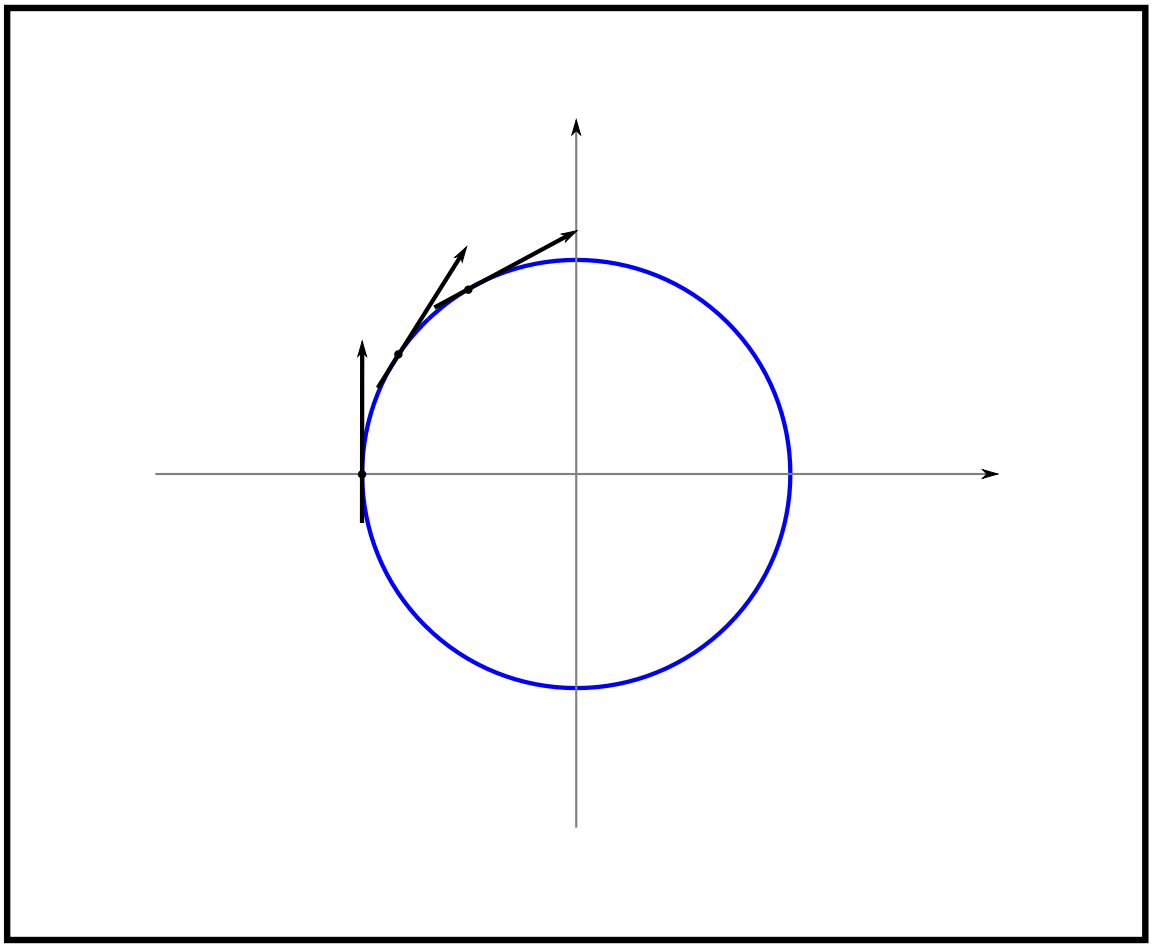
Chains, forms, and duality:

A novice's guide from vector calculus to manifolds

John Kerl

February 3, 2008



Contents

Preface

The art of doing mathematics consists in finding that special case which contains all the germs of generality. — David Hilbert (1862-1943).

To Thales the primary question was not "What do we know?" but "How do we know it?". — Aristotle (384-322 B.C.).

The following is a novice's guide (written by a novice, in mid-journey!) for the often-twisting path from vector calculus to smooth manifolds. The subject matter is roughly that of the University of Arizona Mathematics Department's first-year graduate course in geometry/topology, 534A-B.

This is a supplement, rather than a standalone reference. One may consult Lang's *Algebra*, Spivak's *Calculus on Manifolds*, Boothby's *Introduction*, Frankel's *The Geometry of Physics*, Lee's *Introduction to Topological Manifolds* and *Introduction to Smooth Manifolds*, and other sources as cited throughout.

My intention is to gather together precise definitions and thorough examples in a form that is concise and well-organized, or at least well-organized. I emphasize specific computations and examples that support and motivate theorems which are proved elsewhere, spelling out important details which are all too easily neglected in more formal treatments. It is precisely these omitted, taken-for-granted details which are the primary stumbling block for the first-year graduate student. I believe these stumbling blocks can be identified and smoothed out, becoming paving stones on the student's road to success. (In part, this paper is an as-yet-incomplete attempt to introduce the powerful techniques of reform calculus [**HHGM**], including the Rule of Four, into the graduate environment.)

In particular:

- This paper carefully revisits vector calculus, tying the known to the unknown.
- Substantial attention is given to tensors on finite-dimensional vector spaces, before the development of bundles, sections and forms.
- Topology is put in its proper place before geometry.
- Geometry is emphasized; I have given geometrical motivation for symbolic manipulations whenever I have found it possible to do so.
- Exercises are worked through in full detail. Rigor is the mouthwash of mathematics, not the meat and bread: when we are new to a subject, whether we are six or sixty, we learn by imitation.

This document is a work perpetually in progress.

Goals for the course

We have two main goals in Math 534A-B:

(1) To put vector calculus on a more rigorous footing. In particular, expressions of the form

$$df = \frac{\partial f}{\partial x}\,dx + \frac{\partial f}{\partial y}\,dy + \frac{\partial f}{\partial z}\,dz$$

were historically defined in terms of infinitesimals. Our goal is to *redefine* these familiar symbols in a rigorous way, in terms of linear approximations and vector spaces. In particular, the grand goal is a generalized Stokes Theorem.

(2) To be able to do calculus on non-linear spaces, e.g. the surface of a sphere or torus. In particular, the latter arise as configuration spaces of physical systems.

That is, we are re-thinking everything we thought we understood about vector calculus, *and* we doing it on a donut. As a student, I can attest that this dual goal gives students no small amount of difficulty.

Another irony is that in the undergraduate curriculum, calculus is one of the first subjects taught, whereas our present re-working of calculus requires a lot of math beyond calculus — linear algebra, multilinear algebra, abstract algebra, metric spaces, and quite a bit of new terminology. The payoffs are that (1) old concepts, which we once learned by rote, now become lucid and elegant, and (2) we will be able to apply calculus far beyond the reach of the freshman approach (see for example [**Pen**]).

Goals for the paper

The main, but not sole, purpose of this paper is to organize and annotate the content of Math 534A-B.

If one wishes only to succeed on the geometry-topology qualifier, perhaps one can simply learn examples by rote. Since my goal is complete mastery of these subjects, though, this paper takes a thorough approach toward the abstract algebra which I claim is the majority of the content of the course.

Another goal of this paper is to illuminate and unify the various notions of "tensor" which appear in math and the sciences. I am particularly interested in understanding change of coordinates and "index gymnastics". (That is, it seems that mathematicians speak an entirely different language than scientists do. We can be bilingual; it will benefit all parties involved.) While this is not qualifier material, it is a main goal of my graduate education. Furthermore, this discussion is so intertwined with the main body of this paper that it seemed to me to be foolish to separate it. Non-qualifier sections of this paper are clearly marked as such.

Do the contents of this paper form a redundant replication of pre-existing textbooks? Yes in part, and no:

- I want to explicitly state preliminaries (e.g. linear, abstract, tensor, and homological algebra, and category theory) which were going unstated, or were stated unclearly, in the course.
- One learns by writing: the *selection* of material winnows the important from the unimportant, and the *organization* of material forces a clarity of thought which is otherwise unattainable.
- I focus specifically on topics needed for UA's 534A-B course and geometry-topology qualifier. This invites a best-of approach toward several different texts.
- The contents of this paper are things I *need* to know for the geometry-topology qual, and which are *possible* to recall on qual day. I am *not* interesting in reciting technical details of proofs which I won't remember (details which authors of textbooks can and should include); I *do*, however, want to have a single point of reference.

Structure of the paper

- Necessary **preliminary concepts** from calculus, algebra, analysis, and topology are included in sections 1 through 4, rather than being interleaved throughout the geometry-topology material per se. The intents here are to summarize the key points and to present unified notation.

- **Chains** (which generalize the paths, surfaces, and volumes that appear in integrals in vector calculus) are a *topological* concept. Here (section 5) we need **topological manifolds**; we encounter homotopy, chains, and homology.
- **Forms** (which generalize the integrands and differentials that appear in integrals in vector calculus) are a *differential* concept. Here (section 6) we *add* a **differentiable structure**, or **smooth structure**, to topological manifolds. Here we encounter tangent bundles, forms, and cohomology.
- **Integration** combines homology and cohomology. Here (section 7) we encounter **duality** and **pairings**.
- **Flows** (section 8) are the generalization of ordinary differential equations to manifolds.
- **Lie derivatives** and the **Cartan calculus** (section 9) are, for this paper, a collection of notations and manipulations which facilitate compuations.
- Selected **problem solutions** (section 10) are given for homework and exam problems for Dr. Pickrell's 2005-2006 534A-B.
- Solutions to some old qualifying exams are in a separate paper, which you should find near this file as `prolrevqual.pdf`.

When I say *map*, I mean ...

In the context of group theory, one says merely *homomorphism* in place of *group homomorphism*. Likewise, in differential geometry, one simply says *map* in place of *smooth map*. With reference to section 4.2, one says that one is *working in the category of smooth things*: certain adjectives are omitted but understood. Throughout this paper, functions are assumed to be smooth (infinitely many times differentiable) unless otherwise stated.

Moreover, in this paper I am *working in the category of nice things*. My examples use Euclidean space or submanifolds thereof. I will not discuss infinite-dimensional vector spaces or non-Hausdorff topologies; I am not interested in pathological examples. Straightforward examples provide plenty of work for a first-year course.

Specificity of dimension

I will happily prove a result for arbitrary n by illustating only an $n = 2$ or $n = 3$ case, as long as the $n = 2$ or $n = 3$ case contains all the elements of generality. Often, the general case adds no more detail, and in fact obscures the simplicity of the presentation via its bric-a-brac of subscripts, ...'s, and summation signs. We are low-dimensional creatures; we can see, touch, and feel two and three dimensions. We can use these abilities to help us rather than hinder us.

The exemplar model

I was born not knowing and have had only a little time to change that here and there.
— Richard Feynman (1918-1988).

I am no math historian; in fact, as a second-year graduate student and a teacher with only two years' classroom experience, I am only a budding mathematician. Yet this much seems to me to be a clear pattern in the development of mathematics:

- Real-world problems are studied; mathematical models are developed.
- After some time goes by, wherein mathematicians collaborate and communicate, similarities are discovered between problems which had seemed distinct. (For example, permutations on symbols of letters and Galois' original "substitutions" [**BMA**]; systems of linear equations and systems of differential equations.)
- A system of axioms is developed, encapsulating and abstracting the common features of the once-disparate systems. (In the above examples, we obtained abstract groups and vector spaces, respectively.)
- We then study the abstract systems, proving general results and developing general methods which apply to all the original specific situations, *and* any others which might apply in the future. Herein lies much of the power of mathematics: once we are familiar with, say, abstract groups, whenever we encounter a brand-new object, if we recognize it is a group then we instantly know much about it.
- After enough time has gone by, we begin to teach the abstractions first. In a pure-math environment, one might forget (or forget to mention) the origins entirely.

So much for the *phylogeny* of mathematics — its evolutionary development across the lives of many thinkers. What about its *ontogeny* — the development of ideas within the mind of a single thinker? Like Feynman, none of us were born knowing calculus; every mathematics student and future mathematician must painstakingly relearn these ideas from scratch.

My claim is that **ontogeny recapitulates phylogeny**. This is a biological maxim[1], reflecting the fact that the embryonic development of an individual retraces (albeit inexactly) the evolutionary development of that individual's species. When I was an embryo, very early on I had gill slits — and so did you. One pair of slits became our ears; the rest closed. One might think that we, as modern humans, could reproduce more humans directly, pole-vaulting over the intermediate stages — but we do not.

Likewise in the development of ideas. When we are young, we experience many specific situations. We later learn to abstract and categorize: many once-different objects begin to go by the common name of *fork* or *dog*. (See also the Boas quote in section 4.1.4.) One might think we could pole-vault over our early development, but we do not.

Few would argue that we could hand a two-year-old a dictionary and have him or her skip over those early learning experiences. But in the mathematics classroom — blinded by the power and beauty of modern methods — we often do just that. In my eyes, the key trait distinguishing the mathematics of the 20th and 21st centuries from that of the centuries before is the *axiomatic approach*. This is a rightful component of our discipline. Yet we start the first day of an upper-division or graduate algebra course with a mantra of the form "A *group* is a set endowed with a binary operation ..."; we begin a geometry course with an abstract discussion of atlases and manifolds. This feels good and it feels right, especially to the teacher: we are going straight for the beauty and the elegance. It is especially easy for the teacher, who has already seen these concepts put to use in dozens of concrete situtations: the teacher is ready for abstraction.

But when we do this, we leave our students behind. We attempt to leap over the *absolutely necessary* instantiation of ideas that we ourselves went through (whether we did so consciously or not). If we are to retain students in mathematics, and if modern pure mathematics is to remain relevant to the rest of the world, we must explicitly acknowledge that ontogeny recapitulates phylogeny. This does not mean that we need to retrace the full development of every mathematical concept, with all its false starts and wrong turns. It does mean, however, that we *first* need to tell our students *why* a discipline was invented in the first place.

[1] Attributed to the zoologist Ernst Häckel. See also the Wikipedia article on *Recapitulation theory*.

We need to present the applied problems which gave birth to the subject we are teaching. We must give our students the opportunity to generalize from concrete situations.

One raises the objection: *I don't have time for such trivialities. Graduate students should be able to do this on their own.* I reply, as a graduate student and teacher: It is hard enough for our students to learn what we teach them; it is far harder for them to learn what we do not teach them. The training of a new mind is something worth doing right: it is better that it happen *because* of our teaching methods rather than in spite of them.

This is the one of the reasons I have found it not only entertaining but necessary to write this paper. As happened several times in my engineering career (before graduate school), I could not obtain the kinds of explanations I wanted (my questions were chiefly of the form "*But why?*"). I found myself lacking, and then writing, the references I wished I had had. This in turn led me to discover things I could not have found out otherwise. I have, as yet, found few historical references for the (phylogenetic) development of differential forms. Nonetheless, one of the primary themes of this paper is the (ontogenetic) development of tensor algebra, tangent spaces, etc. from fundamental concepts of area and measurement. I have pieced together most of these insights from my coursework; see also [**Bachman**] who has a similar approach.

Themes

I encourage the reader to keep in mind the following central themes throughout the paper:

- **Geometric** vs. **symbolic** methods[2]. We manipulate symbols because we can do so efficiently, but the reason we invented them the way we did is often geometric. For example, the deteminant can be thought of symbolically in terms of cofactor expansions of a square matrix, but it also has a purely geometric existence in terms of the volume of a parallelepiped. Likewise for the d and wedge operators of differential geometry: we can rattle off axioms for them, but why were those axioms chosen? What geometrical notions do these symbols reflect?
- **Coordinate-dependent** vs. **coordinate-free** definitions. Early differential geometry [xxx cite whom — Spivak 2?] was done concretely, in terms of coordinate charts. Any coordinate-dependent definition needed to be accompanied by a proof that it was invariant under change of coordinates. The modern approach prefers coordinate-independent definitions. These are more elegant, but not as easy for doing specific computations. [xxx back that up or retract it.]
- **Change of coordinates**; **covariance** and **contravariance**. Different objects transform differently on change of coordinates. [xxx xref to some examples within this paper.] In the physics community [xxx see Frankel and Spivak 2?], the transformation rules are taken to *be* the defining characteristics.
- The **chain rule** and the **Jacobian matrix**. These are simply all over the place in this subject. Likewise, **Riemann sums**. (It is ironic that in our first-year analysis class, we learn the beautiful subject of measure theory; Riemann integration becomes a quaint relic. Yet in the geometry course, where everything is smooth or at worst piecewise smooth, the latter works just fine.)

[2] My approach here is reminiscent of the Rule of Four as used in [**HHGM**], although at this point it sounds like merely a Rule of Two. Namely, the rule is that mathematical ideas should be presented symbolically, graphically, numerically, and verbally. See section 12.3 for information about numerics; however, I have much more work yet to do in this area. (See also [**Kerl**].) Lastly, the very length of this paper constitutes a non-empty verbal-description component.

Acknowledgements

Doug Pickrell and Bill Faris provided numerous insights over the course of several discussions. Several problem solutions are due in full or in part to folks in the Monday problem sessions — Andy Lebovitz in particular — as indicated in the problems section of this paper. Also, some ideas are due to the folks in the summer 2007 qualifier-preparation session: in particular, Victor Piercey and Dave Herzog. Donna Harris proofread an early draft. Dan Champion and Gabriela Dumitrascu provided useful feedback. Thanks!

For constructive criticism and/or feedback please contact `kerl at math dot arizona dot edu`.

1 Preliminaries from analytic geometry

To guess what to keep and what to throw away takes considerable skill. Actually it is probably merely a matter of luck, but it looks as if it takes considerable skill. — Richard Feynman (1918-1988).

1.1 Notation

Notation 1.1. Although I may occasionally slip, I endeavor to always write elements of $\mathbb{R}^n$ as **column vectors**. Also, I usually boldface their names. For example,

$$\mathbf{q} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

in $\mathbb{R}^3$. The **transpose** of a column vector is a **row vector**, written

$$\mathbf{q}^t = \begin{pmatrix} x, & y, & z \end{pmatrix},$$

and vice versa.

Convention 1.2. I will identify row vectors with $1 \times n$ matrices and column vectors with $n \times 1$ matrices, respectively.

Notation 1.3. When a row vector $\mathbf{u}^t$ is being thought of as a $1 \times n$ matrix, I will write it as $\mathbf{u}^*$. The reason for this will be seen in section 4.6.7.

Notation 1.4. Throughout this paper, let

$$\hat{\mathbf{x}} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \hat{\mathbf{y}} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \text{and} \quad \hat{\mathbf{z}} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Note that some writers call these $\hat{\boldsymbol{\imath}}$, $\hat{\boldsymbol{\jmath}}$, and $\hat{\mathbf{k}}$. When working in arbitrary dimensions (rather than 2 or 3), I will call them $\hat{\mathbf{e}}_1$ through $\hat{\mathbf{e}}_n$, where $\hat{\mathbf{e}}_j$ has a 1 in the jth slot and zeroes elsewhere.

1.2 Trigonometric functions

xxx incl the wtrig diagram. label and explain.

xxx the necessary trig and hyp-trig identities: pyth, derivs, sum and difference.

$$\sin^2(\alpha) + \cos^2(\alpha) = 1$$
$$\sin(\alpha \pm \beta) = \sin(\alpha)\cos\beta) \pm \cos(\alpha)\sin(\beta) \qquad \cos(\alpha \pm \beta) = \cos(\alpha)\cos\beta) \mp \sin(\alpha)\sin(\beta).$$

Def'ns in terms of exponentials, which make anything possible (esp. choke recovery). [xxx perhaps a section/note on choke recovery ... it does very much matter.] What else? Check against the qual packet.

The key to the castle for spherical and hyperbolic trig functions is the following foursome:

$$\cos(x) = \frac{e^{ix} + e^{-ix}}{2} \qquad \cosh(x) = \frac{e^{x} + e^{-x}}{2}$$
$$\sin(x) = \frac{e^{ix} - e^{-ix}}{2i} \qquad \sinh(x) = \frac{e^{x} - e^{-x}}{2}.$$

xxx also inverse maps.

xxx mention graphs of cosh and sinh.

1.3 The dot product and its applications

Everything in this section is familiar from undergraduate math. In this section my goal is to connect these familiar terms with the new material to come.

