. This means that ϵ' is the Riemannian honest density itself! In other words, the de-capacitization of the Levi-Civita permutation yields the ordinary pseudotensor components of the Riemannian honest density. This kind of identification is allowable for any (nonvanishing) honest density, not just the Riemannian one. Sometimes other structures will yield an honest density, for example, a symplectic structure (but, since symplectic manifolds are always orientable, this often gets combined with orientability).

All of the preceding three ways of identification are, actually, the reason why the theory is confusing---because of the presence of all of these structures in Euclidean space, the tendency was, when everyone thought of things as being their components, to implicitly use these identifications, sometimes without even knowing it. It is only after the formalization of the theory, to see what is doable and what isn’t in general, did this more sophisticated point of view emerge. And it is advantageous, both from a physical and a calculational standpoint, to make distinctions where none was made before. Even though covariant and contravariant quantities can be identified in Riemannian geometry, Riemannian geometers still try to keep the distinctions straight, because sometimes the distinction is useful. It is conceptually useful, for example, to view a 1-form’s action on a vector as putting the arrow through the stack that represents the 1-form, and counting the sheets, than it is to have to take orthogonal projections.

An explanation for why a metric, orientation, or honest density allows for these identifications is that each process is reducing the “admissible” coordinate transformations to a subgroup of $GL_n(\mathbb{R})$ (this is related to the representation theory explanation for all our species of tensors). The metric reduces things to $O(n)$, yielding invariance laws when only length-and-angle-preserving transformations are considered. So, for example, we need not distinguish between length and inverse length. Orientation allows us to restrict to $GL_n^+(\mathbb{R})$ of transformations with positive determinant, thus removing the pseudo-. An orientation along with a metric restricts to $SO(n)$; eliminating the reflections, and an orientation along with an arbitrary nonvanishing honest density (i.e. giving a volume form) allows us to restrict to $SL_n(\mathbb{R})$. And only an honest density restricts us to transformations of determinant ± 1 (I don’t know why this doesn’t have a special name. . .).

3.9. Flux and Integration Over Hypersurfaces. As hinted before, even with pseudoforms, we still will have trouble with integration over submanifolds, because pseudoforms cannot be pulled back, in general. In order to see what is going on here, we revisit one

³⁰Ok, I’m being a little harsh. Practically speaking, for computational or engineering applications, this way of thinking of \mathbf{B} is ok, especially for a first course. One could imagine how much lower attendance to introductory physics courses would be if we threw this at them. But if anyone wishes to dig deeper, this version is better. For example, we may not have to introduce differential forms, but we could say that, for example, Faraday’s field lines are not just another pretty-picture way of thinking about the magnetic field, but a true reinterpretation of the vector concept. See the section on flux for more detail.

of the most venerable physics concepts traditionally associated with vector calculus: flux. Given a fluid with density ρ and velocity field v , they combine to form the flux density field ρv . Then, over an oriented surface in 3-space, flux is the surface integral of the normal component of ρv , which gives the mass of fluid passing through S per unit time. The key is in the words “normal component”---orientability of surfaces embedded in \mathbb{R}^3 is usually specified in terms of being able to find a continuously varying unit normal field on the surface. The Möbius strip does not admit such a field, because if we run a unit normal around a loop, it will come back pointing in the opposite direction (and thus, if we insist on using this normal field, it will have to have a big discontinuity somewhere, where all of them have to flip around).

Actually, the equivalence of orientability of submanifolds and finding continuous normal fields only exists when the ambient space (here, \mathbb{R}^3) is orientable! The condition for finding a continuous normal field is called **two-sidedness**. The Möbius strip is said to be one-sided precisely because of this. But as it turns out, the relationship between orientability and two-sidedness is much like the relation between pseudo- and true tensors.

If you think things through carefully, it should make sense that notion of flux *should* require two-sidedness; otherwise how are we to tell the net flow of fluid through the surface? Let us think things through in the general case. Imagine fluid flowing through a manifold. It turns out that one does not need a metric to talk about the “(mass) density” of a fluid---mass density, if you recall, is something that, when integrated over space, yields the mass of whatever is occupying in the space. So the mass density of a fluid should theoretically be specified as something that can be integrated over a manifold---namely a . . . density! Hence, the name. So a fluid flowing along an n -manifold should have a mass density specified by an n - ψ -form (since M is not necessarily orientable, a ψ -form is the only thing we can guarantee to have a well-defined integral) which we’ll call the **honest mass density**. How does this relate to the ordinary “scalar” quantity called density (and usually denoted ρ ? Well, for Riemannian manifolds, recall that we can turn any function into an honest density by multiplying by $|d\mu_g|$. So the ψ -form describing the fluid density is $\rho|d\mu_g|$. But in general, without a metric, we cannot specify it by a *function*. The total mass of the fluid with honest mass density ω is simply $\int_M \omega$. If ω never vanishes, of course, we actually end up producing something that can be used in place of a Riemannian density (i.e. we can use it to integrate other functions, define the Hodge $*$ operator---we’ll get to that---etc. In particular, if some function is defined pointwise as yielding something per unit mass, we can use the honest mass density to integrate that function to get the total something, without the per unit mass).

Now how do we generalize *flux density*? Recall the motivation in \mathbb{R}^3 . Given a fluid with “scalar” density ρ and velocity field v , one can imagine the flux through a surface to be calculated by breaking it up into lots of little (infinitesimal) parallelograms. The flux through one of these little parallelograms, $\xi \wedge \eta$, can be measured by examining, in a short time δt , the mass contained in the parallelepiped determined by $v\delta t$, ξ and η . The (approximate) flux is then just given by dividing by δt (we let $\delta t \rightarrow 0$ to make it exact).

(PICTURE)

The volume of the parallelepiped is, of course, given by the Riemannian density evaluated on it; multiplying by the density function ρ then gives the mass. In other words, to get the flux through the little parallelogram, we evaluate the density form $\rho dx \wedge dy \wedge dz$ on v , ξ , and η . To get the total flux, then we have to integrate $\rho dx \wedge dy \wedge dz(v, \xi, \eta)$ over all the little parallelograms $\xi \wedge \eta$. This means the 2-form defined by inserting v into the first slot, namely, $i_v(\rho dx \wedge dy \wedge dz)$ is the appropriate integrand, the flux density of the fluid.

So, on a general manifold, it suggests that the **flux density** of a fluid with honest mass density ω and velocity field v should be given by $\Phi_{\omega,v} = i_v\omega$, where ω is its honest mass density form (again, $i_v\omega$ is the $(n-1)$ - ψ -form defined by $i_v(\omega)(X_1, \dots, X_{n-1}) = \omega(v, X_1, \dots, X_{n-1})$).

So this says, in general, any kind of flux density is going to be given by an $(n-1)$ - ψ -form. Can we integrate it over an embedded submanifold (specifically, a hypersurface, because it is one dimension smaller)? As we said, we need a diffeomorphism to pull back a form. Now on a general submanifold, if we don't have a metric, we can still speak of two-sidedness. We don't need a vector field to be *normal* to the hypersurface S , only that it be *transverse* i.e. not tangent to S . We say a hypersurface S is **two-sided in M** if it admits a smooth (or just continuous) transverse vector field N defined on it. This means, in particular, the field must never vanish, since 0 is in every tangent space. Let's see why two-sidedness (transverse orientability) on a Möbius strip S is equivalent to the usual normal field condition. First any normal field is obviously a transverse field. Conversely, any transverse vector field could be projected to the true normal bundle NS , and because $T\mathbb{R}^3 = NS \oplus TS$, a direct sum, so any transverse field must therefore always have a nonzero normal component. So a transverse vector field to the Möbius strip would define a nonvanishing normal field.

The point is, a transverse vector field V to a hypersurface can be used to define a pullback operation for $(n-1)$ -forms. Given η an $(n-1)$ - ψ -form on a manifold M and a 2-sided hypersurface S with transverse field V , we can pull η back to S as follows. If we simply restrict η to S , then η will still have values in ψM rather than ψS . That's what goes wrong with just naive pullback. It makes a difference since the generators of each $\psi_p M$ are the two half-lines in $(\Lambda_n)_p M \setminus 0$, and $\psi_p S$ consists of the connected components of $(\Lambda_{n-1})_p S \setminus 0$, two different spaces.

Now, if S is embedded, at a point p there is a coordinate chart (U, x) in M about p such that $U \cap S = \{p : x^n(p) = 0\}$, the zero set for the last coordinate function x^n . This means, in particular, that the first $n-1$ coordinate vector fields span $T_p S$ for all $p \in U \cap S$. The orientation of each $T_p M$ determined by these first $(n-1)$ coordinate basis vectors, together with V , defines a unit-norm section---we'll just write $o(x, V)$ for it. It must be smooth, otherwise, by the Intermediate Value Theorem, the continuous function defined by the determinant of the change-of-basis matrix from $\frac{\partial}{\partial x^1}, \dots, \frac{\partial}{\partial x^{n-1}}, V$ to $\frac{\partial}{\partial x^1}, \dots, \frac{\partial}{\partial x^{n-1}}, \frac{\partial}{\partial x^n}$ would have to vanish somewhere, meaning that V would be tangent to S at that point). On the other hand, there is a unit-norm section, write it $o(x|_S)$, determined by just the first $(n-1)$ coordinate basis vectors without the V . Thus over this patch, $o(x, V)\eta$ is a true form. It can be pulled back by the inclusion ι . Then we multiply the result by $o(x|_S)$. So in short we define