1.3.1 The dot product and projected length

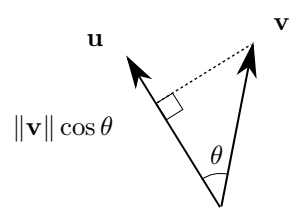
We are familar with the **dot product** of two vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$. *Algebraically*, we write

$$\mathbf{u} \cdot \mathbf{v} = \sum_{i=1}^{n} u_i v_i.$$

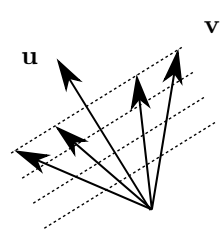
What, though, is the *geometric* interpretation[3]? It can be shown, using the law of cosines and the Pythagorean theorem, that

$$\mathbf{u} \cdot \mathbf{v} = \|\mathbf{u}\| \, \|\mathbf{v}\| \, \cos(\theta)$$

where θ is the angle between the two vectors, with θ taken to be from 0 to π. Now, the dot product is symmetric in $\mathbf{u}$ and $\mathbf{v}$. But suppose that we distinguish, or fix, one vector — say, $\mathbf{u}$ — and let the other one vary. Then $\mathbf{u} \cdot \mathbf{v}$ is the length of the **projection** of $\mathbf{v}$ onto $\mathbf{u}$, times the length of $\mathbf{u}$ itself:



We can think of $\mathbf{u}$ as *measuring length in a specified direction*. [xxx xref fwd to fishbone plots in linear-algebra and form sections.] Holding $\mathbf{u}$ fixed and varying $\mathbf{v}$, we will get different numbers. We can plot vectors $\mathbf{v}$ which all have the same projected length onto $\mathbf{u}$:



[xxx xref fwd to contour plots] Note in particular that if $\mathbf{v}$ is **perpendicular** to $\mathbf{u}$, then $\mathbf{u} \cdot \mathbf{v}$ is zero. [xref to proj/perp; perp space; normal/tangent.] In fact we can *define* perpendicularity: $\mathbf{u}$ is **perpendicular** to $\mathbf{v}$ if $\mathbf{u} \cdot \mathbf{v}$ is zero. (This means that the zero vector is perpendicular to every vector.)

Observe that, using convention 1.2, the dot product $\mathbf{u} \cdot \mathbf{v}$ is the same as the matrix product $\mathbf{u}^t \mathbf{v}$. For example,

[3] See also the nice Java applet at `www.falstad.com/dotproduct`.

in 3 dimensions:

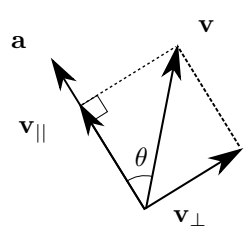
$$\begin{aligned}
\mathbf{u}\cdot\mathbf{v} &= \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = u_1v_1 + u_2v_2 + u_3v_3 \\
\mathbf{u}^t\cdot\mathbf{v} &= \begin{pmatrix} u_1 & u_2 & u_3 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = u_1v_1 + u_2v_2 + u_3v_3.
\end{aligned}$$

This line of thought is continued in section 4.6.4, where we think of $\mathbf{u}^t$ as being a function on vectors $\mathbf{v}$.

Last, note the following: the dot product is a function on two vectors $\mathbf{u}$ and $\mathbf{v}$ which is symmetric with respect to $\mathbf{u}$ and $\mathbf{v}$. That is, $\mathbf{u}\cdot\mathbf{v} = \mathbf{v}\cdot\mathbf{u}$. This is the paradigm whose generalization is the **symmetric tensor**, discussed in more detail in section 4.7.11.

1.3.2 Projection and perpendicular operators

Let $\mathbf{a}$ be a fixed vector in $\mathbb{R}^m$. Often in what follows we will want to **decompose** a vector $\mathbf{v}$ into two **components**, one parallel to $\mathbf{a}$ and one perpendicular to $\mathbf{a}$:



That is, we want to write

$$\mathbf{v} = \mathbf{v}_{\|} + \mathbf{v}_{\perp}$$

where $\mathbf{v}_{\|}$ is the component of $\mathbf{v}$ in the direction of $\mathbf{a}$ and $\mathbf{v}_{\perp}$ is the component of $\mathbf{v}$ perpendicular to $\mathbf{a}$, but still in the plane spanned by $\mathbf{v}$ and $\mathbf{a}$. (Note that if $\mathbf{v}$ is already parallel to $\mathbf{a}$, then $\mathbf{v}$ and $\mathbf{a}$ don't span a plane.) Since

$$\mathbf{v}_{\perp} = \mathbf{v} - \mathbf{v}_{\|},$$

it suffices to find the parallel component.

Given $\mathbf{v}$ and $\mathbf{a}$, we need to construct a vector with magnitude equal to $\|\mathbf{v}\|\cos\theta$ and direction $\mathbf{a}$, where θ is the angle between $\mathbf{v}$ and $\mathbf{a}$. Recall that the unit vector in the direction of $\mathbf{a}$ is

$$\hat{\mathbf{a}} = \frac{\mathbf{a}}{\|\mathbf{a}\|}.$$

So, we want

$$\mathbf{v}_{\|} = \|\mathbf{v}\|\cos\theta\,\frac{\mathbf{a}}{\|\mathbf{a}\|}.$$

Since

$$\mathbf{v}\cdot\mathbf{a} = \|\mathbf{v}\|\,\|\mathbf{a}\|\cos\theta,$$

and

$$\|\mathbf{a}\| = \sqrt{\mathbf{a}\cdot\mathbf{a}},$$

we have

$$\mathbf{v}_{\parallel} = \frac{\mathbf{v}\cdot\mathbf{a}}{\mathbf{a}\cdot\mathbf{a}}\,\mathbf{a} \qquad \text{and} \qquad \mathbf{v}_{\perp} = \mathbf{v} - \frac{\mathbf{v}\cdot\mathbf{a}}{\mathbf{a}\cdot\mathbf{a}}\,\mathbf{a}. \tag{1.1}$$

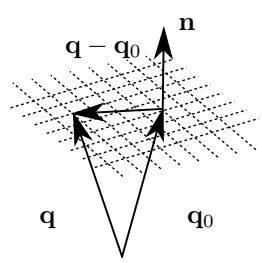
This looks familiar from the Gram-Schmidt process.

This technique will be indispensable when we decompose vectors into components perpendicular and parallel to the normal of a surface, namely, when working with tangent spaces. See for example section 10.7.2.

[xxx xref fwd to divergence/curl/etc.]

1.3.3 Point-normal form

A (hyper)plane is defined by a point $\mathbf{q}_0$ on the plane, and a vector $\mathbf{n}$ normal to the plane:



To get an equation for the plane, let $\mathbf{q}$ be an arbitrary point on the plane. Then $\mathbf{q}-\mathbf{q}_0$ is perpendicular to $\mathbf{n}$. Algebraically,

$$(\mathbf{q}-\mathbf{q}_0)\cdot\mathbf{n} = 0. \tag{1.2}$$

Suppose (in three dimensions) we have

$$\mathbf{q}_0 = \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix}, \mathbf{q} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \quad \text{and} \quad \mathbf{n} = \begin{pmatrix} a \\ b \\ c \end{pmatrix}.$$

Then equation 1.2 becomes

$$\begin{aligned} \begin{pmatrix} x-x_0 \\ y-y_0 \\ z-z_0 \end{pmatrix} \cdot \begin{pmatrix} a \\ b \\ c \end{pmatrix} &= 0 \\ a(x-x_0)+b(y-y_0)+c(z-z_0) &= 0. \end{aligned}$$

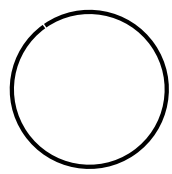
xxx Include the many fwd xrefs.

1.4 Coordinates

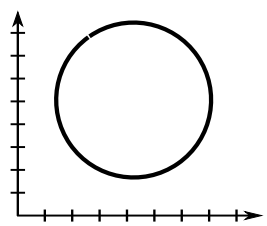
I have an existential map. It has "You are here" written all over it.
— Steven Wright (1955 -).

1.4.1 Coordinates, area, and the determinant

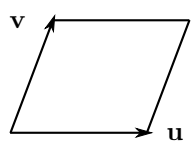
Here is a circle:



How big is it? What is its area? What is its curvature? These things require **rulers**[4]:



I have two points to make here: (1) what is forced upon us when we consider reasonable notions of symmetry and scaling for area, and (2) what happens when we change coordinates. I will address the former point in this section and the latter point in the next section. I'll switch from a circle to a square, or really any parallelogram, since these are easier to measure:



(Parallelograms and parallelepipeds are particularly important since they are regions swept out by systems of not-necessarily-orthonormal basis vectors. [xxx xref fwd.] Also, they are key for computing Riemann sums [xxx xref fwd].)

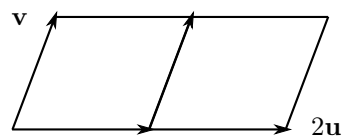
I'd like to think of the **area** of the parallelogram as depending on the lengths of two adjacent sides $\mathbf{u}$ and $\mathbf{v}$, and the angle θ between them. In particular I'd want the area A to be

$$A(\mathbf{u}, \mathbf{v}) = \|\mathbf{u}\| \, \|\mathbf{v}\| \, \sin(\theta).$$

(In a moment I'll inquire about negative θ.)

What properties do we want area to have? Well, it must be **additive**. Two identical parallelograms have twice the area of one of them:

[4]See [**BMA**] for the history of this key development, especially involving Descartes.



Algebraically,

$$A(2\mathbf{u}, \mathbf{v}) = 2A(\mathbf{u}, \mathbf{v}).$$

More generally, we expect

$$A(c\mathbf{u}, d\mathbf{v}) = cdA(\mathbf{u}, \mathbf{v})$$

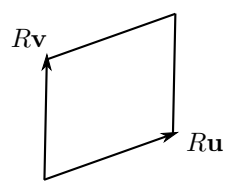
for any zero or positive c and d.

Next, the area of two unit squares should be one less than the area of three unit squares, and so on. From this we are forced to require

$$A(c\mathbf{u}, d\mathbf{v}) = cdA(\mathbf{u}, \mathbf{v})$$

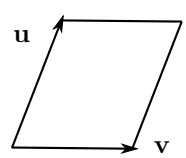
for any real c and d, whether positive, zero, or negative. (This is the **multilinearity** property of area, about which we will have much more to say in section 4.7.) That is, reasonable arithmetic *requires* that we allow the notion of **signed area** — even though I have no idea how much paint it takes to cover a wall whose area is -100 square feet.

Now what if I rotate the parallelogram?



Its area is unchanged; area should be **rotation invariant**. Algebraically, this is nothing new: θ is a *relative* angle betwen $\mathbf{u}$ and $\mathbf{v}$.

What if I pick up the parallelogram and flip it over?



We declare that we want the area to change sign on flips: it is **oriented**. Algebraically,

$$A(\mathbf{v}, \mathbf{u}) = -A(\mathbf{u}, \mathbf{v}).$$

This is the **alternating** or **skew symmetry** property of signed area; we will have more to say about this in section 4.7.11. This also makes sense of negative θ.

These properties, namely multilinearity and skew symmetry, are the defining characteristics of signed area. They are what constitute the determinant.

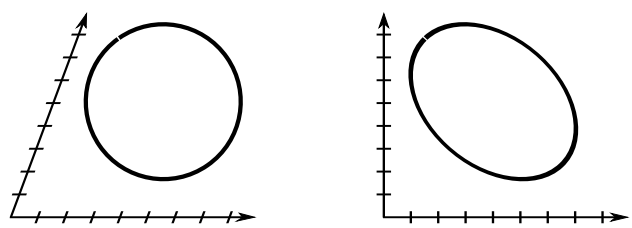
Definition 1.5. The **geometric determinant** of two vectors $\mathbf{u}$ and $\mathbf{v}$ (or three or more) is the oriented signed area of the parallelogram (or parallepiped) spanned by them.

Now, there is a **choice of orientation**: *after* I flip one edge, then the sign of the area is reversed. But what is the sign of the first area? It's a choice and there is a conventional way to decide it. [xxx Mention right-handed and left-handed coordinate systems.]

Showing that this geometric definition is equivalent to the algebraic notion of determinant [xxx below; perhaps type up handwritten notes for 2×2] is a separate task; the geometric notion suffices for now.

1.4.2 Change of coordinates

When we change coordinates, the object being measured is the same; only the underlying measurement system changes. But when we adopt a new measurement system, it becomes the new reality we perceive:



[xxx more here.] This topic will be addressed further in section 2.5.

1.4.3 Some coordinate systems

xxx to do: For $\mathbb{R}^2$ and $\mathbb{R}^3$; $\mathbb{S}^1$ and $\mathbb{S}^2$. Rectangular, polar, cylindrical, spherical with mnemonics. Conventions. Stereographic on both poles with reciprocal transition function.

1.4.4 Stereographic projection on $\mathbb{S}^1$

Questions involving stereographic coordinates are popular on the qualfiying exams. The algebra involved is not too profound, but can throw one for a loop if one has not worked through it at least once beforehand.

[xxx Draw the similar-triangles figures for $p = x/(1-y)$ and $q = x/(1+y)$.]

[xxx note the domains of definition for p and q, and their intersection.]

Proposition 1.6. *One converts from (x, y) coordinates on the circle to p or q coordinates by*

$$p = \frac{x}{1-y} \quad \textit{and} \quad q = \frac{x}{1+y}.$$

Proof. From the first figure, we can see that by similarity of triangles

$$\frac{p}{1} = \frac{x}{1-y}.$$

Likewise, from the second figure,

$$\frac{q}{1} = \frac{x}{1+y}.$$

□

Proposition 1.7. *The coordinates p and q are related by*

$$p = \frac{1}{q}.$$

Algebraic proof. Recalling that $x^2 + y^2 = 1$ on the circle, we have

$$\begin{aligned}
\frac{1}{q} &= \frac{1+y}{x} \\
&= \left(\frac{1+y}{x}\right)\left(\frac{1-y}{1-y}\right) \\
&= \frac{1-y^2}{x(1-y)} \\
&= \frac{x^2}{x(1-y)} \\
&= \frac{x}{(1-y)} \\
&= p.
\end{aligned}$$

□

Geometric proof. [xxx make figure; use similarity of triangles]. [xxx xref fwd to transition functions in the manifold section.] □

Proposition 1.8. *One converts from p or q coordinates back to (x, y) coordinates by*

$$x = \frac{2p}{p^2+1} \qquad \textit{and} \qquad y = \frac{p^2-1}{p^2+1};$$
$$x = \frac{2q}{1+q^2} \qquad \textit{and} \qquad y = \frac{1-q^2}{1+q^2}.$$

Proof. For the first, start with $p = x/(1-y)$. We can solve for y but we need to eliminate x. Our relation between x and y is $x^2 + y^2 = 1$ so we will need to square things. We have

$$p^2 = \frac{x^2}{(1-y)^2} = \frac{1-y^2}{(1-y)^2} = \frac{(1-y)(1+y)}{(1-y)(1-y)} = \frac{1+y}{1-y}.$$

This is an invertible rational function: in general, if $ad - bc \neq 0$ in $(ay+b)/(cy+d)$, then we can invert. Doing as we instructed our freshman students in college algebra, we get

$$\begin{aligned}
p^2 &= \frac{1+y}{1-y} \quad \text{and likewise} \quad q^2 = \frac{1-y}{1+y} \\
(1-y)p^2 &= 1+y \\
p^2 - p^2 y &= 1+y \\
(1+p^2)y &= p^2-1 \\
y &= \frac{p^2-1}{p^2+1}.
\end{aligned}$$

For x, use $p = x/(1-y)$ and solve for p:

$$x = p(1-y) = p\left(1 - \frac{p^2-1}{p^2+1}\right) = p\left(\frac{2}{p^2+1}\right) = \frac{2p}{p^2+1}.$$

Similar algebra proves the formulas for q. $\square$

xxx move to the manifold section. Establish and use the following to illustrate covariance and contravariance:

$$\begin{aligned}
\frac{dp}{dq} &= \frac{d}{dq}\left(\frac{1}{q}\right) = \frac{-1}{q^2} \\
\frac{dq}{dp} &= \frac{d}{dp}\left(\frac{1}{p}\right) = \frac{-1}{p^2}.
\end{aligned}$$

Then (noting that we differentiate $x(p)$ and $y(p)$):

$$\begin{aligned}
\frac{\partial}{\partial p} &= \begin{pmatrix} \partial x/\partial p \\ \partial y/\partial p \end{pmatrix} = \begin{pmatrix} \frac{\partial}{\partial p}\left(\frac{2p}{p^2+1}\right) \\ \frac{\partial}{\partial p}\left(\frac{p^2-1}{p^2+1}\right) \end{pmatrix} = \begin{pmatrix} \frac{2-2p^2}{(p^2+1)^2} \\ \frac{4p}{(p^2+1)^2} \end{pmatrix} \\
\frac{\partial}{\partial q} &= \begin{pmatrix} \partial x/\partial q \\ \partial y/\partial q \end{pmatrix} = \begin{pmatrix} \frac{\partial}{\partial q}\left(\frac{2q}{1+q^2}\right) \\ \frac{\partial}{\partial q}\left(\frac{1-q^2}{1+q^2}\right) \end{pmatrix} = \begin{pmatrix} \frac{2-2q^2}{(q^2+1)^2} \\ \frac{-4q}{(q^2+1)^2} \end{pmatrix}.
\end{aligned}$$

Then (noting that we differentiate the other way around, with $p(x)$):

$$\begin{aligned}
dp &= \frac{\partial}{\partial x}\left(\frac{x}{1-y}\right) dx + \frac{\partial}{\partial y}\left(\frac{x}{1-y}\right) dy \\
&= \frac{dx}{1-y} + \frac{x\,dy}{(1-y)^2} \\
dq &= \frac{\partial}{\partial x}\left(\frac{x}{1+y}\right) dx + \frac{\partial}{\partial y}\left(\frac{x}{1+y}\right) dy \\
&= \frac{dx}{1+y} - \frac{x\,dy}{(1+y)^2}.
\end{aligned}$$

[xxx describe what dx and dy mean here: $(y^2, -xy)$ and $(-xy, x^2)$, respectively. They are scalar multiples of $d\theta$ which is $(-y, x)$. And/or, use x or y as graph coordinates.]

[xxx important remark about how vector fields push forward but forms do not. xref to geometric-tangent-vector section. More work for forms. Edit and type up handwritten notes.]

Then [xxx do the algebra to show] (covariance and contravariance):

$$\begin{aligned}
dp &= \frac{dp}{dq}\,dq. \\
\frac{\partial}{\partial p} &= \frac{dq}{dp}\,\frac{\partial}{\partial q}.
\end{aligned}$$

1.4.5 Stereographic projection on $\mathbb{S}^n$

Here we generalize from section 1.4.4. The pattern becomes clear in $\mathbb{S}^3$.