$$\iota_V^* \eta = o(x|_S) \iota^*(o(x, V)\eta).$$

We have to show this is independent of the submanifold charts (however, it is dependent on the transverse field V). This isn't too difficult: for another coordinate chart (W, y) , such that $W \cap S = \{y^n = 0\}$. The transformation from the x and V frame to that of the y and V frame just looks like, on $U \cap W \cap S$,

$$\begin{pmatrix} \left(\frac{\partial x^\alpha}{\partial y^\beta} \right) & 0 \\ 0 & 1 \end{pmatrix}$$

where α, β runs over 1 to $n-1$. The 0's are present because both the $\frac{\partial}{\partial x^\alpha}$ and $\frac{\partial}{\partial y^\alpha}$ span $T_p S$ and thus have no component of V . This determinant is exactly equal to that of just the transformation from the first $n-1$ components alone (expanding along the final row, say).

Thus in an overlap we once again have two like sign factors which cancel each other, so it is invariant. The ψ -form $\iota_V^* \eta$ is what we can integrate over S .

It of course depends on V ---but that's what we want; intuitively, the choice of V is picking one side of the two-sided hypersurface S ! We should check, however, that there indeed are only two choices (for connected S). This is also easy: given another transverse field W , we check, in a submanifold chart, the sign of the determinant of the change-of-basis matrix between the coordinate frame with the last component replaced with V , and the same frame with the last component W . The pullback only is affected by this sign, and this is a locally constant function, i.e. constant on the connected components of S . In particular, if S is connected, pulling back via the transverse fields V and W can make a difference of at most a sign.

The use of transverse vector fields in integrating pseudoforms over hypersurfaces can, of course, be generalized. This is most nicely described in terms of orientability of *vector bundles* over M , and overcomes the moderately nasty notational problems that have come up here.

3.10. The Cross Product. This is somewhat of a bonus section, because it isn't really essential to the flow of things. It is a demystification of what the cross product really is, and what it isn't. (To be continued).

More work to be done:

- Weinreich's examples: the cross product in \mathbb{R}^3 , curl of a vector field.

3.11. Orientability of Vector Bundles: Bonus Section! Let $\pi : E \rightarrow M$ be a vector bundle. Then each fiber E_p , being a vector space, also has two orientations, namely o_{E_p} . We can thus define ψE_p , the pseudoscalars associated with the orientations of E . Contrast that with all the ψ -scalars we have been talking about, which are really elements of $\psi T_p M$. It turns out we can bundle all the ψE_p , into a bundle $\psi(\pi : E \rightarrow M) = \psi(\pi) = \psi(E, M)$. This doesn't have anything to do with $\psi M \otimes E$. Of course, $\psi M = \psi(TM, M)$.³¹ If $\psi(\pi)$ is trivial, we say E is **orientable as a vector bundle over M** . M is orientable if and only if the tangent bundle orientable as a bundle over M . This concept of orientability can also be described by transition maps---if M admits a trivializing cover for E which have transition functions all with positive determinant. The notion of orientability for vector bundles is important in the theory of characteristic classes, which are highly general measurements of how vector bundles twist in space, i.e. how they can be nontrivial.

Orientability of a bundle is actually one of the ways we can use to succinctly determine whether we can define the integral of a pseudo- k -forms over k -dimensional submanifold of M , where $k < n - 1$.

Let's first motivate things by trying to generalize the situation with the Möbius strip, using normal vector fields. Let S be a submanifold of dimension k , in a Riemannian manifold M . The inner product on each tangent space $T_p M$ gives a splitting $T_p M = T_p S \oplus T_p S^\perp$ where, recall, $T_p S^\perp$ is the orthogonal complement of $T_p S$ in $T_p M$ (recall that choices of direct-sum complements, that is, W such that $T_p S \oplus W = T_p M$ are *not* unique, but an inner product can be used to specify a unique, erm, *canonical* choice). Write $N_p S = T_p S^\perp$. Then the collection of all normal spaces to S is a smooth vector bundle NS over S . We have the following succinct definition: S is **transverse orientable in M** if NS is orientable as a vector bundle over S . Note that since NS is very much dependent on the embedding of S in M , transverse orientability is therefore not an *intrinsic* property of S , but rather a property dependent on