For stereographic projection from the north pole which is $(x, y, z, w) = (0, 0, 0, 1)$, we want a coordinate

chart and a parameterization, [xxx xref fwd or reorg to charts/parameterizations] respectively:

$$\begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} \mapsto \begin{pmatrix} r \\ s \\ t \end{pmatrix} \qquad \text{and} \qquad \begin{pmatrix} r \\ s \\ t \end{pmatrix} \mapsto \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix}.$$

[xxx include figure and argument that this is done slotwise]

Then the desired coordinate chart is

$$r = \frac{x}{1-w}, \quad s = \frac{y}{1-w}, \quad t = \frac{z}{1-w}.$$

For the parameterization, to use the relation $x^2 + y^2 + z^2 + w^2 = 1$ we need to square things; adding them up to collect the numerators we have

$$r^2 + s^2 + t^2 = \frac{x^2 + y^2 + z^2}{(1-w)^2} = \frac{1 - w^2}{(1-w)^2} = \frac{1+w}{1-w}.$$

Now in general, proceeding as in section 1.4.4, to invert such a rational function we have

$$a = \frac{1+b}{1-b} \implies b = \frac{a-1}{a+1}.$$

Here,

$$w = \frac{r^2 + s^2 + t^2 - 1}{r^2 + s^2 + t^2 + 1}.$$

For x, use $x = r(1-w)$. In general

$$1 - \frac{c-1}{c+1} = \frac{2}{c+1}$$

so here

$$x = \frac{2r}{r^2 + s^2 + t^2 + 1}, \quad y = \frac{2s}{r^2 + s^2 + t^2 + 1}, \quad z = \frac{2t}{r^2 + s^2 + t^2 + 1}.$$

Generalizing to the n-sphere, one obtains the following coordinate chart and paramerization for stereographic projection from the north pole:

$$\begin{aligned} p_i &= \frac{x_i}{1 - x_{n+1}}, \ i = 1, \ldots, n; \\ x_i &= \frac{2p_i}{\sum_{j=1}^n p_j^2 + 1}, \ i = 1, \ldots, n, \qquad x_{n+1} = \frac{\sum_{j=1}^n p_j^2 - 1}{\sum_{j=1}^n p_j^2 + 1}. \end{aligned}$$

If one repeats the algebra for the south pole, one obtains

$$\begin{aligned} q_i &= \frac{x_i}{1 + x_{n+1}}, \ i = 1, \ldots, n; \\ x_i &= \frac{2q_i}{1 + \sum_{j=1}^n q_j^2}, \ i = 1, \ldots, n, \qquad x_{n+1} = \frac{1 - \sum_{j=1}^n q_j^2}{1 + \sum_{j=1}^n q_j^2}. \end{aligned}$$

It is evident from the figure [xxx] that the vectors $\mathbf{p} = (p_1, \ldots, p_n)$ and $\mathbf{q} = (q_1, \ldots, q_n)$ are collinear and on the same side of the origin. To see that $\|\mathbf{p}\| = 1/\|\mathbf{q}\|$, it is sufficient to see that $\|\mathbf{p}\|^2 = 1/\|\mathbf{q}\|^2$:

$$\mathbf{p} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \frac{1}{1 - x_{n+1}} \qquad \mathbf{q} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \frac{1}{1 + x_{n+1}}$$

$$\begin{aligned} \|\mathbf{p}\|^2 &= \frac{x_1^2 + \ldots + x_n^2}{(1 - x_{n+1})^2} \\ &= \frac{1 - x_{n+1}^2}{(1 - x_{n+1})^2} \\ &= \frac{1 + x_{n+1}}{1 - x_{n+1}} \end{aligned} \qquad \begin{aligned} \|\mathbf{q}\|^2 &= \frac{x_1^2 + \ldots + x_n^2}{(1 + x_{n+1})^2} \\ &= \frac{1 - x_{n+1}^2}{(1 + x_{n+1})^2} \\ &= \frac{1 - x_{n+1}}{1 + x_{n+1}}. \end{aligned}$$

[xxx xref fwd to transition functions in the manifold section.]

[xxx explicit transition functions. They work out quickly and nicely.]

1.5 Types of functions

We can have functions (linear or not) from $\mathbb{R}^m$ to $\mathbb{R}^n$. However, special cases and special terminology arise when $m = 1$, $n = 1$, or $m = n$. In this section I will discuss these types in the following order:

- paths from $\mathbb{R} \to \mathbb{R}^m$, then
- maps from $\mathbb{R}^m \to \mathbb{R}^n$, then
- scalar functions from $\mathbb{R}^n \to \mathbb{R}$.

This seemingly unimportant editorial decision reflects a natural order of composition:

- Given a path from $\mathbb{R} \to \mathbb{R}^m$, we can post-compose with a map from $\mathbb{R}^m \to \mathbb{R}^n$ to obtain another path.
- Given a scalar function from $\mathbb{R}^n \to \mathbb{R}$, we can pre-compose with a map from $\mathbb{R}^m \to \mathbb{R}^n$ to obtain another scalar function.

$$\begin{tikzcd}
 & \mathbb{R}^m \arrow[r, "\text{maps}"] \arrow[dr] & \mathbb{R}^n \arrow[dr, "\text{scalar functions}"] & \\
\mathbb{R} \arrow[ur, "\text{paths}"] \arrow[urr] & & & \mathbb{R}
\end{tikzcd}$$

This points toward pushforwards and pullbacks as discussed in [xxx xref]. In the differentiation section (section 2.2), though, I will discuss maps after functions. This is simply because maps are made up of functions.

xxx here or elsewhere: parameterized curves vs. graphs ... for VC we prefer the latter and think of the former as a special case; here, the other way around. elaborate.

1.5.1 Paths: $\mathbb{R}$ to $\mathbb{R}^m$

Reasoning draws a conclusion, but does not make the conclusion certain, unless the mind discovers it by the path of experience. — Roger Bacon (c. 1214-1294).

Definition 1.9. A **path**, or **scalar-to-vector function**, is a function $\gamma : \mathbb{R} \to \mathbb{R}^m$, which I will write for $m = 3$ as

$$\gamma(t) \quad \text{or} \quad \begin{pmatrix} \kappa(t) \\ \lambda(t) \\ \mu(t) \end{pmatrix}.$$

A scalar-to-vector function has m single-variable **component functions**.

Paths will lead us to **tangent vectors** [xxx xref fwd] and [xxx xref to topology, homotopy classes, first homology classes, etc.].

Visualizing paths is easy. xxx include a figure with a starting point and an arrow.

Example 1.10. ▷ Let

$$\gamma(t) = \begin{pmatrix} \cos(t) \\ \sin(t) \end{pmatrix}.$$

[xxx matlab figure.] ◁

1.5.2 Maps: $\mathbb{R}^m$ to $\mathbb{R}^n$

Definition 1.11. A **map**, or **vector-to-vector function**, is a function $\mathbf{F} : \mathbb{R}^m \to \mathbb{R}^n$. For $m = n = 3$, I will often write such a function as

$$\mathbf{F}(\mathbf{q}) \quad \text{or} \quad \begin{pmatrix} f(x,y,z) \\ g(x,y,z) \\ h(x,y,z) \end{pmatrix}.$$

A vector-to-vector function has n **component functions** of m variables. These component functions are vector-to-scalar functions (section 1.5.3).

Vector-to-vector functions will appear as **diffeomorphisms** [xxx xref fwd], in [xxx xref fwd] **change of variables**, and **vector fields** and **flows** [xxx xref fwd].

Maps are a bit harder to visualize than paths. There are at least two ways:

- **Grid images**. We think of each point $\mathbf{q}$ of $\mathbb{R}^m$ as being **carried** by $\mathbf{F}$ to a point $\mathbf{F}(\mathbf{q})$ of $\mathbb{R}^n$.
- **Quiver plots**. We think of $\mathbf{F}$ as **attaching** to each point $\mathbf{q}$ of $\mathbb{R}^m$ a vector $\mathbf{F}(\mathbf{q})$ which is **footed** not at the origin but at $\mathbf{q}$.

Example 1.12. ▷ Let $\mathbf{F} : \mathbb{R}^2 \to \mathbb{R}^2$ be given by

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x - y^3 \\ x^3 + 2y \end{pmatrix}.$$

A grid-image plot of $\mathbf{F}$ is as follows: [xxx matlab figures]. The first figure is the original plane: the positive x-axis is marked in heavy red and the positive y-axis is marked in heavy blue. The second figure shows the image of the grid under $\mathbf{F}$. [xxx xref to matlab section: section 12.3.]

A quiver plot shows $\mathbf{F}$ as a vector field: [xxx matlab figure] ◁

Remark 1.13. Generally we write vector-to-vector functions in the form

$$\mathbf{x} = \mathbf{F}(\mathbf{s}) \qquad \text{or} \qquad \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} f(s,t,u) \\ g(s,t,u) \\ h(s,t,u) \end{pmatrix}$$

using separate letters for input, function name, and output. A **change of coordinates** is nothing more than a vector-to-vector function (which we generally want to be invertible), but we usually omit the function name. (I call these **nameless functions**, although that terminology is non-standard.) We have simply

$$\mathbf{x} = \mathbf{x}(\mathbf{s}) \qquad \text{or} \qquad \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x(s,t,u) \\ y(s,t,u) \\ z(s,t,u) \end{pmatrix}.$$

1.5.3 Scalar functions: $\mathbb{R}^n$ to $\mathbb{R}$

Definition 1.14. A **scalar function**, or **vector-to-scalar function**, is a function $G : \mathbb{R}^n \to \mathbb{R}$, written for $n = 3$ as

$$G(x, y, z) \quad \text{or} \quad G(\mathbf{q}).$$

Scalar functions will appear as [xxx write and xref fwd] [xxx submanifolds] [xxx xref to regular-value theorem] [xxx coefficients in forms].

Scalar functions may be visualized in at least two ways:

- xxx **surfaces** [xxx figure here]
- xxx **contours** [xxx figure here]

Example 1.15. $\triangleright$ The **height function**

$$h(x, y, z) = z$$

or

$$h(x, y) = y.$$

[xxx elaborate.] [xxx figures.] $\triangleleft$

Example 1.16. $\triangleright$ $G(x, y) = x^2 + y^2$. [xxx elaborate] $\triangleleft$

[xref fwd (not far!) to level sets]

1.5.4 Level sets

Definition 1.17. Let $f : \mathbb{R}^3 \to \mathbb{R}$ be a vector-to-scalar function. The **level set** of f and c is the set of points

$$\{\mathbf{q} : f(\mathbf{q}) = c\}.$$

Example 1.18. $\triangleright$ Let $f(x, y) = x^2 + y^2$. It requires 3 dimensions to visualize this (non-flat) 2D surface, with $z = f(x, y)$:

[xxx figure]

It requires 2 dimensions to visualize f as a contour plot:

[xxx figure]

Now take the level set of f and 1. This result is a circle, which is a (non-flat) 1D object in 2D space:

[xxx figure]

A single level set *selects out* a one-dimensional space. We will see more of this in theorem 6.17. In short, $f(x, y) = c$ is one equation in two unknowns; there are two variables with one constraint and thus one degree of freedom. $\triangleleft$

xxx xref to tangent/normal dichotomy section.

1.5.5 Images of paths under maps: pushforwards

Let $\gamma : \mathbb{R} \to \mathbb{R}^m$ and $\mathbf{F} : \mathbb{R}^m \to \mathbb{R}^n$. Then $\mathbf{F} \circ \gamma$ is a path from $\mathbb{R}$ to $\mathbb{R}^n$. The path γ has been **pushed forward** from $\mathbb{R}^m$ to $\mathbb{R}^n$.

$$\begin{tikzcd}
 & \mathbb{R} \arrow[dl, "\gamma"'] \arrow[dr, "\mathbf{F} \circ \gamma"] & \\
\mathbb{R}^m \arrow[rr, "\mathbf{F}"'] & & \mathbb{R}^n
\end{tikzcd}$$

Example 1.19. ▷ Combining examples 1.12 and 1.10, let

$$\mathbf{F}(x, y) = \begin{pmatrix} x - y^3 \\ x^3 + 2y \end{pmatrix}.$$

and

$$\gamma(t) = \begin{pmatrix} \cos(t) \\ \sin(t) \end{pmatrix}.$$

[xxx matlab plots]. ◁

[xxx xref fwd to functors]

[xxx xref fwd to pushforward of tangent vectors]

[xxx xref fwd to DF and F_* section; 1st xref to v.c. pushf/deriv section.]

1.5.6 Images of maps under functions: pullbacks

Let $G : \mathbb{R}^n \to \mathbb{R}$ and $\mathbf{F} : \mathbb{R}^m \to \mathbb{R}^n$. Then $G \circ \mathbf{F}$ is a function from $\mathbb{R}^m$ to $\mathbb{R}$. The function G has been **pulled back** from $\mathbb{R}^n$ to $\mathbb{R}^m$.

$$\begin{tikzcd}
\mathbb{R}^m \arrow[rr, "\mathbf{F}"] \arrow[dr, "G \circ \mathbf{F}"'] & & \mathbb{R}^n \arrow[dl, "G"] \\
 & \mathbb{R} &
\end{tikzcd}$$

Example 1.20. ▷ TBD. Make it punchy; make sure it connects to other example(s) later. ◁

[xxx xref fwd to functors]

[xxx xref fwd to pullback of forms]

1.6 A gallery of curves and surfaces, part 1

xxx:

- $\mathbb{S}^1$ $(x^2 + y^2 = 1)$

- $\mathbb{S}^2$ $(x^2 + y^2 + z^2 = 1)$
- $\mathbb{T}^2$ in coordinates
- Paraboloids
- Hyperboloid of one sheet $(x^2 + y^2 - z^2 = 1)$
- Hyperboloid of two sheets $(z^2 - x^2 - y^2 = 1)$
- Surface of revolution for $g(y)$
- ...search 534 HW and quals for others.

2 Preliminaries from vector calculus

Hobbits delight in such things, if they are accurate: they like to have books filled with things that they already know, set out fair and square with no contradictions. — J.R.R. Tolkien, *The Fellowship of the Ring.*

Here we review notation which is old-fashioned in two senses: (1) it is 19th-century notation, and (2) it is what we were all taught as undergraduates. Our aim in this course is to replace, generalize, and extend such things, while still having the symbols *look* largely the same. Thus, it is worth taking the time to review what it is that we want to replace.

Also, I have found that much of the supposedly new material in the geometry-topology course is really vector calculus, being viewed in a new light. Viewing new things in new ways requires the learner to make a larger mental leap, so to ease that leap I will discuss old things in new ways for a while.

2.1 Goals for vector calculus

xxx list out classes of problems solved. xxx note that many of these arise in physics.

- Rate of change of functions.
- Linear approximation of functions.
- Optimization.
- Volume, mass, center of mass.
- Path and surface integrals: used for computing **work**, etc.
- Find the evolution of a system given specified initial/boundary conditions: ordinary differential equations, partial differential equations. In particular, configuration space of mechanical systems. (E.g. one-holed torus vs. double pendulum.)
- Understand how results change when different coordinate systems are introduced.

2.2 Differentiation

Certainly he who can digest a second or third fluxion need not, methinks, be squeamish about any point in divinity. — George Berkeley, 1734.

We will see that almost everything in this differentiation section involves the **chain rule**, which in turn is encapsulated in the **Jacobian matrix**. The Jacobian matrix will continue to reappear [xxx xref to COV for integrals, and elsewhere].

2.2.1 Derivatives of functions from $\mathbb{R}$ to $\mathbb{R}$

Recall from single-variable calculus that we use the derivative function, evaluated at a point x_0, to form a linear approximation to f at x_0. This is nothing more than **point-slope** form:

$$\begin{aligned} y &= y_0 + m(x - x_0) \\ f(x) &\approx f(x_0) + f'(x_0)(x - x_0). \end{aligned}$$

Note that the derivative itself is not a linear transformation. It is a function; evaluated at a point, it gives a slope. Linear transformations send zero to zero and this approximation need not: consider $f(x) = \sin(x)$ at $x_0 = 2\pi$. Such approximations are *affine*: a linear transformation with translation. But if we rearrange terms, then we do have a linear transformation:

$$f(x) - f(x_0) \approx f'(x_0)(x - x_0)$$

i.e.

$$\Delta f \approx f'(x_0)\Delta x.$$

[xxx two figures here, the latter with translated origin.]

[xxx more about coord charts don't have the origin at each point but the tangent spaces do. xref to position/velocity.]

Definition 2.1. Let $f : \mathbb{R} \to \mathbb{R}$. We say that x is a **critical point** of f if $f'(x)$ is zero or undefined. If x is a critical point for f, we say that $y = f(x)$ is a **critical value** for f.

Example 2.2. $\triangleright$ Let $f(x) = x^2 + 2x + 3$. We compute $f'(x) = 2x + 2$. This is a polynomial and so is not undefined anywhere; it is zero at $x = -1$. Therefore -1 is the only critical point for f. The critical value is $f(-1) = 2$. $\triangleleft$

See definition 6.16 in section 6.1.5 for the more general definition.

Remark 2.3. Remember that relative extrema happen at critical points, but not necessarily vice versa. The canonical examples are $f(x) = x^2$ and $g(x) = x^3$. Both have critical points at $x = 0$; f has a minimum there but g does not.

2.2.2 Derivatives of paths

Paths are simply differentiated componentwise.

Definition 2.4. Let

$$\gamma(t) = \begin{pmatrix} \kappa(t) \\ \lambda(t) \\ \mu(t)) \end{pmatrix}$$

be a path. Then

$$\gamma'(t) = \begin{pmatrix} \kappa'(t) \\ \lambda'(t) \\ \mu'(t)) \end{pmatrix}.$$

Notation 2.5. Looking ahead to section 2.2.4, I will write

$$\gamma' = D\gamma.$$

Note that γ is from $\mathbb{R}^1 \to \mathbb{R}^m$ and $D\gamma$ is an $m \times 1$ matrix.

[xxx figure]

We form tangent-line approximations for paths, much as in the single-variable case as described in section 2.2.1:

$$\gamma(t) \approx \gamma(t_0) + \gamma'(t_0)(t - t_0).$$

Example 2.6. $\triangleright$ Continuing example 1.10, let

$$\gamma(t) = \begin{pmatrix} \cos(t) \\ \sin(t) \end{pmatrix}.$$

Then

$$\gamma'(t) = \begin{pmatrix} -\sin(t) \\ \cos(t) \end{pmatrix}.$$

At $t_0 = \pi/2$, we have an equation for the tangent line $\ell(t)$ given by

$$\begin{aligned} \gamma(t) &\approx \gamma(t_0) + \gamma'(t_0)(t - t_0) \\ &= \begin{pmatrix} \cos(\pi/2) \\ \sin(\pi/2) \end{pmatrix} + \begin{pmatrix} -\sin(\pi/2) \\ \cos(\pi/2) \end{pmatrix} (t - \pi/2) \\ &= \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \begin{pmatrix} -1 \\ 0 \end{pmatrix} (t - \pi/2) \\ &= \begin{pmatrix} \pi/2 - t \\ 1 \end{pmatrix}. \end{aligned}$$

[xxx matlab figure.] $\triangleleft$

If a path $\gamma(t)$ represents **position** of an object, then the derivative $\gamma'(t)$ represents its **velocity**. When the velocity is multiplied by the object's mass, we have **momentum** (physicists write $\mathbf{p} = m\mathbf{v}$). [xxx elaborate: a physicist might in fact think of a tangent bundle [xxx xref fwd] as a space of positions and momenta.]

[xxx mnem: gamma as car driving on curve; gamma prime as the headlight beam.]

[xxx note: we lose the linear-transformation part of it when we put it to work for the tangent-line approximation. The match is at $t = t_0$ not $t = 0$.]

See definition 6.16 in section 6.1.5 for the definition of critical point. For paths from $\mathbb{R}$ into higher-dimensional spaces, critical points are not interesting. (The derivative can never be surjective.)

2.2.3 Derivatives of functions from $\mathbb{R}^n$ to $\mathbb{R}$

Notation 2.7. For single-variable functions we write the derivative as

$$\frac{dG}{dx} \qquad \text{or} \qquad G'(x);$$

for multi-variable functions we write

$$\frac{\partial G}{\partial x} \qquad \text{or} \qquad \frac{\partial G}{\partial x}(x,y,z) \qquad \text{or} \qquad G_x(x,y,z).$$

When we evaluate at a point, we write

$$\left.\frac{dG}{dx}\right|_{x_0} \qquad \text{or} \qquad \frac{dG}{dx}(x_0) \qquad \text{or} \qquad G'(x_0)$$

in the single-variable case and

$$\left.\frac{\partial G}{\partial x}\right|_{\left(\begin{smallmatrix} x_0 \\ y_0 \\ z_0 \end{smallmatrix}\right)} \qquad \text{or} \qquad \frac{\partial G}{\partial x}(x_0,y_0,z_0) \qquad \text{or} \qquad G_x(x_0,y_0,z_0)$$

in the multivariable case.

Intuition 2.8. We can use partial derivatives to form a **linear approximation** to a scalar function:

$$G(x,y) \approx G(x_0,y_0) + G_x(x_0,y_0)(x-x_0) + G_y(x_0,y_0)(y-y_0).$$

The best way I know to motivate this is by using single-variable linear approximations, one variable at at time. Starting at (x_0,y_0) we can make a line to (x,y_0) and from there to (x,y):

$$\begin{aligned} G(x,y_0) &\approx G(x_0,y_0) + G_x(x_0,y_0)(x-x_0) \\ G(x,y) &\approx G(x,y_0) + G_y(x,y_0)(y-y_0) \\ &\approx G(x_0,y_0) + G_x(x_0,y_0)(x-x_0) + G_y(x,y_0)(y-y_0) \end{aligned}$$

Alternatively, starting at (x_0,y_0) we can make a line to (x_0,y) and from there to (x,y):

$$\begin{aligned} G(x_0,y) &\approx G(x_0,y_0) + G_y(x_0,y_0)(y-y_0) \\ G(x,y) &\approx G(x_0,y) + G_x(x_0,y)(x-x_0) \\ &\approx G(x_0,y_0) + G_y(x_0,y_0)(y-y_0) + G_x(x_0,y)(x-x_0) \\ &= G(x_0,y_0) + G_x(x_0,y)(x-x_0) + G_y(x_0,y_0)(y-y_0). \end{aligned}$$

Comparing these two expressions, we see that the difference is

$$G_x(x_0,y_0)(x-x_0) + G_y(x,y_0)(y-y_0)$$

vs.

$$G_x(x_0,y)(x-x_0) + G_y(x_0,y_0)(y-y_0).$$

If G is linear, then

$$G_x(x_0,y_0) = G_x(x_0,y) \qquad \text{and} \qquad G_y(x_0,y_0) = G_y(x,y_0);$$

if G is approximately linear then we can replace one with the other to obtain the desired result

$$\begin{aligned} G(x,y) &\approx G(x_0,y_0) + G_x(x_0,y_0)(x-x_0) + G_y(x_0,y_0)(y-y_0) \\ &= G(x_0,y_0) + G_x(x_0,y_0)\Delta x + G_y(x_0,y_0)\Delta y. \end{aligned}$$

A careful use of the word "approximately" requires a limit proof; see [**Rudin**], [**Spivak1**], or the differentiation section of any graduate analysis text.

xxx figure

xxx now note it looks like a plane form.

Just as in section 2.2.1, this approximation isn't a linear transformation — it only becomes one when we rearrange terms:

$$\begin{aligned} f(x,y) - f(x_0,y_0) &\approx f_x(x_0,y_0)\Delta x + f_y(x_0,y_0)\Delta y \\ \Delta f &\approx \begin{pmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \end{pmatrix}\Bigg|_{\left(\begin{smallmatrix} x_0 \\ y_0 \end{smallmatrix}\right)} \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix}. \end{aligned}$$

We can think of this product in two ways:

- As the dot product of two column vectors;
- As the matrix product of a row vector by a column vector.

[xxx foreshadow ominously that this seemingly minor distinction leads into covariance and contravariance. numerous xrefs.]

Definition 2.9. Let $G : \mathbb{R}^n \to \mathbb{R}$. Define the **column vector**

$$\nabla G = \begin{pmatrix} \frac{\partial G}{\partial x_1}(x_1,\ldots,x_n) \\ \vdots \\ \frac{\partial G}{\partial x_n}(x_1,\ldots,x_n) \end{pmatrix}.$$

(This inverted triangle is pronounced "grad" or "nabla". [xxx Anton ref w/ page number for the etymology.]) In particular, for $n = 3$ we have

$$\nabla G = \begin{pmatrix} \partial G/\partial x \\ \partial G/\partial y \\ \partial G/\partial z \end{pmatrix}.$$

Notice that we started with a vector-to-scalar function $G : \mathbb{R}^n \to \mathbb{R}$. Then we obtained a vector-to-vector function $\nabla G : \mathbb{R}^n \to \mathbb{R}^n$. This suggests that the ∇ may be thought of as a **differential operator** ∇ which takes vector-to-scalar functions into vector-to-vector functions [xxx xref to functors].

Definition 2.10. The **gradient** is the following **differential operator**, written as a **column vector**:

$$\nabla = \begin{pmatrix} \partial/\partial x_1 \\ \vdots \\ \partial/\partial x_n \end{pmatrix}.$$

For $n = 3$ we have

$$\nabla = \begin{pmatrix} \partial/\partial x \\ \partial/\partial y \\ \partial/\partial z \end{pmatrix}.$$

xxx figure: contour and quiver.

Remark 2.11. Since ∇ takes a scalar function G into a vector field function $\mathbf{F} = \nabla G$, it is natural to ask if we can go the other way. That is, given $\mathbf{F}$, can we find G such that $\nabla G = \mathbf{F}$? If so, we say that G is a **potential function** for $\mathbf{F}$. [xxx write and xref to where we can answer that question.] [xxx figures]

Notation 2.12. Since the gradient of a scalar function is a column vector, its transpose is a row vector. Looking ahead to section 2.2.4, and recalling section 2.2.2, I might write this as

$$\nabla^t G = DG = \begin{pmatrix} \frac{\partial G}{\partial x} & \frac{\partial G}{\partial y} & \frac{\partial G}{\partial z} \end{pmatrix}.$$

Note that G is from $\mathbb{R}^n \to \mathbb{R}^1$ and DG is a $1 \times n$ matrix. Then we have

$$\Delta f \approx DG \Delta \mathbf{x}.$$

xxx:

$$f(x,y,z) - f(x_0,y_0,z_0) \quad \approx \quad \begin{pmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} & \frac{\partial f}{\partial z} \end{pmatrix}\Bigg|_{\left(\begin{smallmatrix} x_0 \\ y_0 \\ z_0 \end{smallmatrix}\right)} \begin{pmatrix} \Delta x \\ \Delta y \\ \Delta z \end{pmatrix}.$$

[xxx dir'nal deriv, and grad as direction of greatest change? figures. this leads to normals.]

Definition 2.13. Let $G : \mathbb{R}^m \to \mathbb{R}$. We say that $\mathbf{q}$ is a **critical point** of G if $DG(\mathbf{q})$ is zero or undefined. If $\mathbf{q}$ is a critical point for G, we say that $y = G(\mathbf{q})$ is a **critical value** for G.

Remark 2.14. See definition 6.16 in section 6.1.5 for the more general definition: namely, for functions $\mathbf{F} : \mathbb{R}^m$ to $\mathbb{R}^n$, a point $\mathbf{q}$ is critical if $D\mathbf{F}(\mathbf{q})$ is not surjective. Here, where $n = 1$, DG is an $1 \times m$ matrix (which looks like a row vector). For a $1 \times m$ matrix not to be surjective means it must have rank less than 1. But its rank can be at most 1. For a $1 \times m$ matrix to have rank zero means it must *be* zero. This is the same as the gradient (remember $\nabla^t G = DG$) being zero.

Example 2.15. $\triangleright$ Let $G(x,y) = x^2 - y^2 + 3$. We compute $DG(x,y) = (2x, -2y)$. This has polynomial components and so is not undefined anywhere. It is zero only when $2x = -2y = 0$, i.e. at the single point $x = 0, y = 0$. The critical value is $G(0,0) = 3$. $\triangleleft$

Remark 2.16. Just as in the scalar case (see remark 2.3), relative extrema occur at critical points but not necessarily vice versa. For the example just given, the point $x = 0, y = 0$ is a critical point, but it is a saddle point rather than a relative minimum or relative maximum. (If we had defined $G(x,y) = x^2 + y^2 + 3$ we would have had a minimum at the origin.)

[xxx include figures here]

2.2.4 Derivatives of maps from $\mathbb{R}^m$ to $\mathbb{R}^n$

In section 1.5.2 I wrote a vector-to-vector function as

$$\mathbf{F}(x,y,z) = \begin{pmatrix} f(x,y,z) \\ g(x,y,z) \\ h(x,y,z) \end{pmatrix}.$$

That is, a vector-to-vector function is simply a *stack* of the component vector-to-scalar functions. To make a linear approximation for a vector-to-vector function, then, we can simply apply the gradient as defined in

section 2.2.3 to each component function:

$$\begin{aligned}
f(x,y,z) - f(x_0,y_0,z_0) &\approx \begin{pmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} & \frac{\partial f}{\partial z} \end{pmatrix}\Bigg|_{\left(\begin{smallmatrix} x_0 \\ y_0 \\ z_0 \end{smallmatrix}\right)} \begin{pmatrix} \Delta x \\ \Delta y \\ \Delta z \end{pmatrix} \\
g(x,y,z) - g(x_0,y_0,z_0) &\approx \begin{pmatrix} \frac{\partial g}{\partial x} & \frac{\partial g}{\partial y} & \frac{\partial g}{\partial z} \end{pmatrix}\Bigg|_{\left(\begin{smallmatrix} x_0 \\ y_0 \\ z_0 \end{smallmatrix}\right)} \begin{pmatrix} \Delta x \\ \Delta y \\ \Delta z \end{pmatrix} \\
h(x,y,z) - h(x_0,y_0,z_0) &\approx \begin{pmatrix} \frac{\partial h}{\partial x} & \frac{\partial h}{\partial y} & \frac{\partial h}{\partial z} \end{pmatrix}\Bigg|_{\left(\begin{smallmatrix} x_0 \\ y_0 \\ z_0 \end{smallmatrix}\right)} \begin{pmatrix} \Delta x \\ \Delta y \\ \Delta z \end{pmatrix}.
\end{aligned}$$

But this is the same as saying

$$\begin{pmatrix} f(x,y,z) - f(x_0,y_0,z_0) \\ g(x,y,z) - g(x_0,y_0,z_0) \\ h(x,y,z) - h(x_0,y_0,z_0) \end{pmatrix} \approx \begin{pmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} & \frac{\partial f}{\partial z} \\ \frac{\partial g}{\partial x} & \frac{\partial g}{\partial y} & \frac{\partial g}{\partial z} \\ \frac{\partial h}{\partial x} & \frac{\partial h}{\partial y} & \frac{\partial h}{\partial z} \end{pmatrix}\Bigg|_{\left(\begin{smallmatrix} x_0 \\ y_0 \\ z_0 \end{smallmatrix}\right)} \begin{pmatrix} \Delta x \\ \Delta y \\ \Delta z \end{pmatrix}.$$

Thus we have motivated the following definition.

Definition 2.17. Let $\mathbf{F} : \mathbb{R}^3 \to \mathbb{R}^3$ be

$$\mathbf{F}(x,y,z) = \begin{pmatrix} f(x,y,z) \\ g(x,y,z) \\ h(x,y,z) \end{pmatrix}.$$

Then we define the **Jacobian matrix** for $\mathbf{F}$ to be the matrix with rows equal to the gradients of the component functions:

$$D\mathbf{F} = \begin{pmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} & \frac{\partial f}{\partial z} \\ \frac{\partial g}{\partial x} & \frac{\partial g}{\partial y} & \frac{\partial g}{\partial z} \\ \frac{\partial h}{\partial x} & \frac{\partial h}{\partial y} & \frac{\partial h}{\partial z} \end{pmatrix}.$$

In general, for $\mathbf{F} : \mathbb{R}^m \to \mathbb{R}^m$ given by

$$\mathbf{F}(x_1,\ldots,x_m) = \begin{pmatrix} f_1(x_1,\ldots,x_m) \\ \vdots \\ f_m(x_1,\ldots,x_m) \end{pmatrix}$$

the Jacobian matrix has entries

$$\partial f_i / \partial x_j.$$

Notation 2.18. In case we have a nameless function $\mathbf{y} = \mathbf{y}(\mathbf{x})$, as in remark 1.13, it is common to write the Jacobian as either

$$\frac{\partial(y_1,\ldots,y_n)}{\partial(x_1,\ldots,x_m)}$$

or

$$\frac{\partial \mathbf{y}}{\partial \mathbf{x}}.$$

Remark 2.19. Note that the Jacobian matrix, as presented, is a **matrix-valued function**. Only when we evaluate it at a point, e.g.

$$D\mathbf{F}|_{\mathbf{q}},$$

does it become a matrix.

Example 2.20. ▷ Let

$$\mathbf{F}(x,y) = \begin{pmatrix} x - y^3 \\ x^3 + 2y \end{pmatrix}$$

as in example 1.12. Then

$$D\mathbf{F}(x,y) = \begin{pmatrix} 1 & -3y^2 \\ 3x^2 & 2 \end{pmatrix}.$$

Evaluated at $x = 0, y = 1$, this is

$$D\mathbf{F}(0,1) = \begin{pmatrix} 1 & -3 \\ 0 & 2 \end{pmatrix}.$$

We have

$$\mathbf{F}(0,1) = \begin{pmatrix} -1 \\ 2 \end{pmatrix}$$

so near $x_0 = 0$, $y_0 = 1$ we have

$$\begin{aligned} \mathbf{F}(x,y) &\approx \mathbf{F}(x_0,y_0) + D\mathbf{F}(x_0,y_0)\begin{pmatrix} x - x_0 \\ y - y_0 \end{pmatrix} \\ &\approx \begin{pmatrix} -1 \\ 2 \end{pmatrix} + \begin{pmatrix} 1 & -3 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} x - 0 \\ y - 1 \end{pmatrix} = \begin{pmatrix} x - 3y + 2 \\ 2y \end{pmatrix}. \end{aligned}$$

[xxx figures: superimpose eucl grid on grid image] ◁

Note that $\gamma' = D\gamma$ and $\nabla^t G = DG$ from notations 2.5 and 2.12 are both simply special cases of the Jacobian. [xxx xref fwd to DF section.]

See definition 6.16 in section 6.1.5 for the definition of critical point. The very definition requires linear algebra, which is why we don't teach it at the vector-calculus level.

2.2.5 Derivatives for images of paths under maps

Let

$$\gamma(t) = \begin{pmatrix} \kappa(t) \\ \lambda(t) \\ \mu(t) \end{pmatrix}$$

be a path and let

$$\mathbf{F}(x,y,z) = \begin{pmatrix} f(x,y,z) \\ g(x,y,z) \\ h(x,y,z) \end{pmatrix}.$$

What is the derivative of the composition $\mathbf{F} \circ \gamma$? Using the **chain rule**, we have

$$\begin{aligned} \frac{d}{dt}\mathbf{F}(\gamma(t)) &= \begin{pmatrix} \frac{d}{dt} f(\kappa(t),\lambda(t),\mu(t)) \\ \frac{d}{dt} g(\kappa(t),\lambda(t),\mu(t)) \\ \frac{d}{dt} h(\kappa(t),\lambda(t),\mu(t)) \end{pmatrix} \\ &= \begin{pmatrix} \frac{\partial f}{\partial x}(\kappa(t))\kappa'(t) + \frac{\partial f}{\partial y}(\lambda(t))\lambda'(t) + \frac{\partial f}{\partial z}(\mu(t))\mu'(t) \\ \frac{\partial g}{\partial x}(\kappa(t))\kappa'(t) + \frac{\partial g}{\partial y}(\lambda(t))\lambda'(t) + \frac{\partial g}{\partial z}(\mu(t))\mu'(t) \\ \frac{\partial h}{\partial x}(\kappa(t))\kappa'(t) + \frac{\partial h}{\partial y}(\lambda(t))\lambda'(t) + \frac{\partial h}{\partial z}(\mu(t))\mu'(t) \end{pmatrix} \\ &= \begin{pmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} & \frac{\partial f}{\partial z} \\ \frac{\partial g}{\partial x} & \frac{\partial g}{\partial y} & \frac{\partial g}{\partial z} \\ \frac{\partial h}{\partial x} & \frac{\partial h}{\partial y} & \frac{\partial h}{\partial z} \end{pmatrix}\Bigg|_{\left(\begin{smallmatrix} \kappa(t) \\ \lambda(t) \\ \mu(t) \end{smallmatrix}\right)} \begin{pmatrix} \kappa'(t) \\ \lambda'(t) \\ \mu'(t) \end{pmatrix}. \end{aligned}$$

That is to say,

$$\frac{d}{dt}(\mathbf{F}(\gamma(t)) = D\mathbf{F}(\gamma(t))\gamma'(t).$$

We can use this rule to produce an equation for the tangent line to the image of a path under a map.

Example 2.21. ▷ Continuing examples 1.10, 1.12, 1.19, 2.6, and 2.20, let

$$\mathbf{F}(x,y) = \begin{pmatrix} x - y^3 \\ x^3 + 2y \end{pmatrix},$$

$$\gamma(t) = \begin{pmatrix} \cos(t) \\ \sin(t) \end{pmatrix},$$

and $t_0 = \pi/2$. Then from example 2.6 we have an equation for the tangent line to the original path:

$$\begin{aligned} \gamma(t) &\approx \gamma(t_0) + \gamma'(t_0)(t - t_0) \\ &\approx \begin{pmatrix} \pi/2 - t \\ 1 \end{pmatrix}. \end{aligned}$$

Then

$$\begin{aligned} \mathbf{F}(\gamma(t)) &\approx \mathbf{F}(\gamma(t_0)) + D\mathbf{F}(\gamma(t_0))\gamma'(t_0)(t - t_0) \\ &= \mathbf{F}(0,1) + D\mathbf{F}(0,1)\begin{pmatrix} -1 \\ 0 \end{pmatrix}(t - \pi/2) \\ &= \begin{pmatrix} -1 \\ 2 \end{pmatrix} + \begin{pmatrix} 1 & -3 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} -1 \\ 0 \end{pmatrix}(t - \pi/2) \\ &= \begin{pmatrix} -1 \\ 2 \end{pmatrix} + \begin{pmatrix} -1 \\ 0 \end{pmatrix}(t - \pi/2) \\ &= \begin{pmatrix} \pi/2 - 1 - t \\ 2 \end{pmatrix}. \end{aligned}$$

[xxx matlab figures] ◁

Note:

- The not-necessarily-linear function $\mathbf{F}$ pushes forward the not-necessarily-linear path γ.
- The linear approximation $D\mathbf{F}$ pushes forward the tangent line defined by γ'.

[xxx xref fwd to pushforward of tangent vectors]

2.2.6 Derivatives for images of maps under functions

[xxx example]

[xxx xref fwd to pullback of forms]

2.2.7 The chain rule and the Jacobian

xxx write out:

$D((B \circ A)(\mathbf{q})) = DB(A(\mathbf{q}))DA(\mathbf{q})$.

xref back/fwd that prevs are special cases of this.

2.2.8 Divergence and curl

xxx intuit first. Follow [**HHGM**]

xxx the det formula seems quite arbitrary. xref to d section where it falls out naturally. also intuit ....

Definition 2.22. Let $\mathbf{F}$ be a vector-to-vector function. The **curl** of $\mathbf{F}$, written $\mathrm{curl}(\mathbf{F})$ or $\nabla \times \mathbf{F}$, is the cross product written as

$$\nabla \times \mathbf{F} = \begin{vmatrix} \hat{\mathbf{x}} & \hat{\mathbf{y}} & \hat{\mathbf{z}} \\ \partial/\partial x & \partial/\partial y & \partial/\partial z \\ f(x,y,z) & g(x,y,z) & h(x,y,z) \end{vmatrix} = \begin{pmatrix} \partial h/\partial y - \partial g/\partial z \\ \partial f/\partial z - \partial h/\partial x \\ \partial g/\partial x - \partial f/\partial y \end{pmatrix} = \begin{pmatrix} h_y - g_z \\ f_z - h_x \\ g_x - f_y \end{pmatrix}.$$

Note that the middle expression is just a silly mnemonic; it is not a determinant in any formal sense.

Definition 2.23. Let $\mathbf{F}$ be a vector-to-vector function. The **divergence** of $\mathbf{F}$, sometimes written $\mathrm{div}(\mathbf{F})$, is the dot product

$$\nabla \cdot \mathbf{F} = \begin{pmatrix} \partial/\partial x \\ \partial/\partial y \\ \partial/\partial z \end{pmatrix} \cdot \begin{pmatrix} f(x,y,z) \\ g(x,y,z) \\ h(x,y,z) \end{pmatrix} = \partial f/\partial x + \partial g/\partial y + \partial h/\partial z.$$

Note several things:

- The gradient takes vector-to-scalar functions to vector-to-vector functions.
- The curl takes vector-to-vector functions to vector-to-vector functions.
- The divergence takes vector-to-vector functions to vector-to-scalar functions.

Since the gradient of a vector-to-scalar function is a vector-to-vector function, it makes sense to take the curl of it; also it makes sense to take divergence of the curl of a vector-to-vector function.

Proposition 2.24. *As operators,* $\nabla \times \nabla = 0$ *and* $\nabla \cdot \nabla \times = 0$.

Proof. For the first, let G be a scalar function of three variables. Then

$$\begin{aligned} \nabla \times \nabla G &= \nabla \times \begin{pmatrix} \partial G/\partial x \\ \partial G/\partial y \\ \partial G/\partial z \end{pmatrix} = \begin{vmatrix} \hat{\mathbf{x}} & \hat{\mathbf{y}} & \hat{\mathbf{z}} \\ \partial/\partial x & \partial/\partial y & \partial/\partial z \\ \partial G/\partial x & \partial G/\partial y & \partial G/\partial z \end{vmatrix} \\ &= \begin{pmatrix} \partial^2 G/\partial y \partial z - \partial^2 G/\partial z \partial y \\ \partial^2 G/\partial z \partial x - \partial^2 G/\partial x \partial z \\ \partial^2 G/\partial x \partial y - \partial^2 G/\partial y \partial x \end{pmatrix} = 0. \end{aligned}$$

The last step is true because all our functions are assumed smooth: since the first and second partials are continuous, the mixed partials are equal.