³¹I don't think there's a standard notation for $\psi(E \rightarrow M)$, but the essential information is encoded in the relation between E and M , not M itself. One definitely should not write $\psi(E)$ alone, because that would mean $\psi(TE, E)$ which is different from $\psi(E, M)$.

the embedding of S in M . For hypersurfaces, of course, NS is one-dimensional over S , and is thus orientable as a bundle if and only if it has a global nonvanishing cross-section. As an example, a Möbius strip $M\ddot{o}$ is actually transverse-orientable (or two-sided) in the manifold $M\ddot{o} \times \mathbb{R}$, because, using, say, the product metric, the global coordinate t on the \mathbb{R} factor gives a smooth global frame $\frac{\partial}{\partial t}$, and hence, since it is orthogonal to $T_p M\ddot{o}$, it represents a global section of $N_p M\ddot{o}$.

Since every manifold has a (noncanonical) Riemannian metric, we can define transverse orientability for any submanifold of any manifold: we simply consider the orientability of the normal bundle from a metric constructed using a partition of unity. Despite the arbitrary choice involved, we can show the property of transverse orientability is independent of the normal bundle chosen. This is simply because all normal spaces are bundle-isomorphic to the quotient space $T_p M/T_p S$ at every point. That space does not depend on a metric for its definition, so in general, topologists tend to define *the* normal bundle NS to be the quotient bundle TM/TS . We will stick to our definition, though. Dude, a totally awesome!!!! theorem in differential topology says NS is diffeomorphic to a **tubular neighborhood** of S in M . Hang ten!

A sufficient (but it turns out, not a necessary---it is only necessary in codimension 1) condition for NS to be orientable is if S admits a global smooth transverse frame in TM which span NS at every point. The reason why it is not necessary in general is because we actually only need S to be covered by trivializing patches for which all the transition functions have positive determinant (the definition of orientability of the bundle), and that the frames defining the trivializations do not have to smoothly patch together; only the determinant of the transition functions have to. Just as in an orientable tangent bundle (M itself being orientable), a specific choice of orientation in an orientable bundle E over M is specified by a smooth unit-length section into $\psi(E, M)$. So in particular, most succinctly, S is transverse-orientable if and only if S has a nonvanishing (or unit-length) section $\tau : S \rightarrow \psi(NS, S)$. This definition works even for the abstract definition $NS = TM/TS$, without worrying about any Riemannian metrics at all. An **orientation of NS over S** , or **pseudo-orientation of S in M** is then given as a particular choice of section into $\psi(NS, S)$.

So, if η is a ψ - k -form on M and S is a transverse-orientable submanifold with transverse orientation $\tau : S \rightarrow \psi(NS, S)$, we can pull η back to S , and thus, in particular, integrate η over S . This is done as before. At p , choose a submanifold chart at p as before. Then the first k vectors of the chart combine with $n - k$ vectors representing $\tau(p)$ to form a parallelepiped in $T_p M$, whose orientation class we'll multiply by η , to de-pseudoize η , which we can then pull back. Then the pseudoness is put back on by multiplying again by, now, the orientation class the first k vectors (parallelepiped in $T_p S$).

(more details).

4. DUALS, DOUBLE-DUALS, AND CANONICAL TRANSFORMATIONS

“Insert witty quote here.”---Anonymous

Ok so mathematicians and physicists love to talk about all these vector dualities and such. Duals pervade the theory of mathematics, anyhow. Basically the realization of a dual concept is almost always in the notion of *adjoint*. Studied in its primitive form, i.e. on Hilbert spaces, for example, we have an inner product and study when can we get an operator H^* , given H , that does something like $\langle Hv, w \rangle = \langle v, H^*w \rangle$. Much more general notions of adjoint exist. On suitable spaces, when evaluating a function on something, like $f(x)$, it could in theory be written more symmetrically as $[f, x]$ for example. Especially when

the f 's live in a similarly dimensioned space as the x 's do. Then we might be able to define operators on f 's as adjoints of operators on x , that is $[H^* f, x] = [f, Hx]$. The transpose of a linear transformation is an example: traditionally it is given $T'(g) = g \circ T$ or for all v , $T'(g)(v) = g(Tv)$. But written in our spiffy new notation, it really is $[T'g, v] = [g, Tv]$. An adjoint, see? The metric identifies V and V^* by simply declaring $[G, C] = \langle G^b, C \rangle$ i.e. it converts $[,]$ to \langle, \rangle via the metric. But b is adjoint to \sharp and so $\langle G^b, C \rangle = \langle G, C^\sharp \rangle$ (ok I chose G and C just so that it would be more plausible, music-wise).