For the second, let $\mathbf{F} : \mathbb{R}^3 \to \mathbb{R}^3$. Then

$$\begin{aligned}
\nabla \cdot \nabla \times \mathbf{F} &= \nabla \cdot \begin{pmatrix} \partial h/\partial y - \partial g/\partial z \\ \partial f/\partial z - \partial h/\partial x \\ \partial g/\partial x - \partial f/\partial y \end{pmatrix} \\
&= (\partial^2 h/\partial x \partial y - \partial^2 g/\partial x \partial z) \\
&+ (\partial^2 f/\partial y \partial z - \partial^2 h/\partial y \partial x) \\
&+ (\partial^2 g/\partial z \partial x - \partial^2 f/\partial z \partial y) \\
&= 0
\end{aligned}$$

again due to equality of mixed partials. □

Notes:

- Since $\nabla \times \mathbf{F}$ is vector-to-vector, it makes sense to consider $\nabla \times \nabla \times \mathbf{F}$. I am leaving this topic untouched. (See [**Hildebrand**], section 6.9.)
- What do we do in higher dimensions (e.g. phase spaces of complex mechanical systems)? The gradient and the divergence seem to make sense for higher dimensions, but there is no clear notion here of how to extend the concept of curl. [xxx xref to 2D and 4D $d^2 = 0$ complexes, when I write that. Or to the d section and let them do it themselves.]

2.3 A first pass at tangent spaces

[xxx s2tp]

circle. level set and not. two views of tangent line: kernel of linear transformation. show by example that we get the same thing either way; question is how.

same with the two-sphere. relate to plane forms ....

main point: normal is perp to tangent.

2.4 Integration

But just as much as it is easy to find the differential of a given quantity, so it is difficult to find the integral of a given differential. Moreover, sometimes we cannot say with certainty whether the integral of a given quantity can be found or not. — Johann Bernoulli (1667-1748).

2.4.1 Integrals of paths

If $\mathbf{v}(t)$ is an expression for the velocity of a particle, i.e. $\mathbf{v}(t) = \gamma'(t)$ for some path $\gamma(t)$, then we can recover $\gamma(t)$ via

$$\gamma(t) - \gamma(t_0) = \int_{t_0}^{t} \mathbf{v}(\tau) d\tau.$$

Definition 2.25. The integral of a path is done componentwise: if

$$\mathbf{v}(t) = \begin{pmatrix} u(t) \\ v(t) \\ w(t) \end{pmatrix}$$

then

$$\int_{t_0}^{t} \mathbf{v}(t) d\tau = \int_{t_0}^{t} \begin{pmatrix} u(t) \\ v(t) \\ w(t) \end{pmatrix} d\tau = \begin{pmatrix} \int_{t_0}^{t} u(t) d\tau \\ \int_{t_0}^{t} v(t) d\tau \\ \int_{t_0}^{t} w(t) d\tau \end{pmatrix}$$

Example 2.26. $\triangleright$ Let

$$\mathbf{v}(t) = \begin{pmatrix} -\sin(t) \\ \cos(t) \end{pmatrix}$$

and $t_0 = 0$. Then

$$\begin{aligned} \int_{t_0}^{t} \mathbf{v}(t) d\tau &= \int_{0}^{t} \begin{pmatrix} -\sin(t) \\ \cos(t) \end{pmatrix} d\tau = \begin{pmatrix} \int_0^t -\sin(t) d\tau \\ \int_0^t \cos(t) d\tau \end{pmatrix} \\ &= \begin{pmatrix} \cos(t) - 1 \\ \sin(t) \end{pmatrix} = \begin{pmatrix} \cos(t) \\ \sin(t) \end{pmatrix} - \begin{pmatrix} 1 \\ 0 \end{pmatrix} \end{aligned}$$

which is $\gamma(t) - \gamma(0)$ for the path of example 2.6. $\triangleleft$

There is not much more to say here that we haven't already seen as undergraduates; the graduate course is silent on this subject.

2.4.2 Integrals of scalar functions

The integral of a scalar function $G(x)$, $G(x, y)$, or $G(x, y, z)$ is the familiar area/volume under the curve: if the circus grounds are on a region R of the (x, y) plane and G is an expression for the height of the circus tent, then $\int_R G(x, y)\, dx\, dy$ is the volume enclosed by the tent.

[xxx figure]

Such integrals are easy to understand and (maybe) easy to compute. Thus it is perhaps no surprise that we have integration theorems which turn integrals of vector-to-vector functions into integrals of scalar functions. [xxx xref.]

2.4.3 Integrals of maps

Line integrals: integral of a vector field over a path ....

xxx

[xxx antideriv vs. solving differential equations?]

curl and div; xref fwd to the classical integration thms. Point out the interior/boundary pattern, including the FTC, and xref to DG stuff.

[xxx include $\mathbf{F} \cdot d\mathbf{r}$.]

Definition 2.27. Given a path

$$\gamma(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}$$

for $t \in [a, b]$ and a vector field

$$\mathbf{F}(x, y) = \begin{pmatrix} f(x, y) \\ g(x, y) \end{pmatrix}$$

we write

$$\int_\gamma f(x, y)dx + g(x, y)dy \quad = \quad \int_\gamma f(x, y)dx + \int_\gamma g(x, y)dy$$

where

$$\int_\gamma f(x, y)dx \quad = \quad \int_{t=a}^{t=b} f(x(t), y(t))x'(t)dt$$

and

$$\int_\gamma g(x, y)dy \quad = \quad \int_{t=a}^{t=b} g(x(t), y(t))y'(t)dt.$$

Notation 2.28. Writing

$$d\mathbf{r} = \begin{pmatrix} dx \\ dy \end{pmatrix}$$

we have

$$\mathbf{F} \cdot d\mathbf{r} = f(x, y)dx + g(x, y)dy$$

and so we write

$$\int_\gamma f(x, y)dx + g(x, y)dy = \int_\gamma \mathbf{F} \cdot d\mathbf{r}.$$

xxx rem about **work** done by a **force**, justifying the use of the dot.

Remark 2.29. An expression of the form

$$\int_{x=0}^{x=1} f(x, y)dy$$

may appear shady. However, in terms of the definition of the path integral, this is zero: for the parameterization path along the x axis from 0 to 1, $y(t)$ is constant and so $y'(t)$ is zero along that path. Thus,

$$\int_{x=0}^{x=1} f(x, y)dy = 0.$$

2.5 Change of coordinates

In this section we examine what happens to derivatives and integrals when we change coordinates. This foreshadows **transition functions** on manifolds, as discussed in section [xxx xref].

xxx xref to tensor stuff as well. Several more xrefs ....

2.5.1 Change of coordinates for derivatives

Whether we change coordinates in the domain or the range of a function, we use the chain rule as expected to compute the derivative in the new coordinates.

Let $\gamma(t) : \mathbb{R} \to \mathbb{R}^m$ be a path. Let $x = x(t)$ be a nameless change-of-coordinate function, in the sense of remark 1.13. Then we get the derivative of the reparameterized path using the **chain rule**:

$$\frac{d}{dt}\gamma(x(t)) = \gamma'(x(t))x'(t).$$

Changing coordinates in the range is the same as section 2.2.5:

$$\frac{d}{dt}\mathbf{F}(\gamma(t)) = D\mathbf{F}(\gamma(t))\gamma'(t).$$

xxx note how paths change coordinates. Use the jacobian.

xxx connect to pushforward. The COC map is just an $\mathbf{F}$.

xxx note (cite frankel and who else?) that one can define a vector field as *anything* which changes coordinates in this way.

If $G(x, y)$ is a scalar function, with x and y reparameterized using s and t, then the chain rule gives us

$$\frac{d}{ds}G(x(s,t), y(s,t)) = \frac{\partial G}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial G}{\partial y}\frac{\partial y}{\partial s}$$

etc. Likewise for maps from $\mathbb{R}^m$ to $\mathbb{R}^n$.

2.5.2 Change of coordinates for single-variable integrals

We all know how to do u-substitution, but let's take a fresh look at it. We can attach some new terminology which will help us in section [xxx xref]. I will start with an absurdly simple example — it contains (for me at least) a few surprises. Compute

$$\int_{x=0}^{x=1} x\,dx = \left[\frac{x^2}{2}\right]_{x=0}^{x=1} = \frac{1}{2}.$$

A u-substitution is unnecessary but let's do one for the sake of discussion: say, $u = 2x$. With reference to section [xxx xref], we are measuring the area under the curve with a new ruler — one which measures half-inches rather than inches. We expect the new ruler to overcount area, and so we expect to compensate by dividing by four. We have

$$u = 2x, \quad du = 2\,dx$$

so

$$\frac{du}{2} = dx$$

and so

$$\int_{x=0}^{x=1} x\,dx = \int_{u=0}^{u=2} \frac{u}{2}\,\frac{du}{2} == \left[\frac{u^2}{8}\right]_{u=0}^{u=2} = \frac{1}{2}.$$

I don't know about you, but when I do a u-substitution, I think of x as the "old" variable and u as the "new" variable. But let's trace back through what we did in that example, using symbols this time. We started with

$$\int_{x=a}^{x=b} f(x)\,dx.$$

We chose a new variable as a function of the old,

$$u = u(x),$$

then found

$$du = \frac{du}{dx}\,dx.$$

But when we substituted in the integrand we wrote $x = x(u)$:

$$x = \frac{u}{2} \quad \text{and} \quad dx = \frac{dx}{du}\,du.$$

That is, we mapped from u to x and obtained

$$\int_{x=a}^{x=b} f(x)\,dx = \int_{u=u(a)}^{u=u(b)} f(x(u))\,\frac{dx}{du}\,du.$$

Note that the integrand uses the function $x = x(u)$, while the limits of the integral use the function $u = u(x)$. In summary, when we transition from u coordinates to x coordinates, the behavior of the integrand $f(x)\,dx$ looks reminiscent of the **pullback** diagram in section 1.5.6:

$$\begin{tikzcd}
f(x(u))\,\frac{dx}{du}\,du \arrow[rr, "x(u)"] \arrow[dr] & & f(x)\,dx \arrow[dl] \\
& \mathbb{R} &
\end{tikzcd}$$

$[x^{-1}(a), x^{-1}(b)]$ $\qquad$ $[a, b]$

We will revisit this picture in [xxx xref fwd. to transition functions, $\omega = f(x)\,dx$ is a form, $[a, b]$ is a chain, integral is pairing, $x^*(\omega) = \omega \circ x$ is a pullback of a form, etc. etc.]

* * *

Here is another way to think of the change-of-coordinate formula for single-variable integrals. Again start with

$$\int_{x=a}^{x=b} f(x)\,dx.$$

xxx picket figure (inkscape).

xxx form a Riemann sum and examine a single cell.

Old: x from a to b. Area:

$$(b-a)f(a).$$

New: u from c to d, where $a = x(c)$ and $b = x(d)$. Area:

$$\left(\frac{b-a}{d-c}\right)\,(d-c)\,f(x(c))$$

But

$$\frac{(b-a)}{(d-c)} = \frac{\Delta x}{\Delta u}$$

so the area becomes

$$f(a)\,\Delta x = \frac{\Delta x}{\Delta u}\,(d-c)\,f(x(c)) = \frac{\Delta x}{\Delta u}\,\Delta u\,f(x(c)).$$

Returning to the limit we have

$$f(x)dx \approx f(x(u))\,\frac{dx}{du}\,du.$$

Result:

$$\int_{x=a}^{x=b} f(x)\,dx = \int_{u=x^{-1}(a)}^{u=x^{-1}(b)} f(x(u))\,\frac{dx}{du}\,du.$$

2.5.3 Change of coordinates for integrals of scalar functions

For scalar functions of more than one variable, we encounter once again the Jacobian matrix.

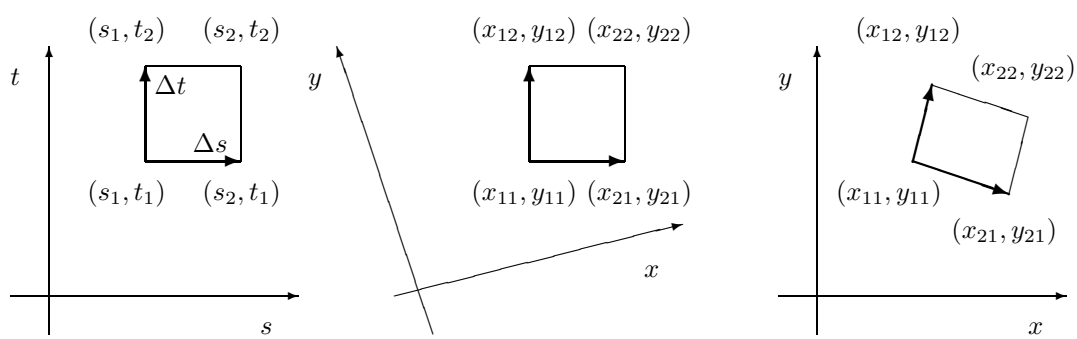
I want to show that

$$\int_R G(\mathbf{x})\,d\mathbf{x} = \int_{\mathbf{x}^{-1}(R)} G(\mathbf{x}(\mathbf{s}))\left|\frac{\partial \mathbf{x}}{\partial \mathbf{s}}\right| d\mathbf{s}.$$

Here is a picture of the function:

xxx tent figure here first. Then:

We form a Riemann sum, as in [xxx xref], then reduce to consideration of a single rectangle where G is approximately linear. We want to consider what happens to the floor under this cell of G in the new coordinates. [xref back to section 1.4.1.] Then we have:



As in section 1.4.1, the area of the rectangle in the original coordinates is the area of the parallelogram spanned by the (perpendicular) vectors

$$\begin{pmatrix} \Delta s \\ 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 \\ \Delta t \end{pmatrix}$$

where

$$\Delta s = s_2 - s_1 \qquad \text{and} \qquad \Delta t = t_2 - t_1.$$

As in section 1.4.1, this is the determinant

$$\det \begin{pmatrix} \Delta s & 0 \\ 0 & \Delta t \end{pmatrix}$$

which simply the product

$$\Delta s \, \Delta t.$$

The question is, what is the area of the parallepiped spanned by the images of Δs and Δt, in (x, y) coordinates? For brevity, I will write

$$\begin{aligned} x_{11} &= x(s_1, t_1), & y_{11} &= y(s_1, t_1), \\ x_{12} &= x(s_1, t_2), & y_{12} &= y(s_1, t_2), \\ x_{21} &= x(s_2, t_1), & y_{21} &= y(s_2, t_1), \quad \text{and} \\ x_{22} &= x(s_2, t_2), & y_{22} &= y(s_2, t_2). \end{aligned}$$

The images of the two edges in question are

$$\begin{pmatrix} \Delta s \\ 0 \end{pmatrix} \mapsto \begin{pmatrix} x_{21} - x_{11} \\ y_{21} - y_{11} \end{pmatrix} \qquad \text{and} \qquad \begin{pmatrix} 0 \\ \Delta t \end{pmatrix} \mapsto \begin{pmatrix} x_{12} - x_{11} \\ y_{12} - y_{11} \end{pmatrix}.$$

We have a nameless change-of-coordinate function $\mathbf{x} = \mathbf{x}(\mathbf{s})$, with Jacobian written as in notation 2.18 $\partial \mathbf{x}/\partial \mathbf{s}$. We know from section 2.2.4 that this Jacobian approximates differences:

$$\begin{pmatrix} x_{21} - x_{11} \\ y_{21} - y_{11} \end{pmatrix} \approx \begin{pmatrix} \partial x/\partial s & \partial x/\partial t \\ \partial y/\partial s & \partial y/\partial t \end{pmatrix} \Bigg|_{\left(\begin{smallmatrix} s_0 \\ t_0 \end{smallmatrix}\right)} \begin{pmatrix} \Delta s \\ 0 \end{pmatrix} = \begin{pmatrix} \partial x/\partial s \Delta s \\ \partial y/\partial s \Delta s \end{pmatrix}$$

and

$$\begin{pmatrix} x_{12} - x_{11} \\ y_{12} - y_{11} \end{pmatrix} \approx \begin{pmatrix} \partial x/\partial s & \partial x/\partial t \\ \partial y/\partial s & \partial y/\partial t \end{pmatrix} \Bigg|_{\left(\begin{smallmatrix} s_0 \\ t_0 \end{smallmatrix}\right)} \begin{pmatrix} 0 \\ \Delta t \end{pmatrix} = \begin{pmatrix} \partial x/\partial t \Delta t \\ \partial y/\partial t \Delta t \end{pmatrix}.$$

So, just as in section 1.4.1, the area of the parallelogram spanned by these two vectors is the determinant

$$\det \begin{pmatrix} \partial x/\partial s \Delta s & \partial y/\partial s \Delta s \\ \partial x/\partial t \Delta t & \partial y/\partial t \Delta t \end{pmatrix}.$$

But due to the multilinearity of the determinant, we can factor out Δs from the first column and Δt from the second to obtain

$$\Delta s \, \Delta t \det \begin{pmatrix} \partial x/\partial s & \partial y/\partial s \\ \partial x/\partial t & \partial y/\partial t \end{pmatrix} \Bigg|_{\left(\begin{smallmatrix} s_0 \\ t_0 \end{smallmatrix}\right)} = \Delta s \, \Delta t \left| \frac{\partial \mathbf{x}}{\partial \mathbf{s}} \right|.$$

Then the ratio of the area of the image parallelogram to the area of the original rectangle is

$$\frac{\Delta s \, \Delta t \left| \frac{\partial \mathbf{x}}{\partial \mathbf{s}} \right|}{\Delta s \, \Delta t} = \left| \frac{\partial \mathbf{x}}{\partial \mathbf{s}} \right|.$$

Forming the Riemann sum of such cells and taking the limit as Δs and Δt go to zero gives the result claimed at the top of this section.

Note that I have discussed this situation only for $G : \mathbb{R}^2 \to \mathbb{R}$, for clarity of presentation, but in fact nothing here is dependent on $n = 2$.

2.6 Integration theorems

Here we recall several theorems and constructions, stated here for ready reference. See [**Anton**] or [**HHGM**] for more information. These theorems will be proved here: such questions occur from time to time on the geometry qualifying exam.

We will see in section 7.2 that the generalized Stokes theorem, which is the main point of the geometry course, subsumes all of the following as special cases:

- The fundamental theorem of calculus (section 2.6.2).
- Green's theorem (section 2.6.4).
- Classical Stokes (section 2.6.5).
- Divergence theorem (section 2.6.6).
- Cauchy's theorem (section 2.6.8).

There are two main points here:

- Most of the theorems in this section convert one kind of integral into another.
- In particular, they follow the pattern of the fundamental theorem of calculus (section 2.6.2): the integral of a derivative of a function over a region is equated to the integral of the original function over the boundary of the region. We will discuss this situation in greater generality in section 7.

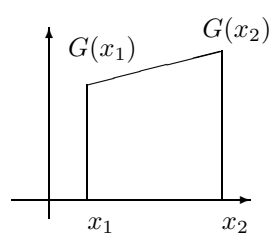
xref fwd: point out things we currently *cannot* do.

2.6.1 Pieces for the integration theorems

Remark 2.30. Given a function $G : \mathbb{R} \to \mathbb{R}$, note that the left-hand approximation to $G'(x)$ at x_1 is the same as the right-hand approximation to $G'(x)$ at x_2, i.e.

$$G'(x_1) \approx \frac{G(x_2) - G(x_1)}{\Delta x} \approx G'(x_2)$$

where $\Delta x = x_2 - x_1$:

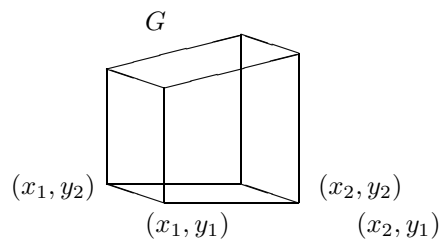


This seems obvious, but it is one of the tricks to remember in the proof of Green's theorem in section 2.6.4.