The notion of Hodge duality uses the Weinreichian duality construction along with a Riemannian ψ -density to generate what is called the **Hodge star operator**

$$* : \Lambda^m \rightarrow \Psi \Lambda^{n-m}$$

by $\eta \wedge \omega = \tilde{g}(\eta, *\omega)|d\mu_g|$, where \tilde{g} is the metric on Λ^n induced by the usual metric g , and $|d\mu_g| = o(x) \sqrt{g} dx^1 \wedge \cdots \wedge dx^n$ the volume form. The nondegeneracy of both \wedge and inner product allows us to define it in this non-explicit form.

More work to be done:

- Why the double-dual isomorphism is canonical and the single dual is not.
- Categorical bollocks on canonical transformations in general.
- Transformation laws are preserved by canonical transformations.

5. BILINEAR FORMS, AND THEIR RELATION TO METRICS AND EIGENSTUFF

APPENDIX A. A NOTE ON TEXTS

“One way to learn a lot of mathematics is by reading the first chapters of many books.”---Paul R. Halmos

Ah, the dreaded discussion of texts for tensor analysis. I am addicted to collecting math books (I also often have 10 books checked out from the library simultaneously) and reading the first 20 pages of them. Very occasionally I make it through farther than that. Despite the proliferation of bad tensor analysis texts (some would say all of them are bad), I have to admit I have gleaned everything I have learned about Tensor Analysis from reading these books, collecting the knowledge into a gigantic patchwork. The fact is, each text actually has a gem of insight or two that is not presented in any others. There is much overlap in the bad parts, and some in the good parts too, but of course, it's always hard to consult so many references, since I often forget the transformation laws on those overlaps. . . not to mention also the transformation laws that tell how the notation changes---regarding this, we have the following

A.1. **Joke.** Differential geometry is the study of those things invariant under change of notation.

Another problem is also that it's very hard to strike a balance between being intuitive in the explanations of what these things are---and hence being vague---and also developing a precise, formal theory that is the real deal---hence being obfuscatory. The need for contravariance and covariance seems to be a point that does in fact get explained well---you're made fully aware of these things, and perhaps, like me, become really paranoid about when you're implicitly using that damn identification.

All right, so let's talk about texts anyway. I'll say what I like, but there's no guarantee that you'll like them. After all, tensor analysis texts are bad, right?

I must mention, first and foremost, Weinreich [We97]. I don't think this text is very well-known; I'd only happened to see it while online one day back when I was learning

vector calc. Without this text I probably would not have gotten into this in the first place. I can't overestimate the impact this text has had on me, because without it, *I might have never actually become a mathematician*. Weinreich gives geometric interpretations of many of the missing concepts in tensor analysis texts---the fancy-schmancy stuff, and applications of this pictorial approach to electromagnetism. Unfortunately it is very brief and only describes the rank-0 and rank-1 case---scalars, vectors, and covectors. Also, although he has the right ideas about the pseudo-business, namely with axial and polar stuff (heck, it's the only book I've ever seen such definitions even given), even he, working in \mathbb{R}^3 , doesn't emphasize the importance of separating the concepts very well; he ends up using the right-hand rule a lot anyway.

It is this text that set me on a quest to gain a deeper understanding of these concepts in a more general setting. Unfortunately I have found no analogous text which *completely* develops this mystical general setting! A text that comes close in visualization (Weinreich's text seems to have drawn some inspiration from this one. . .) is the much-lauded and also much-hated (in short, much-polarized) Bible of Gravity, [MTW] which uses colorful expressions such as "egg crates" to talk about these things. However its heftiness sort of is off-putting---it may take them forever to get to the point of some explanation (especially hard on the never-getting-past-page-twenty folk), and it might not go with your style. There are nice pretty pictures and descriptions, but there are still some unanswered questions. Maybe I am picky or too demanding. The astute reader has probably noticed that I also do not have such a great ability to get to the point.

Eric Michelsen has produced some nice notes introducing the reader to a sorely needed big picture of tensor analysis, and some down-to-earth applications. His notes are the inspiration for these.

For a good workout in the formal development of the modern theory I suggest [Lee02], which tries to be as geometric as possible while not sacrificing the formal developments. Lee goes into detail in computations, the special case of differential forms and their integration, and gives a brief overview of Riemannian metrics (but doesn't get too far into curvature---that's discussed in [Lee97]---this text, by the way gives an important exercise on an important equivalent characterization of tensor fields which we'll get to). Anyway the value of this text is that it gets the reader a good *working* knowledge of modern tensors as they are used in the mathematical theory. Once again, the knowledge gained is a more practical (of course important!) slant---tensors are these useful tools we must use, so it leaves some deeper meaning unanswered. One doesn't need to know *all* the vagaries of tensors to actually get work done with manifolds. He says nothing about fancy-schmancies although he does develop a theory of integration on nonorientable manifolds using the nontensorial "absolute value of the volume form" business. John Baez insists in [BBF] that this is strictly incorrect and that there is a true way of regarding this as a true tensor quantity and not something violently abused via the horrid absolute value signs.