Remark 2.31. I can re-use the same picture as in remark 2.30 for one cell of a Riemman sum, using the **trapezoid rule**:

$$\int_{x_1}^{x_2} f(x)\, dx \approx \frac{f(x_1) + f(x_2)}{2}\, \Delta x.$$

Remark 2.32. Likewise, we can use the trapezoid rule for a 2D cell:



$$\begin{aligned}
\int_{y_1}^{y_2} \int_{x_1}^{x_2} G(x,y)\, dx\, dy &= \int_{y_1}^{y_2} \left[\int_{x_1}^{x_2} G(x,y)\, dx \right] dy \\
&\approx \int_{y_1}^{y_2} \frac{G(x_1,y) + G(x_2,y)}{2}\, \Delta x\, dy \\
&\approx \frac{G(x_1,y_1) + G(x_2,y_1) + G(x_1,y_2) + G(x_2,y_2)}{4} \Delta x\, \Delta y.
\end{aligned}$$

Notation 2.33. On a cell with corners (x_1, y_1), (x_1, y_2), (x_2, y_1), and (x_2, y_2), write $G_{11} = G(x_1, y_1)$ and so on. Then we have the more streamlined notation

$$\int_{y_1}^{y_2} \int_{x_1}^{x_2} G(x,y)\, dx\, dy \approx \frac{G_{11} + G_{12} + G_{21} + G_{22}}{4} \Delta x\, \Delta y.$$

That is, we find the height of the cell by averaging the heights of the four posts.

Remark 2.34. Likewise, for a single cell of a 3D Riemann sum:

$$\int_{z_1}^{z_2} \int_{y_1}^{y_2} \int_{x_1}^{x_2} G(x,y)\, dx\, dy\, dz \approx \frac{G_{111} + G_{112} + G_{121} + G_{122} + G_{211} + G_{212} + G_{221} + G_{222}}{8}\, \Delta x\, \Delta y\, \Delta z.$$

2.6.2 The fundamental theorem of calculus

Theorem 2.35 (Fundamental theorem of calculus)**.** *If G is an antiderivative of g, then*

$$\int_a^b g(x)dx = G(b) - G(a).$$

Proof. Since G is an antiderivative of g, g is the derivative of G, i.e. $g = G'$. We form a Riemann sum for the left-hand side as

$$\int_a^b G'(x)dx \approx \sum_{k=0}^{n-1} G'(x_k)\Delta x_k$$

where $a = x_0 < x_1 < \ldots < x_{n-1} < x_n = b$ and $\Delta x_k = x_{k+1} - x_k$. Then

$$G'(x_k) \approx \frac{G(x_{k+1}) - G(x_k)}{\Delta x}$$

so the sum becomes

$$\sum_{k=0}^{n-1} G'(x_k)\Delta x \approx \sum_{k=0}^{n-1} \frac{G(x_{k+1}) - G(x_k)}{\Delta x} \Delta x = \sum_{k=0}^{n-1} G(x_{k+1}) - G(x_k).$$

But this is a **telescoping sum**

$$G(x_1) - G(x_0) + G(x_2) - G(x_1) + G(x_3) - G(x_2) + \ldots + G(x_n) - G(x_{n-1}) = G(x_n) - G(x_0) = G(b) - G(a).$$

The lack of rigor here is in my use of the symbol $\approx$. More precisely, one replaces $G'(x_k)$ with $G'(c_k)$ where c_k is the point in the interval (x_k, x_{k+1}) whose existence is guaranteed by the **mean value theorem**, such that

$$G'(c_k) = \frac{G(x_{k+1}) - G(x_k)}{\Delta x}.$$

□

Regardless of the approach, the three key points are (1) **linearization** of G, (2) cancellation of the integral's numerator Δx's by the derivative's denominator Δx's, and **telescoping**. These three points all reappear in the proof of Green's theorem in section 2.6.4.

[xxx note connected domain. point out what happens if not. xref/foreshadow.]

2.6.3 The second fundamental theorem of calculus

Theorem 2.36 (Second fundamental theorem of calculus)**.** *If*

$$F(x) = \int_a^x f(t)dt,$$

then

$$F'(x) = f(x).$$

Proof heuristic. The left-hand side is

$$\begin{aligned}
\frac{d}{dx} F(x) &= \frac{d}{dx} \int_a^x f(t)\, dt \\
&\approx \frac{F(x + \Delta x) - F(x)}{\Delta x} \\
&= \frac{\int_a^{x+\Delta x} f(t)\, dt - \int_a^x f(t)\, dt}{\Delta x} \\
&= \frac{\int_x^{x+\Delta x} f(t)\, dt}{\Delta x} \\
&\approx \frac{f(x)\Delta x}{\Delta x} \\
&= f(x).
\end{aligned}$$

[xxx make a picture w/ **linearized** extra strip at the right edge.] □

2.6.4 Green's theorem

Green's theorem: Let σ be a region of $\mathbb{R}^2$ with oriented boundary C. Let $f : \sigma \to \mathbb{R}$.

$$\int_C f(x, y)dx + g(x, y)dy = \int\int_\sigma (\partial g/\partial x - \partial f/\partial y)dA.$$

Theorem 2.37. *Let* $\mathbf{F} = (P, Q)$ *[xxx stack vert]* $: \mathbb{R}^2 \to \mathbb{R}^2$ *with continuous partials. Let* U *be a simply connected [xxx define above and xref backward] region of* $\mathbb{R}^2$*. Let* C *be a counterclockwise closed path in* U*, with interior* D*. Then*

$$\oint_C P(x,y)\,dx + Q(x,y)\,dy = \int\int_D \left(\frac{\partial Q}{\partial y} - \frac{\partial P}{\partial x}\right)\,dx\,dy.$$

Proof. xxx grids. Mention **linearization** and **telescoping**. xref back to proof of FTC. Use G_{12} notation. □

2.6.5 Classical Stokes

Theorem 2.38. *Copy and paste into here.*

Proof. [xxx goes here]. □

Classical Stokes: Let σ be a two-dimensional surface in $\mathbb{R}^3$ with oriented boundary C which in turn has outward-pointing normal $\hat{\mathbf{n}}$. (To quote Ben Poletta, if σ is a potato chip then C is the peel on the potato chip.) Let $\mathbf{F}$ be a vector-to-vector function, i.e. $\mathbf{F} : \mathbb{R}^3 \to \mathbb{R}^3$, defined on σ.

$$\int_C \mathbf{F} \cdot d\mathbf{r} = \int\int_\sigma (\nabla \times \mathbf{F}) \cdot \hat{\mathbf{n}}\,dS.$$

(Note that this reduces to Green's theorem when σ is confined to the plane, since then

$$\mathbf{F} = \begin{pmatrix} f \\ g \\ h \end{pmatrix}, \qquad \hat{\mathbf{n}} = \hat{\mathbf{z}}, \qquad \text{and} \qquad d\mathbf{r} = dx + dy + dz.$$

Alternatively, classical Stokes is the extension of Green's theorem when we allow R to levitate off the plane, and perhaps warp in the breeze.)

2.6.6 Divergence theorem

Theorem 2.39. *Copy and paste into here.*

Proof. Mimic proof of Green as much as possible. Use G_{122} notation. Use $(\mathbf{V} \cdot \hat{\mathbf{n}})dA$ notation. □

Divergence theorem: Let G be a solid region in $\mathbb{R}^3$ (e.g. a solid potato) with oriented boundary σ (the peel). Let $\hat{\mathbf{n}}$ be the outward-pointing normal. Let $\mathbf{F}$ be a vector-to-vector function, i.e. $\mathbf{F} : \mathbb{R}^3 \to \mathbb{R}^3$, defined on G.

$$\int_\sigma (\mathbf{F} \cdot \hat{\mathbf{n}})\,dS = \int\int\int_G (\nabla \cdot \mathbf{F})\,dV.$$

2.6.7 Fundamental theorem of calculus for line integrals

Theorem 2.40. *Goes here.*

Proof. xxx Use Green's theorem / classical Stokes on closed loops to get path independence. Type up handwritten notes. □

2.6.8 Cauchy's theorem

Cauchy's theorem

2.7 Lagrange multipliers

Lagrange multipliers are used to solve the following problem (stated here for $\mathbb{R}^3$): Maximimize the function $f(x, y, z)$, subject to the constraint that $g(x, y, z) = 0$.

When I first took calculus, this seemed to be a quite arbitrary thing to want. In differential geometry, though, it's the natural thing to use when restricting a function from Euclidean space (e.g. $\mathbb{R}^3$) to a manifold (e.g. $\mathbb{S}^2$), when that manifold is the level set of some function. (E.g. $\mathbb{S}^2$ is the subset of $\mathbb{R}^3$ subject to the constraint that $g(x, y, z) = x^2 + y^2 + z^2 - 1 = 0$.)

Proposition 2.41 (Anton's theorem 16.10.1)**.** *Let $f, g : \mathbb{R}^3 \to \mathbb{R}$ with continuous first partials on some open set containing the constraint surface $g(x, y, z) = 0$. Further suppose that g has non-zero gradient on this surface. If f has a constrained relative extremum, it occurs at a point (x_0, y_0, z_0) where the gradients of f and g are parallel, i.e.*

$$\nabla f(x_0, y_0, z_0) = \lambda \nabla g(x_0, y_0, z_0)$$

for some real number λ.

Remark 2.42. Recall from remark 13.4 that ∇g is a normal vector to the surface defined by $g = 0$. Thus, the proposition is that if f has an extremum on the surface, it occurs when f's gradient is normal to the surface.

Remark 2.43. Not all the points where ∇f is parallel to ∇g are necessarily extrema for f on g's zero set. We have the following:

- Solving $\nabla f = \lambda \nabla g$ will give you points at which the two gradients are parallel.
- As explained in remark 2.45 below, these will all be critical points.
- Of the critical points, some may be extrema; others may be saddle points.
- Remember that extrema of f occur at critical points of f or on the boundary (if any) of the intersection of the domains of f and g.

Example 2.44. $\triangleright$ Problem: maximize $f(x, y, z) = y^2 - z$ on $\mathbb{S}^2$, i.e. subject to the constraint $g(x, y, z) = x^2 + y^2 + z^2 - 1 = 0$.

First compute the gradients:

$$\nabla f = \begin{pmatrix} 0 \\ 2y \\ -1 \end{pmatrix} \qquad \text{and} \qquad \nabla g = \begin{pmatrix} 2x \\ 2y \\ 2z \end{pmatrix}.$$

Observe that ∇g doesn't vanish on $\mathbb{S}^2$, since it only vanishes at $(x, y, z) = (0, 0, 0)$ which is not a point on $\mathbb{S}^2$. Thus the proposition applies. Next, find when ∇f and ∇g are parallel:

$$\begin{cases} 0 &= 2\lambda x \\ 2y &= 2\lambda y \\ -1 &= 2\lambda z \end{cases}$$

$$\begin{cases} \lambda x &= 0 \\ (1-\lambda)y &= 0 \\ \lambda z &= -1/2 \end{cases}$$

$$\begin{cases} \lambda = 0 & \text{or} & x = 0 \\ \lambda = 1 & \text{or} & y = 0 \\ \lambda z = -1/2 & & \end{cases}$$

- Case $\lambda = 0$ and $\lambda = 1$: Absurd.
- Case $\lambda = 0$ and $y = 0$: The third equation cannot be satisfied.
- Case $x = 0$ and $\lambda = 1$: $x = 0$ and $z = -1/2$, from which $y = \pm\sqrt{1 - x^2 - z^2} = \pm\sqrt{3/4} = \pm\sqrt{3}/2$.
- Case $x = 0$ and $y = 0$: $z = \pm 1$, with $\lambda = \mp 1/2$.

This gives us four candidate extremum points for f on $\mathbb{S}^2$:

$$\begin{pmatrix} 0 \\ \sqrt{3}/2 \\ -1/2 \end{pmatrix}, \begin{pmatrix} 0 \\ -\sqrt{3}/2 \\ -1/2 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}.$$

Evaluating f at these four points yields, respectively,

$$\frac{5}{4}, \frac{5}{4}, -1, 1.$$

Thus, the maxima of f are at

$$\begin{pmatrix} 0 \\ \sqrt{3}/2 \\ -1/2 \end{pmatrix} \qquad \text{and} \qquad \begin{pmatrix} 0 \\ -\sqrt{3}/2 \\ -1/2 \end{pmatrix},$$

while the minimum of f (which was not asked for) is at the north pole. Note that $\mathbb{S}^2$ has no boundary so there are no extrema at boundaries.

[xxx incorporate Matlab figures from `y2z.m`.] $\triangleleft$

Remark 2.45. It may appear at first blush that something is wrong. Take the first point (call it $\mathbf{q}$), for example: If we evaluate ∇f there, we don't get zero:

$$\mathbf{q} = \begin{pmatrix} 0 \\ \sqrt{3}/2 \\ -1/2 \end{pmatrix}; \quad \nabla f = \begin{pmatrix} 0 \\ 2y \\ -1 \end{pmatrix}; \quad \nabla f(\mathbf{q}) = \begin{pmatrix} 0 \\ \sqrt{3} \\ -1 \end{pmatrix}.$$

However, it is not the gradient of f but rather the gradient of f restricted to $\mathbb{S}^2$ which vanishes. (We might write this $\nabla^{\mathbb{S}^2} f$.) How can we compute this restricted gradient? We could do the following:

- Compute the tangent plane to the sphere at $\mathbf{q}$.
- Use the projection operator (1.3.2) to project $\nabla f(\mathbf{q})$ down onto that plane.
- Then see if that projected vector is zero.

Alternatively, we can simply recall [xxx xref bkwd] that the normal to the sphere is given by ∇g. The normal is perpendicular to the tangent plane at each point, and we've found that ∇f is parallel to ∇g, i.e. normal to the tangent plane. So we know that if we project $\nabla f(\mathbf{q})$ down to the tangent plane at $\mathbf{q}$, we will get zero. This sanity check gives us reassurance that we've really found critical points using this method.

2.8 ODEs in one variable

Consult your favorite ODE text for full information, of course ... for qualifier problems, the only types of single-variable ODEs you are likely to need to solve are separable first-order ones, or linear homogeneous second-order with constant coefficients. That is, you don't need to re-learn a thick volume worth of ODE techniques.

Here an example of a first-order equation:

$$\begin{aligned}
\frac{dy}{dx} &= xy \\
\frac{dy}{y} &= dx \\
\int \frac{dy}{y} &= \int dx \\
\ln(y) &= x + C \\
y &= e^{x+C} = ke^x.
\end{aligned}$$

Then use the initial conditions to determine C.

Here is the general technique for linear homogeneous second-order with constant coefficients. We start with

$$ax'' + bx' + c = 0.$$

Guess $x = e^{kt}$ for k to be determined. Then $x' = ke^{kt}$ and $x'' = k^2e^{kt}$. Plug these back in and recall that e^{kt} is non-vanishing:

$$\begin{aligned}
(ak^2 + bk + c)e^{kt} &= 0 \\
ak^2 + bk + c &= 0 \\
k &= \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}
\end{aligned}$$

In the case there are two distinct k's, say k_1 and k_2, write

$$x = Ae^{k_1 t} + Be^{k_2 t}.$$

Then apply the initial conditions and solve for A and B. If the quadratic formula gives one k, then use

$$x = Ae^{kt} + Bte^{kt}.$$

If k is imaginary, then you can directly convert the complex exponentials to sine and cosine using the identities in section 1.2. Or, you can directly take your solution to be in terms of sines and cosines. This is shown by example:

$$\begin{aligned}
x'' + 4x &= 0 \\
k^2 + 4 &= 0 \\
k &= \pm 2i \\
x &= Ae^{2it} + Be^{-2it}.
\end{aligned}$$

Now

$$\cos(2t) = \frac{e^{2it} + e^{-2it}}{2} \qquad \text{and} \qquad \sin(2t) = \frac{e^{2it} - e^{-2it}}{2i}$$

so

$$\begin{aligned} x &= Ae^{2it} + Be^{-2it} \\ &= (A+iB)\frac{e^{2it}+Be^{-2it}}{2} \;+\; (A-iB)\frac{e^{2it}-Be^{-2it}}{2i} \\ &= C\cos(2t) \;+\; D\sin(2t). \end{aligned}$$

That is, since the constants are to be determined anyway, we can either write down the solution as

$$x = Ae^{2it} + Be^{-2it}$$

or as

$$x = C\cos(2t) \;+\; D\sin(2t).$$

If k is real with distinct roots, you can again convert the exponentials to sinh and cosh, or take the solution directly in terms of sinh and cosh as follows:

$$\begin{aligned} x'' - 4x &= 0 \\ k^2 - 4 &= 0 \\ k &= \pm 2 \\ x &= Ae^{2t} + Be^{-2t} \qquad \text{or} \\ x &= C\cosh(2t) + D\sinh(2t). \end{aligned}$$

If k has distinct roots which aren't purely real or imaginary, proceed as follows:

$$\begin{aligned} x'' - 4x' + 5x &= 0 \\ k^2 - 4k + 5 &= 0 \\ k &= 2 \pm i \\ x &= Ae^{2t}e^{it} + Be^{2t}e^{-it} \qquad \text{or} \\ x &= Ce^{2t}\cos(t) + De^{2t}\sin(t). \end{aligned}$$

2.9 ODEs in two variables

As in previous sections, this material is review. However, we want to be sure that what we do later on, with new terminology, is consistent with what we already know.

Systems of ODEs on the quals tend to be separable second-order. One eliminates variables, then uses techniques from section 2.8. This is best illustrated by example.

In $\mathbb{R}^2$, consider the system of ODEs given by

$$\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = \begin{pmatrix} -y \\ x \end{pmatrix}.$$

Taking second derivatives, we get

$$\begin{pmatrix} \ddot{x} \\ \ddot{y} \end{pmatrix} = \begin{pmatrix} -\dot{y} \\ \dot{x} \end{pmatrix} = \begin{pmatrix} -x \\ -y \end{pmatrix}.$$

This gives two ODEs, each in one variable alone:

$$\begin{aligned} \ddot{x} + x &= 0 \\ \ddot{y} + y &= 0 \end{aligned}$$

These have the solutions from section 2.8, namely,

$$\begin{aligned} x &= a\cos t + b\sin t \\ y &= c\cos t + d\sin t. \end{aligned}$$

Putting initial conditions $(x(0), y(0)) = (x_0, y_0)$ gives

$$x(0) = a, \quad y(0) = c$$

so we have

$$\begin{aligned} x &= x_0\cos t + b\sin t \\ y &= y_0\cos t + d\sin t. \end{aligned}$$

But these two equations are not independent; the original system was **coupled** in the sense that $\dot{x}$ depends on y and vice versa:

$$\begin{array}{rclclcl} \dot{x} &=& -x_0\sin t + b\cos t &=& -y &=& -y_0\cos t + d\sin t \\ \dot{y} &=& -y_0\sin t + d\cos t &=& x &=& x_0\cos t + b\sin t. \end{array}$$

from which

$$\begin{array}{rclclcl} \dot{x} &=& -x_0\sin t + b\cos t &=& -y &=& -y_0\cos t + d\sin t \\ \dot{y} &=& -y_0\sin t + d\cos t &=& x &=& x_0\cos t + b\sin t. \end{array}$$

which gives

$$b = -y_0, \quad d = x_0$$

and so the solution is

$$\begin{aligned} x &= x_0\cos t - y_0\sin t \\ y &= y_0\cos t + x_0\sin t \end{aligned}$$

or

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}.$$

Note that this matrix is the **rotation matrix** for angle t. That is, any point (x_0, y_0) will be revolved counterclockwise about the origin as t increases. The origin itself stays fixed.

2.10 PDEs

xxx TBD. Lee makes a cryptic reference to $d\eta = \omega$ as a PDE .... Include some specific examples, with relevance.

2.11 Limitations of vector calculus

2.11.1 xxx 0

FTC with disconnected domain. Note this is no big deal: it just splits up.

2.11.2 xxx 1

Cite and use Massey example: non-existence of potential function, given a 1-hole.

2.11.3 xxx 2

Cite and use Massey example: function with a 2-hole.

3 Preliminaries from topology and analysis

Analysis does not set out to make pathological reactions impossible, but to give the patient's ego freedom to decide one way or another. — Sigmund Freud (1856-1939).

xxx emph that much of this is *used* in vector calculus, but not stated as such. Now, in graduate school, we use these things explicitly.

3.1 Connectedness?

3.2 Compactness?

Warrants any separate discussion?

3.3 Homeomorphisms and diffeomorphisms

Recall the following:

Definition 3.1. A function from one topological space to another is a **homeomorphism** iff it is a bijection which is continuous with continuous inverse.

This is the isomorphism in the category of topological spaces. On Euclidean spaces (more generally, on Banach spaces) we have the notion of differentiability.