Robert Wald's GR text [Wa84] contains a very fast-paced development of modern tensor notation in the context of general relativity (he uses Penrose's abstract index notation). No fancy-schmancy here.

The text by Dodson and Poston [DP90] may actually be the Tensor Intuition Holy Grail, but which I have not actually had the time to really read through (I'm swamped enough as it is), but whose insights may eventually make it into these notes. From the little I have read of it, it has some very interesting geometric discussions and emphasizes the reason why covariant differentiation is something necessary: differentiation of a vector field leads to a vector. . . "out of this world." One thing it does is formulate tangent spaces as

affine spaces---a concept which gauge-symmetry-happy physicists might find familiar. The reason for this is that in order to differentiate, it is not so much the ability to add that is important---it is the ability to *subtract* (and divide by small numbers) that is important. It ain't called *differential* calculus for nothin'. The index provides no evidence of the existence of fancy-schmancy tensors, however.

Another candidate bridge-book is one that a physicist might well love: Theodore Frankel's book [Fr04], who goes through pains to get down to things invariantly. There's a lot of good stuff in this book, and I would recommend it just for its overall great explanation of the general stuff. In terms of tensor analysis, though, he gets happy about all contra- and co-business and is takes pains to separate and indicate their use. In addition, the volume form is treated as a structure-giving entity in its own right. He mentions only the pseudo- part of the fancy-schmancy tensors and makes somewhat of an argument for their true necessity, but the explanation is a bit confusing and it is not really developed in enough detail, only sort of glossed over. He does mention that the correct way to do integration on a nonorientable manifold is by pseudo-forms (it's really the volume *pseudo*-form, we learn) and not using absolute value business. But, once again, it seems that he still sweeps some of it under the rug, because most of the time he just refers to the volume *form* and mumbles something about tacitly assuming some orientability. It is amazing, though, that pseudo-stuff actually gets mentioned as a genuine phenomenon.

How long could I go without mentioning Spivak's comprehensive reference on differential geometry? I had discovered that densitization was possible via tensor products in his book, although only in an end-of-chapter exercise, and the concept subsequently left undeveloped. However this is only in Volume 1 of the book. I have not quite had the time to really dig in to his other volumes where they could conceivably be really put to good use. Because, if that's all he has to say about tensor densities, then perhaps it does not live up to its title, *Differential Geometry: A Comprehensive Introduction*. Ok maybe that's a bit harsh, and I'm really picky. . . because the text is otherwise very praiseworthy again for the general stuff.

I have to give some accolades to the book by Bott and Tu [BT82] which is a very nice book in general. Raoul Bott is (or was, for he passed on during the writing of these notes) an electrical engineer turned mathematician. So he had to have some appreciation of all the pseudo business. In the book they give a brief overview of it, in the context of generalizing some very fancy (and important) topological theorems such as Poincaré duality to non-orientable manifolds. They construct pseudotensors in the same way Spivak does, namely by declaring that pseudoscalars are a vector bundle with transition given the sign of the determinant of a coordinate change. This is much better than using absolute values on the tensors themselves; it is impossible to define the de Rham theory (necessary for Poincaré duality) without it (absolute values, last time I checked, weren't differentiable at 0).

Finally, I spent all my Borders gift-card money on Roger Penrose's enormous new tome [Pen05], supposedly written for the "serious lay reader." That is, if you define "lay reader" to mean "7th year mathematics and physics graduate student." Ok maybe not that bad. But seriously, some of the concepts in there are things that I learned just last year, and would have been severely put off from reading. I haven't learned about gauge connections and fibre bundles at all yet (though of course I have some idea of what they are). He goes through everything---trying to intuitively and geometrically explain calculus, complex analysis, manifolds, and of course tensors---and mentions densitization. But of course, such a good geometric overview must come at the price of precision and, of course, the deeper development we seek.

All right. That is enough babbling about the books. Enjoy tensor analysis!

APPENDIX B. ON HEARTBREAK: A PARODY ON PHYSICS TEXTS AND TENSORS

B.1. Introduction. Heartbreak... a miserable, miserable thing. One of the long-standing questions of many scientists in love is: *How do you quantify love? And the heartbreak that's supposed to come with it?* Of course many would scream loudly at this point that many aspects of the human experience should never be disgraced by assigning a number to it, for it would condemn human qualities to analysis by that cold, hard, soulless subject known as mathematics. Who would think of doing something so wretched?