Definition 3.2. A function from $\mathbb{R}^m$ to $\mathbb{R}^n$ is C^k if it is k-times differentiable. A function which may be differentiated arbitrarily many times is said to be C^∞, or **smooth**.

Most of the functions we deal with in this course are C^∞. This is a convenience which frees us from having to track how many times a given function is needs to be differentiable in each situation.

Definition 3.3. A function from $\mathbb{R}^m$ to $\mathbb{R}^n$ is a **diffeomorphism** iff it is a bijection which is smooth with smooth inverse.

Note that this is distinct from a definition used in other contexts, namely, that a diffeomorphism is a bijection from $\mathbb{R}^m$ to $\mathbb{R}^n$ which is differentiable with differentiable inverse. When need be, we may distinguish these two concepts by saying that the latter is a C^1 diffeomorphism, whereas the definition given above describes a C^∞ diffeomorphism.

Example 3.4. $\triangleright$ The canonical example of a homeomorphism which is not a diffeomorphism is $f : \mathbb{R} \to \mathbb{R} : x \mapsto x^3$, the inverse of which is not differentiable at 0. $\triangleleft$

3.4 Implicit function theorem

Theorem 3.5 (Implicit function theorem)**.** *Let E be an open subset of $\mathbb{R}^{m+n}$, and let $\mathbf{f} : E \to \mathbb{R}^n$ be a vector-to-vector function which is continuously differentiable, i.e. of type C^2. Let*

$$(\mathbf{a}, \mathbf{b}) = (a_1, \ldots, a_m, b_1, \ldots, b_n) \in E$$

be such that

$$\mathbf{f}(\mathbf{a}, \mathbf{b}) = 0.$$

Let D be the $n \times n$ submatrix of the Jacobian of f given by $\partial f_i / \partial x_j$, for $i = 1, \ldots, n$ and $j = m+1, \ldots, n$, evaluated at $\mathbf{b}$. If $\det(D) \neq 0$, then there exists a neighborhood U of $\mathbf{a}$ and a unique C^2 function $\mathbf{g} : U \to \mathbb{R}^n$ such that

$$\mathbf{g}(\mathbf{a}) = \mathbf{b}, \quad \textit{i.e.} \quad \mathbf{f}(\mathbf{a}, \mathbf{g}(\mathbf{a})) = 0,$$

and for all $\mathbf{a}' \in U$,

$$\mathbf{f}(\mathbf{a}', \mathbf{g}(\mathbf{a}')) = 0.$$

That is, we can **solve for** *the* $\mathbf{b}$ *variables at and near* $\mathbf{a}$.

Proof. See any of the geometry texts in the bibliography. □

Remark 3.6. The point is that, given a system of equations (for this course, usually a single equation), we have an easy criterion for when we can solve for some variables in terms of the others. Note however that the implicit function theorem ensures existence and uniqueness; actually *finding* the function $\mathbf{g}$ is another matter.

Remark 3.7. The theorem, as stated, has the to-be-solved-for variable(s) written last. In practice, this may not be the case. E.g. given a function $f(v, w, x, y, z) : \mathbb{R}^5 \to \mathbb{R}^2$, we might want to solve for, say, v and x. In that case, we would need to check the submatrix formed by the first and third columns of the Jacobian of f. Furthermore, we might not know ahead of time which variables to solve for, until we apply the implicit function theorem to various submatrices of the Jacobian.

Remark 3.8. We make use of this theorem when we use **graph coordinates** for a manifold. See section 6.1.2.

Example 3.9. ▷ Let

$$f : \mathbb{R}^3 \to \mathbb{R} : \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto x^2 + y^2 + z^2 - 1.$$

Here we have $m = 2$ and $n = 1$. Then the kernel of f is the sphere $\mathbb{S}^2$. Consider the north pole, $(0, 0, 1)$, written as

$$(\mathbf{a}, \mathbf{b}) = (a_1, a_2, b_1) = (x, y, z) = (0, 0, 1).$$

The Jacobian of f is

$$\begin{pmatrix} 2x & 2y & 2z \end{pmatrix}$$

which evaluated at the north pole is

$$\begin{pmatrix} 0 & 0 & 1. \end{pmatrix}$$

Now, there is only one 1×1 submatrix of this which is non-zero, namely, the last. So, there is a neighborhood U of $(0, 0)$ and a unique function $g : U \to \mathbb{R}$ such that $z = g(x, y)$. Here, it's clear what this is: take

$$z = g(x, y) = \sqrt{1 - x^2 - y^2}.$$

◁

Another example is in section 10.7.1.

3.5 Inverse function theorem

Theorem 3.10 (Inverse function theorem)**.** *Let E be an open subset of $\mathbb{R}^m$, and let $\mathbf{f} : E \to \mathbb{R}^m$ be a vector-to-vector function which is continuously differentiable, i.e. of type C^2. If the Jacobian of $\mathbf{f}$ is nonsingular at a point $\mathbf{q}$ of E, then there exist open neighborhoods U of $\mathbf{q}$ and V of $\mathbf{f}(\mathbf{q})$ such that $\mathbf{f}$ is a diffeomorphsim from U into V.*

Proof. See any of the geometry texts in the bibliography. □

Remark 3.11. The point is that, even if such a function $\mathbf{f}$ is wildly non-linear, to check for invertibility at a point it suffices to make the much simpler check of the invertibility of the linearization of $\mathbf{f}$.

4 Preliminaries from algebra

Structures are the weapons of the mathematician. — Nicholas Bourbaki

4.1 Algebraic axioms

Here I gather together needed axioms. This material is certainly review. My intention is not to insult the reader, but rather to ensure that terminology is uniform throughout this paper.

4.1.1 Groups and semigroups

Following [**Herstein**], I include closure as an axiom. (This is worth doing since it is often lack of closure which keeps a set from being a subgroup. In particular, the even integers are a subgroup of the integers with respect to addition, whereas the odd integers fail to be closed under addition.) Thus, I say that there are four group axioms, rather than the usual three. A **group** G is a set with a binary operation (written using juxtaposition) satisfying the following four axioms:

- G is **closed** under the operation, the operation is **associative**, there is a (unique) **identity**, and every element has a (unique) **inverse**.
- For an **abelian group**, there is a fifth axiom: the operation is **commutative**.

Definition 4.1. A **semigroup**, which will be referred to briefly in section 4.3.2, lives up to the prefix *semi* in that it is *halfway* to being a group: it satisfies the first two of the four group axioms. Namely, it is closed under the semigroup operation, and that operation is associative.

Example 4.2. $\triangleright$ The canonical example of semigroup is strings of letters over some fixed alphabet, with the operation being concatenation. E.g. "ab" pasted together with "c" is "abc". $\triangleleft$

Remark 4.3. If we include the empty string as a letter, then we get an identity. The resulting structure is 3/4 of the way to being a group, and is called a **monoid**. (Mnemonic: the Greek root *monos* means single (or one), and we often write an identity as 1 — which is precisely what a monoid has beyond a semigroup.)

4.1.2 Normalizer and normal closure of a subgroup

xxx re-do this with subsets, not subgroups.

Definition 4.4. Let G be a group with subgroup H. The **normalizer** of H in G, written $N_G(H)$, or just $N(H)$, is the *largest* subgroup of G in which H is normal.

Definition 4.5. Let G be a group with subgroup H. The **normal closure** or **conjugate closure** of H in G, written $\langle H \rangle_N$ or $\overline{H}$, is the *smallest* subgroup of G in which H is normal.

Remark 4.6. In particular, if H is already normal in G (e.g. if G is abelian, in particular if G is cyclic) then $N(H)$ is all of G.

xxx examples

xxx How to actually *compute* such things?

rmk: H in center of G implies H normal in G.

[xxx xref fwd: SvK uses normal closure; action of π_1 on fibers uses normalizer.]

4.1.3 Group actions

Vision without action is a daydream; action without vision is a nightmare.
— Japanese proverb.

xxx left actions most useful elsewhere, and probably are more familiar. for this course, though, we mostly need *right* actions (xref fwd to π_1 acting on fibers).

xxx xref fwd to action of π_1 on fibers.

Definition 4.7. Let G be a group and let S be a set. Let e denote the identity element of G. A map $\mu : S \times G \to S$, written $\mu(s, g) = s \cdot g$ or $\mu(s, g) = s^g$ is said to be a **right group action** if the following two axioms (roughly speaking, identity and associativity) are satisfied:

(i) For all $s \in S$, $s \cdot e = s$, and

(ii) For all $g_1, g_2 \in G$ and for all $s \in S$, $(s \cdot g_1) \cdot g_2 = s \cdot (g_1 g_2)$.

standard examples.

xxx note the standard consequences and definitions (faithful and transitive). Also the orbit-index formula. xref to [**Grove**].

note that each $g \in G$ permutes S (1-1), but that different group elements need not give the *same* permutation (not 1-1).

G mod isotropy *is* 1-1 and onto, if the action is transitive.

note that the action axioms put some requirements on the structure (orbits, total size) of what might otherwise seem an arbitrary set.

massey's "E is a homogenous G-space" simply means that G has a transitive group action on E.

def auts: homs are set maps ϕ such that for all $g \in G$ and all $s \in S$,

$$\phi(s \cdot g) = \phi(s) \cdot g.$$

write as commutative diagram w/ a nice picture. then auts are bijective self-homs.

normalizer in group section? own section?

4.1.4 Group action example

Suppose that you want to teach the "cat" concept to a very young child. Do you explain that a cat is a relatively small, primarily carnivorous mammal with retractible claws, a distinctive sonic output, etc.? I'll bet not. You probably show the kid a lot of different cats, saying "kitty" each time, until it gets the idea. To put it more generally, generalizations are best made by abstraction from experience.
— R. P. Boas

make a big deal of this. standard example.

$\mathbb{Z}$ acting on the vertices of an n-gon (take an octagon for specificity in the picture). (Kind of like $\mathcal{D}_n$ acting on the n-gon, but not a faithful action, and no flips.)

- The group G is $\mathbb{Z}\rho$ where ρ is $1/n$ of the way around.
- The stability subgroup of a point is $H = n\mathbb{Z}$.
- G is cyclic hence abelian hence H is normal. So $N(H) = G$.
- Auts of the action are $N(H)/H = \mathbb{Z}/n\mathbb{Z}$.

4.1.5 Rings

A **ring** is a set R with addition and multiplication satisfying the following 8 axioms:

- R is an **abelian group** with respect to addition.
- R is **closed** with respect to multiplication.
- Multiplication is **associative**.
- Addition and multiplication are related by left and right **distributivity**.

4.1.6 Fields

A **field** is a set F with addition and multiplication satisfying the following 11 axioms:

- F is a **ring**.
- Multiplication is **commutative**; there is a (unique) **multiplicative identity** for all non-zero elements; for each non-zero element, there is a (unique) **multiplicative inverse**.

Another way to clump the 11 field axioms is:

- F is an **abelian group** with respect to addition.
- The non-zero elements of F form an **abelian group** with respect to multiplication.
- Addition and multiplication are related by left and right **distributivity**.

4.1.7 Modules and vector spaces

There are 21 axioms for a vector space V over a field F:

- F is a **field**.
- V is an **abelian group** with respect to vector addition.
- V is **closed** with respect to scalar multiplication (from either side), scalar multiplication is **associative** ($(ab)\mathbf{v} = a(b\mathbf{v})$), the field's **multiplicative identity** 1_F satisfies $1_F\mathbf{v} = \mathbf{v}$ for all vectors, **scalar distributivity** ($a(\mathbf{u}+\mathbf{v}) = a\mathbf{u} + a\mathbf{v}$), **vector distributivity** ($(a+b)\mathbf{v} = a\mathbf{v} + b\mathbf{v}$).

For a module M over a ring R, here I mean a **unital module** [**DF**]: I require R to have a multiplicative identity 1_R and I require $1_R\mathbf{v} = \mathbf{v}$ for all $\mathbf{v} \in M$. Otherwise, the axioms are the same as those for a vector space over a field. There are 19 axioms for a unital module:

- R is a **ring** with multiplicative identity 1_R.
- V is an **abelian group** with respect to vector addition.
- V is **closed** with respect to scalar multiplication (from either side), scalar multiplication is **associative** ($(ab)\mathbf{v} = a(b\mathbf{v})$), the field's **multiplicative identity** 1_F satisfies $1_F\mathbf{v} = \mathbf{v}$ for all vectors, **scalar distributivity** ($a(\mathbf{u}+\mathbf{v}) = a\mathbf{u} + a\mathbf{v}$), **vector distributivity** ($(a+b)\mathbf{v} = a\mathbf{v} + b\mathbf{v}$).

Remark 4.8. Any abelian group is already a $\mathbb{Z}$-module. To see this, let G be an abelian group, and let $x, y \in G$. When we write the group operation multiplicatively, we write products as xy and repeated multiplication as x^n. However, when we write the group operation additively, we write sums at $x + y$ and repeated addition as nx. Thus, scalar multiplication, by integer scalar n, is nothing more than repeated addition n times, using the existing group operation. You can check off the axioms for a $\mathbb{Z}$-module and verify that they are in fact satisfied.

4.1.8 Algebras

There are several ways to think of an algebra. There are the same number of axioms; the difference is only in the way in which those axioms are clumped together. One point of view is taken by [**Lang**]. Let R be a commutative ring with identity. Then an R-algebra is a ring A along with a homomorphism $\phi : R \to A$ such that $\phi(R)$ is contained in the center of A. The other point of view, and the one I prefer here, is taken by [**Hungerford**]:

Definition 4.9. Let R be a commutative ring with identity. An unital R-module A is an **algebra** if it is a ring and if $r(ab) = (ra)b = a(rb)$ for all $r \in R$ and $a, b \in A$.

When the base ring is a field, we can say the following (which is equivalent, but phrased in more familiar terms of fields and vector spaces):

Definition 4.10. Let F be a field. An F-vector space V is an **algebra** if there is a multiplication operation on V making V a **ring**, and such that $r(\mathbf{uv}) = (r\mathbf{u})\mathbf{v} = \mathbf{u}(r\mathbf{v})$ for all $r \in F$ and $\mathbf{u}, \mathbf{v} \in V$.

There are three ideas here:

(i) We already have a field.

(ii) We have a vector space and we want to be able to multiply vectors.

(iii) The old scalar multiplication and the new vector multiplication should be compatible with each other.

There are 25 axioms for an algebra A over a field F:

- F is a field.
- A is a vector space.
- A is **closed** under vector multiplication, and the vector multiplication is **associative**.

- Vector multiplication **distributes** over vector addition.
- For all $r \in F$ and $\mathbf{u}, \mathbf{v} \in V$, $r(\mathbf{uv}) = (r\mathbf{u})\mathbf{v} = \mathbf{u}(r\mathbf{v})$.

Examples of algebras include:

- The ring of polynomials $\mathbb{R}[x]$ is an $\mathbb{R}$-algebra. The vector-space point of view is illustrated by the standard basis $\{1, x, x^2, x^3, \ldots\}$. When we add two polynomials, we add them coordinatewise, by equating like terms. Of course, we can multiply polynomials as well.
- The ring $\mathbb{R}[x, y]$ of real bivariate polynomials has basis $\{x^i y^j : i, j = 0, 1, 2, \ldots\}$.
- The ring of polynomials $\mathbb{R}[x_1, \ldots, x_n]$ is likewise an $\mathbb{R}$-algebra.
- The ring of $m \times n$ matrices with entries in $\mathbb{R}$ is a vector space over $\mathbb{R}$, but if $m = n$, then we can also multiply two such matrices.
- Extension fields: for example, take the base field $\mathbb{Q}$ and the extension field $\mathbb{Q}(i)$, which is the rational complex numbers. We can treat $\mathbb{Q}(i)$ as a 2-dimensional vector space over $\mathbb{Q}$, with elements of the form (a, b) and (c, d), but of course we can also multiply complex numbers using $(a + bi)(c + di)$.
- Likewise, the complex numbers $\mathbb{C}$ are just $\mathbb{R}^2$, along with a special rule for multiplying vectors. In a similar way, the real quaternions are an $\mathbb{R}$-algebra, namely $\mathbb{R}^4$. Unlike for the complexes, though, the vector-times-vector multiplication is not commutative.

Note that an algebra is a vector space with vector-vector multiplication. So, given an algebra we can obtain a vector space by forgetting about the vector-vector multiplication. In turn, a vector space is an abelian group with multiplication by elements of a field. So, given a vector space we can obtain an abelian group by forgetting about the field.

4.1.9 Graded algebras

Definition. An abelian group A is the **direct sum** of abelian groups B and C, written $A = B \oplus C$, if:

(i) $A = B + C$, i.e. for all $a \in A$, there are $b \in B$ and $c \in C$ such that $a = b + c$.

(ii) $B \cap C = \{0\}$.

Note in particular that this definition applies to rings and vector spaces, which are simply abelian groups with some more structure added.

Definition. A ring S is **graded** ([**DF**]) if it is the direct sum of additive subgroups

$$S = S_0 \oplus S_1 \oplus S_2 \oplus \ldots$$

such that for all $i, j \geq 0$, $S_i S_j \subseteq S_{i+j}$. In particular, $s \in S_i$ and $t \in S_j$ implies $st \in S_{i+j}$.

Example 4.11. $\triangleright$ The ring of $\mathbb{R}[x]$ of real univariate polynomials is graded. Each additive subgroup S_i consists of the zero polynomial along with all monomials with degree i. $\triangleleft$

The main examples for us (in fact the only reason for introducing graded rings in this paper) are tensor algebras. These are discussed in section 4.7.9.

4.2 Categories

Mathematics is the art of giving the same name to different things. — Jules Henri Poincaré (1854-1912).

Like several sections of this paper, this section is certainly optional. What I want to do here is to give a thorough, Faris-esque grounding in fundamental terminology. The primary reason for presenting category theory in this paper is as a mnemonic device: for example, it helps us remember the difference between f_* and f^* in differential geometry. If you find that it only adds confusion, you can certainly do differential geometry without it.

4.2.1 Definitions

Category theory is sometimes ([**Lang**]) referred to as *abstract nonsense*. However, it is simply an abstraction of some very concrete things. Namely, we are abstracting the concept of structures (groups, vector spaces, topological spaces, etc.) and mappings (homomorphisms, linear transformations, homeomorphisms, etc.). One oddity is that the definitions below are specifically designed to avoid mentioning the *elements* of those structures.

References are Appendix II of [**DF**], section I.11 of [**Lang**], and section I.7 and chapter X of [**Hungerford**].

Definition 4.12. A **category** $\mathcal{C}$ consists of the following (formally, it is an ordered triple of the following three things):

- There is a class of **objects**.
- For each pair (A, B) of objects, there is a set of **morphisms** or **arrows** from A to B. This set is called a **hom set** and is written $\mathrm{Hom}_{\mathcal{C}}(A, B)$, or more often simply $\mathrm{Hom}(A, B)$. For an element f of $\mathrm{Hom}(A, B)$, we often write $f : A \to B$, or $A \xrightarrow{f} B$.
- There is a **law of composition**: if $f \in \mathrm{Hom}(A, B)$ and $g \in \mathrm{Hom}(B, C)$ (which is to say $f : A \to B$ and $g : B \to C$), then there is an element $g \circ f$ of $\mathrm{Hom}(A, C)$.

The objects, hom sets, and composition law must satisfy the following three axioms:

- The hom sets are **disjoint**: If A_1 and A_2 are different, or if B_1 and B_2 are different, then $\mathrm{Hom}(A_1, B_1)$ and $\mathrm{Hom}(A_2, B_2)$ have no morphisms in common.
- Each object A has an **identity morphism**, written 1_A. Since we are defining morphisms without reference to specific elements of an object, we can't say something like "...such that $1_A(x)$ for all $x \in A$". Rather, an identity morphism on an object A is characterized by the property that for all other objects B, and for all $f : A \to B$ and all $g : B \to A$,
$$f \circ 1_A = f \quad \text{and} \quad 1_A \circ g = g.$$
- Composition of morphisms is **associative**: if $f : A \to B$, $g : B \to C$, and $h : C \to D$, then $h \circ (g \circ f) = (h \circ g) \circ f$.

Category theory is carefully constructed to avoid references to elements of an object. This is a good and powerful thing. However, for this paper, I use categories mostly as a handy way to refer to concepts which apply equally well to rings, vector spaces, modules, etc. (This applies particularly to section 4.5 on sequences of complexes.) I always *do* have in mind categories which *are* sets, and which *do* at least have a zero element.