Me, of course. Mathematics is neither soulless, nor does assigning numbers to things or measuring things necessarily dehumanize them. Instead mathematics should be part of the human experience. Anyway enough “mathematical moralizing.” Here we describe a way to quantify heartbreak. (Remember, students, that this is *only a theory*, and hence you are encouraged to approach this subject with an open mind...)

We define the standard unit of heartbreak, the **harmony** (symbol Hm) to be the amount of heartbreak dealt to me by Harmony. All the standard prefixes apply. For example, Harmony dealt me 1 harmony = 1000 milliharmonies = 10^6 microharmonies of heartbreak.

Small units of heartbreak (milli-, micro-, nano-, and so forth) are useful for dealing with ordinary offenses such as insults, getting cut off in traffic, getting dirty looks, and so forth. For example, some random girl gave me the finger the other day for turning from the wrong lane. That constitutes perhaps 1 nanoharmony of heartbreak.

B.2. Stendhal's Law and Sensitivity. As it stands the unit of heartbreak is virtually useless to the casual reader. Few would know who Harmony even is. Thus, as it stands, one would have no idea how to actually measure heartbreak on their own terms. However an important result in the theory of heartbreak, which alleviates this difficulty, is **Stendhal's Law**:

B.1. Theorem. *For ideal, linear, and monotropic lovers (such lovers are called quixotic), the amount of time to get over a heartbreak is directly proportional to the amount of heartbreak dealt. Quantitatively, we can write this as:*

$$(B.1) \quad T = 4\pi sH$$

where T is the time to get over the heartbreak, H is the amount of heartbreak dealt, and s is a constant of proportionality called the coefficient of sensitivity, and the 4π is there solely for the purpose of making the formula have a π in it.

The coefficient s is characteristic to the lover. Why is s called the sensitivity? It is roughly because more sensitive individuals take longer to get over heartbreaks, i.e. for fixed H , a large s means a larger T .³² A common stereotype is, then, men have small s , while s is relatively large for women.

Now this defines a natural standard unit of sensitivity: the **ch0ndawg**. 1 ch0ndawg is defined to be the amount of sensitivity that makes 1 harmony of heartbreak correspond to 4 months of time, i.e. the time it took me to get over Harmony. It therefore follows that 1 ch0ndawg = $(1/\pi)$ month/harmony. This is still not quite useful for the reader, however. This is because the collection of three units, (harmony, ch0ndawg, month) form a type of

³²In strongly curved spacetime, the so-called “constant” 4π can vary. Not the π , of course, but rather an expression involving the Riemann curvature tensor. The formula is quite ugly. One can think of this in terms of the following example: in the vicinity of a black hole, one's heart can be assumed to be literally breaking from holding up against strong tidal gravitational forces (i.e. dealing with the large curvature of spacetime), so of course it would result in a larger coefficient. Since gravity and hence spacetime curvature in the neighborhood of our solar system is relatively weak, we can safely ignore this aspect.

coordinate system of heartbreak called an **egocentric** coordinate system. Perhaps a more useful coordinate system could be based on, say, national surveys or averages. Indeed, one of the most useful coordinate systems is based on the notion of “coolness.” That is we could establish standard units for the average person considered to be “cool” since, well, people like to be cool. This unit system is called the **kewliocentric** coordinate system.

B.3. Explanation of the Terms. We have not explained the terminology *ideal*, *linear*, and *monotropic*. Let us treat each of these conditions in turn, and point out their flaws. *Ideal* means, just that, ideal, which means the heartbreak dynamics of any real lover cannot be described perfectly via Stendhal’s Law. *Linear* means that the law is a proportion in the first place, i.e. s is not an explicit function of H (one’s sensitivity is not affected by the heartbreak one receives). Finally, *monotropic* means that the lover only obsesses over one person (henceforth referred to as a *victim*) at a time.

We can swiftly take care of the issue of ideality by saying that in most cases, possibly with the modifications we will later mention in conjunction with the other terms, Stendhal’s law is a very good approximation.

A deviation from monotropicity, i.e. lovers with multiple victims (*polytropic* lovers), is easily treated. Lovers who treat their victims equitably can still be considered monotropic but we must make H a vector quantity and the equation is now just involving a scalar multiplication. More usually the case, polytropic lovers treat each victim differently, and moreover love of one victim may have some influence on the others (i.e. the victims are not independent of one another). In this type of scenario, a variant of Stendhal’s Law still holds (still assuming linearity). The situation here is more complex and we shall, for the rest of this discussion apart from this remark, only consider monotropic lovers (without loss of too much generality). The variant of Stendhal’s Law which holds for polytropic lovers is given by simply considering s to be a rank-2 tensor called the *sensitivity tensor* of the lover. For lovers who become more aggressive as their victims protest, this tensor is rank-2 contravariant, and lovers who back off from such protests have a rank-2 covariant sensitivity tensor. Finally the lover that persists unchangingly has a sensitivity tensor of mixed variance. So the law would be

$$T^j = s_i^j H^i$$

Now for the linearity itself. In real lovers, of course, sensitivity is affected by heartbreak---individuals become bitter, cynical, etc. Severe effects of course are hopefully temporary. . . at least within a certain “linearity regime” of heartbreak. However time evolution does introduce permanent changes, not necessarily due to the various negative things listed above---it could be something as simple as becoming more mature and more adept at handling situations such as heartbreak. Nevertheless it is useful to fix a period in one’s life as a universal basis to judge and compare heartbreaks, if for nothing other than historical interest. For the egocentric (harmony, ch0ndawg, month) system the standard is the **Epoch C1999.08 Egocentric** system, which refers to the date in which the unit of harmony was established.³³

³³A tacit assumption, aside from the three mentioned, is that all our lovers are in fact *sexually orientable*. It is important to be able to formulate the theory without this assumption because, in fact, non-sexually orientable lovers do exist, e.g. bisexuals. It is rumored that someone who had an enormous crush on the mathematician August Ferdinand Möbius in fact was non-sexually orientable. Stendhal’s Law still holds, with appropriate modifications, for nonorientable lovers, but theory is made more complicated because we must consider pseudo-quantities---quantities that have an extra sign-reversal under coordinate inversions (the main culprit in the formulation of the various laws involves the devakian/michellian $*$ operator). However since most computations take place in one coordinate chart, I shall, like most physicists and mathematicians when dealing with pseudo-quantities, happily sweep this issue under the rug.

REFERENCES

- [BBF] John C. Baez, Toby Bartels, Eric Alan Forgy, et al. “Densitized Pseudo Twisted Forms.”
http://groups.google.com/group/sci.physics.research/browse_frm/thread/6a231426b3a313c0/3904dcfcc4043739.
 Usenet newsgroup `sci.physics.research`: 2001-2003.
- [BT82] Raoul Bott and Loring W. Tu, *Differential Forms in Algebraic Topology*, Graduate Text in Mathematics **82**. New York: Springer-Verlag, 1982.
- [DP90] Kit Dodson and Tim Poston, *Tensor Geometry*, 2nd Ed., Graduate Texts in Mathematics **130**. New York: Springer-Verlag, 1990.
- [Fr04] Theodore Frankel, *The Geometry of Physics: An Introduction*. 2nd Ed. Cambridge: Cambridge University Press, 2004.
- [Lee97] John M. Lee, *Riemannian Manifolds: An Introduction to Curvature*, Graduate Texts in Mathematics **176**. New York: Springer-Verlag, 1997.
- [Lee02] John M. Lee, *Introduction to Smooth Manifolds*, Graduate Texts in Mathematics **218**. New York: Springer-Verlag, 2002.
- [MaT97] Ib Madsen and Jørgen Tornehave, *From Calculus to Cohomology: de Rham Cohomology and Characteristic Classes*. Cambridge: Cambridge University Press, 1997. This title inserted for the purpose of having an author whose name has an o in it.
- [Mic06] Eric L. Michelsen, *Funky Mathematical Physics Concepts*. Spectacular set of notes by a UCSD Physics grad student, 2006.
- [MTW] Charles Misner, Kip S. Thorne, and John A. Wheeler. *Gravitation*. New York: W.H. Freeman, 1973.
- [Pen05] Roger Penrose, *The Road to Reality*. New York: Alfred Knopf, 2005. Chapters 12-15.
- [Scy97] Harry M. Schey, *div, grad, curl and All That: An Informal Text on Vector Calculus*, 3rd Ed. New York: W. W. Norton & Co., 1997
- [Sp03] Michael Spivak, *A Comprehensive Introduction to Differential Geometry*, Volume 1, Chapter 7. Publish or Perish, 2003.
- [Wa84] Robert Wald, *General Relativity*, Chapters 1-2. Chicago: University of Chicago Press, 1984.
- [We97] Gabriel Weinreich, *Geometrical Vectors*, Chicago Lectures in Physics. Chicago: University of Chicago Press, 1997.
- E-mail address: ctiee@math.ucsd.edu*