Definition 4.13. For purposes of this paper, a **concrete category** has objects which are sets, each containing a zero element; the morphisms are homomorphisms.

4.2.2 Examples of categories

Some familiar examples of categories are:

- Objects are sets and morphisms are arbitrary functions.
- Objects are groups and morphisms are group homomorphisms.
- Objects are rings and morphisms are ring homomorphisms.
- Objects are vector spaces and morphisms are linear transformations.

Example 4.14. $\triangleright$ A perhaps less familiar example is a partially ordered set. Let P be a partially ordered set, with operation $\preccurlyeq$. Then the poset P is itself the entire category. Objects of P are the elements of P. Morphisms of P are quite literally arrows: there is an arrow from a to b if $a \preccurlyeq b$. Composition of arrows is nothing more than the transitivity property of posets. The identity morphisms are provided by the reflexivity property of posets, namely, $a \preccurlyeq a$ for all $a \in P$. $\triangleleft$

Here is a specific instance of that example: the divisibility lattice on factors of 12. Here I am not showing all the arrows (morphisms): (1) I am omitting the arrows which may be obtained by composition, e.g. the arrow from 2 to 12 since $2 \mid 4$ and $4 \mid 12$. (2) I am omitting arrows from objects (integers) to themselves, e.g. $2 \mid 2$. I sometimes refer to these, the identity morphisms which objects in any category must possess, as **self-loops**: if we draw arrows as a graph, the identity morphisms appear as loops from each object to itself.

$$\begin{tikzcd}
 & 12 & \\
4 \arrow[ur] & & 6 \arrow[ul] \\
2 \arrow[u] \arrow[urr] & & 3 \arrow[u] \\
 & 1 \arrow[ul] \arrow[ur] &
\end{tikzcd}$$

Example 4.15. $\triangleright$ Another example is a single group G. There is a single object which is all of G; the hom set $\mathrm{Hom}_G(G, G)$ is all of G. That is, the morphisms are the elements of G. The composition of two morphisms x and y is the usual product xy. The identity morphism is the identity element of G. $\triangleleft$

4.2.3 Functors

A **functor** is a map between categories which takes objects to objects and morphisms to morphisms.

Definition 4.16. Let $\mathcal{C}$ and $\mathcal{D}$ be categories. A **covariant functor** T is a map from $\mathcal{C}$ to $\mathcal{D}$ such that:

- T takes objects of $\mathcal{C}$ to objects of $\mathcal{D}$.
- If $f : A \to B$ in $\mathcal{C}$, then $T(f) : T(A) \to T(B)$ in $\mathcal{D}$. (This is the covariance property: arrows point the same way.)

$$\begin{tikzcd} A \arrow[d, "f"] & & T(A) \arrow[d, "T(f)"] \\ B & \xrightarrow{T} & T(B) \end{tikzcd}$$

- Identity morphisms map to identity morphisms: for all objects A of $\mathcal{C}$, $T(1_A) = 1_{T(A)}$. A self-loop in $\mathcal{C}$ must map to a self-loop in $\mathcal{D}$.
- Composition is respected: $T(g \circ f) = T(g) \circ T(f)$.

Definition 4.17. A **contravariant functor** T is almost the same, except for the direction of arrows. It is a map from $\mathcal{C}$ to $\mathcal{D}$ such that:

- T takes objects of $\mathcal{C}$ to objects of $\mathcal{D}$.
- If $f : A \to B$ in $\mathcal{C}$, then $T(f) : T(B) \to T(A)$ in $\mathcal{D}$. (This is the contravariance property: arrows point the other way.)

$$\begin{tikzcd} A \arrow[d, "f"] & & T(A) \\ B & \xrightarrow{T} & T(B) \arrow[u, "T(f)"'] \end{tikzcd}$$

- Identity morphisms map to identity morphisms: for all objects A of $\mathcal{C}$, $T(1_A) = 1_{T(A)}$. A self-loop in $\mathcal{C}$ must map to a self-loop in $\mathcal{D}$.
- Composition is respected (with arrows reversed): $T(g \circ f) = T(f) \circ T(g)$.

Notation 4.18. For a given functor T and a morphism f, sometimes one writes

$$\begin{aligned} f_* &= T(f), \quad \text{if } T \text{ is a covariant functor;} \\ f^* &= T(f), \quad \text{if } T \text{ is a contravariant functor.} \end{aligned}$$

Some simple examples are as follows.

Example 4.19. $\triangleright$ The **identity functor** on any category, sending objects to themselves and morphisms to themselves, is covariant. $\triangleleft$

Example 4.20. $\triangleright$ The **forgetful functor** is best illustrated by example. Take $\mathcal{C}$ to be the category of groups, where morphisms are group homomorphisms, and take $\mathcal{D}$ to be the category of sets, where morphisms are arbitrary functions. For any groups G and H and any homomorphism $\phi : G \to H$, send G and H to the sets of elements of G and H, respectively. Have ϕ be the same map sending elements of G to elements of H, and simply forget the homomorphism ploperty. We also have a forgetful functor from vector spaces to abelian groups, which forgets about scalar multiplication. The forgetful functor is covariant. $\triangleleft$

Example 4.21. $\triangleright$ Let the category A be the poset consisting of the elements 1, 2, and 3 as shown. There are arrows (morphisms) from 1 to 3 and 2 to 3, along with the self-loops from 1 to 1, 2 to 2, and 3 to 3. Next to that category in the figure are two possible images of the category.

The first, say a functor named f, sends 3 to c and collapses 1 and 2 into a. This functor is covariant since the arrows $1 \to 3$ and $2 \to 3$ are mapped to the (identical) arrows $f(1) \to f(3)$ and $f(2) \to f(3)$, which are both $a \to c$.

The second, say a functor named g, sends 1 to a and collapses 2 and 3 into c. This functor is contravariant since the arrow $1 \to 3$ is mapped to $g(3) \to g(1)$, which is $b \to a$, while the other non-identity arrow $2 \to 3$ is mapped to the self-loop $b \to b$.

$$\begin{tikzcd}
\mathbf{A} & & 3 & & \mathbf{B} & c & & \mathbf{C} & & b \arrow[dl] \\
1 \arrow[ur] & & & 2 \arrow[ul] & & a \arrow[u] & & & a &
\end{tikzcd}$$

◁

4.2.4 Hom functors

[xxx perhaps exclude this section]

Note: Perhaps this section introduces more abstraction than is worthwhile.

Here I want to formalize a familiar concept. This will be revisited in sections 4.6.8 and 4.7.8. First, I will consider vector spaces as a motivating example, followed by categories in general.

Fix a vector space V. For any other vector space W, the hom set $\mathrm{Hom}(V, W)$ is what we call $\mathcal{L}(V, W)$, namely, the set of all linear transformations from V to W. If we have a linear transformation f from vector spaces W_1 to W_2, and a linear transformation g from V to W_1, then we can compose f and g to get a linear transformation $f \circ g$ from V to W_2:

$$\begin{tikzcd}
V \arrow[r, "g"] \arrow[dr, "f \circ g"'] & W_1 \arrow[d, "f"] \\
& W_2
\end{tikzcd}$$

We can use this to make a functor as follows:

- It will be a functor from the category of vector spaces to the category of sets. We have to say what it does to objects (vector spaces) and morphisms (linear transformations).
- A vector space W is sent to $\mathcal{L}(V, W)$. That is, a given W is sent to the collection of all linear transformations from the fixed vector space V to that W.
- A linear transformation $f : W_1 \to W_2$ is sent to a function which converts a linear transformation $g : V \to W_1$ into $f \circ g : V \to W_2$. That is,

$$(f : W_1 \to W_2) \mapsto f_* : \mathcal{L}(V, W_1) \to \mathcal{L}(V, W_2) \quad \text{where} \quad f_*(g : V \to W_1) = (f \circ g : V \to W_2)$$

This functor is written

$$\mathrm{Hom}(V, \_).$$

It looks like this:

$$\begin{tikzcd} W_1 \arrow[d, "f"] & & \mathcal{L}(V, W_1) \arrow[d, "f_*"] \\ W_2 & \xrightarrow{\mathrm{Hom}(V, \_)} & \mathcal{L}(V, W_2) \end{tikzcd}$$

Since it sends W_1 to $\mathcal{L}(V, W_1)$, W_2 to $\mathcal{L}(V, W_2)$, and $f : W_1 \to W_2$ to $f_* : \mathcal{L}(V, W_1) \to \mathcal{L}(V, W_2)$, i.e. it preserves the direction of the arrows, $\mathrm{Hom}(V, \_)$ is said to be a **covariant hom functor**.

* * *

Since we have a covariant hom functor, it is natural to ask about a contravariant hom functor. Now fix a vector space W. If we have a linear transformation f from vector spaces V_1 to V_2, and a linear transformation g from V_2 to W, then we can compose g and f to get a linear transformation $g \circ f$ from V_1 to W:

$$\begin{tikzcd} V_1 \arrow[d, "f"'] \arrow[dr, "g \circ f"] & \\ V_2 \arrow[r, "g"] & W \end{tikzcd}$$

This time, the functor associated to W is written

$$\mathrm{Hom}(\_, W).$$

It behaves as follows:

- A vector space V is sent to $\mathcal{L}(V, W)$. That is, a given V is sent to the collection of all linear transformations from that V to the fixed vector space W.
- A linear transformation $f : V_1 \to V_2$ is sent to a function which converts a linear transformation $g : V_2 \to W$ into $g \circ f : V_1 \to W$. That is,

$$(f : V_1 \to V_2) \mapsto f^* : \mathcal{L}(V_2, W) \to \mathcal{L}(V_1, W) \quad \text{where} \quad f^*(g : V_2 \to W) = (g \circ f : V_1 \to W)$$

It looks like this:

$$\begin{tikzcd} V_1 \arrow[d, "f"] & & \mathcal{L}(V_1, W) \\ V_2 & \xrightarrow{\mathrm{Hom}(\_, W)} & \mathcal{L}(V_2, W) \arrow[u, "f^*"'] \end{tikzcd}$$

Since $\mathrm{Hom}(\_, W)$ sends V_1 to $\mathcal{L}(V_1, W)$, V_2 to $\mathcal{L}(V_2, W)$, and $f : V_1 \to V_2$ to $f^* : \mathcal{L}(V_2, W) \to \mathcal{L}(V_1, W)$, i.e. it reverses the direction of the arrows, $\mathrm{Hom}(\_, W)$ is said to be a **contravariant hom functor**.

Now, here I used vector spaces as a concrete example, since they are familiar to all of us. However, the same goes for categories in general: given an object A of a category $\mathcal{C}$, we have the covariant hom functor $\mathrm{Hom}(A, \_)$; given an object B of $\mathcal{C}$, we have the contravariant hom functor $\mathrm{Hom}(\_, B)$.

4.3 Freeness

Everything that is really great and inspiring is created by the individual who can labor in freedom.
— Albert Einstein (1879-1955).

The concept of freeness is central in algebra. The formal definition — which is undeniably necessarily for mathematical rigor — obscures the simple intuition. Thus, in this section I focus primarily on the intuition.

4.3.1 Definitions

Definition 4.22. Let $\mathcal{C}$ be a category. An object F in $\mathcal{C}$ is **free** with respect to a set S and with respect to a morphism i from S to F if for any other object A of $\mathcal{C}$ and a morphism $f : S \to A$ there is a unique morphism h from F to A such that $h \circ i = f$.

This sounds too abstract to be of any use. To rescue it, here is a commutative diagram:

$$\begin{tikzcd} S \arrow[r, "i"] \arrow[dr, "f"'] & F \arrow[d, "h"] \\ & A \end{tikzcd}$$

This still looks overly abstract. What the definition is saying, though, makes sense when we consider the familiar example of vector spaces.

Example 4.23. $\triangleright$ Let β be a basis for a finite-dimensional vector space V. Then i is (as is always the case) the inclusion map. Let f be a transformation sending the basis vectors to another vector space W. The existence and uniquess of h simply means that when we specify the images of the basis vectors under f, there is *no more choice* in where to send any other element of V. This is precisely what we mean when we write a linear transformation as a matrix: we are specifying the images of the *basis* vectors, and that gives us the image of *any* vector in V by linearity. $\triangleleft$

This example motivates the following intuitive notion of freeness.

Intuition 4.24. An object F (group, abelian group, module, etc.) is **free** on a generating set S (called the **generating set** or **basis**) if, for a morphism (homomorphism, linear transformation, etc.), specifying the image of the basis elements uniquely specifies the image of all of F.

For the formal definition of generators and relations, see [**Grove**]. Here, the following suffices.

Intuition 4.25. In any algebraic structure with an identity (group, abelian group, module, etc.) a **relation** between two or more elements is an expression which is equal to the identity element. (When there is more than one operation, e.g. in rings, we mean the additive identity.)

Example 4.26. $\triangleright$ In an abelian group G, for all $a, b \in G$, we have $aba^{-1}b^{-1} = 1$. The string $aba^{-1}b^{-1}$ is equal to the group identity, so it is a relation. $\triangleleft$

This gives us another way to think of freeness.

Intuition 4.27. An algebraic structure F (group, abelian group, module, etc.) is **free** on a generating set S if the elements of S have no relations other than the minimal ones necessary for the operations appropriate to the structure.

Example 4.28. $\triangleright$ For vector spaces, this is the familiar notion of **linear independence**: a linear combination of basis vectors summing to zero (which is the additive identity of the vector space) is simply a relation. When the generating set is linearly independent, any linear combination of basis vectors summing to zero has all coefficients zero, which is the trivial relation $0 = 0$. $\triangleleft$

More examples of free objects appear in the following sections.

4.3.2 Free groups

We will need free groups in algebraic topology. First, we have the definition of free semigroup. This is not needed for the rest of this paper (nor anywhere in the geometry-topology course) but it helps motivate the concept of free group. Semigroups were defined in section 4.1.1.

Intuition 4.29. A **free semigroup** on a set S is the smallest semigroup containing S in which there are no relations between the elements of S.

The canonical example is the same as the semigroup in example 4.2:

Example 4.30. $\triangleright$ Let $S = \{a, b, c\}$. Then sample elements of the free semigroup on S are the strings a, abb, $abbbbcccc$, $abab$, etc. No string simplifies into anything else. $\triangleleft$

Intuition 4.31. A **free group** F on a generating set S is the smallest group containing S along with all inverses of elements of S. As well, elements of F are completely distinct *except* that inverses cancel, and there must be an identity.

The formal construction of free groups, which makes this intuitive notion precise, always take some fuss. The shortest presentation I've seen is in [**Grove**], and even there it takes several pages. The simple concept behind that fuss, though, is that a free group F on a generating set S is as much like a free semigroup as possible: elements of the free group are strings composed of letters from S, with no simplification possible *except* for the cancelling of inverses. We may think of the empty string as the identity element.

Example 4.32. $\triangleright$ Let $S = \{a, b, c\}$. Then a sample element of the free group on S is

$$ababababccc.$$

Note however that

$$ac = abb^{-1}c = ab^2b^{-2}c.$$

$\triangleleft$

4.3.3 Free abelian groups

We need free abelian groups to construct tensor products, starting in section 4.7.1, and we will use them to help define integrals on manifolds in section 5.13. Free abelian groups are also used in free modules (coming next) which permeate this course.

Intuition 4.33. In the **free abelian group** on a generating set S, all elements of S are distinct as with a free group, but instead of forming strings of elements, we keep track of how many elements we have.

Example 4.34. $\triangleright$ Again take the generating set S to be $\{a, b, c\}$. When we form the free abelian group A on S, we have elements of the form

$$3a - 2b + c.$$

Another element might be $3a + 0b + c$ (which we abbreviate as $3a + c$), and their sum would be $6a - 2b + 2c$. The sum $a + a$ simplifies to $2a$, but the sum $a + b$ stays $a + b$: a and b are distinct; they are apples and oranges. All we can do is count how *many* apples and how many oranges, etc. $\triangleleft$

To make that last sentence a bit more precise, I note the following:

Remark 4.35. A free abelian group on a set S is isomorphic to $\mathbb{Z}^r$ where r is the cardinality of S. For each generator g in S there is one copy of $\mathbb{Z}$, which counts the number of g's. For example, the free abelian group on $\{a, b, c\}$ is isomorphic to $\mathbb{Z}^3$: $3a - 2b + c$ maps to the triple $(3, -2, 1)$.

4.3.4 Free product of groups

xxx stress intuition and examples, following the pattern above.

4.3.5 Free modules

Next we have the notion of free module, which generalizes the previous concept while also bringing us back to vector spaces.

Intuition 4.36. Let R be a commutative ring with identity. A **free R-module** M on a generating set S is the smallest module containing all R-linear combinations of elements of S.

Remark 4.37. A free $\mathbb{Z}$-module on a set S is the precisely the same thing as a free abelian group on S. (Similarly, recall from remark 4.8 that any abelian group is a $\mathbb{Z}$-module.)

We use this concept in two main ways for this course:

- We can start with a *small* set of elements which, much as in free groups and free abelian groups, are defined to be unmixable. We then take all R-linear combinations of elements of those basis elements, and we obtain a *big* module.
- We can start with a *big* R-module M, then look inside of it to find a *small* **basis** S, or set of generators. (We call the cardinality of S the dimension of M.) Every element of M is uniquely expressible as an R-linear combination of basis elements. This is precisely what we do with vector spaces. Every vector space is a free module, but we can also form free modules when R is not a field. We can think of a free module as being as close to a vector space as we can get when the base ring is not a field.

Example 4.38. $\triangleright$ Here is an example of the first notion, i.e. starting with a generating set and forming a free module on it. As in example 4.34, take S to consist of the letters a, b, c. For the ring R, this time use the rational numbers. Then elements of the free $\mathbb{Q}$-module on S are $\mathbb{Q}$-linear combinations of a, b, c, e.g.

$$\frac{3a}{4} - \frac{2c}{7}.$$

$\triangleleft$

Now for the second notion, i.e. finding a generating set inside a given module. Just as a free abelian group is isomorphic to $\mathbb{Z}^r$, where r is the cardinality of the generating set, a similar statement holds for free modules.

Remark 4.39. Let R be a commutative ring with identity. A free R-module M is isomorphic with R^m for some integer m. The integer m is called the **dimension** of M. (For each generator g of M there is one copy of R, which holds the coefficient on g.) We are familiar with this for vector spaces: we know from linear algebra that every finite-dimensional real vector space V is isomorphic to $\mathbb{R}^m$, where m is the size of any basis for V.

4.4 Quotients, projections, and sections

Definition 4.40. A function $f : X \to X$ from a set to itself is **idempotent** if $f(f(x)) = f(x)$ for all $x \in X$, i.e. if $f^2 = f$.

This simply means that when we apply a function twice, we don't get anything new, i.e. it acts like the identity on its image.

Definition 4.41. Let E and B be sets. A **projection** $\pi : E \to B$ is simply an epimorphism. If $E = B$, then we further require that π be idempotent. Usually E and B are something more than sets; we require that π be a homomorphism for whatever category E and B belong to (e.g. groups, rings, vector spaces).

The set E is called the **total space** and B is called the **base space**.

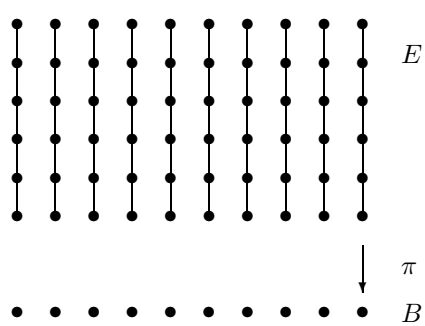
Definition 4.42. Let $\pi : E \to B$ be a projection. Let $b \in B$. The **fiber** of b is $\pi^{-1}(b)$.

This is thought of as the preimage of b, or everything **lying above** b.

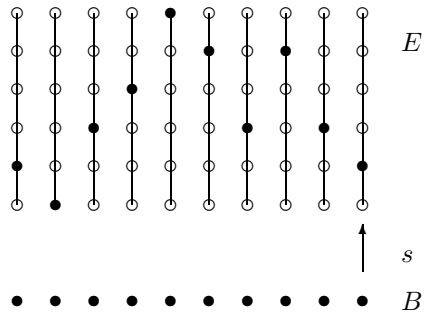
Definition 4.43. Let π be a projection from a set E to a set B. A **section** of π (or a **cross-section**) is a function $s : B \to E$ with the property that $\pi \circ s$ is the identity on B.

The $\pi \circ s = 1_B$ property simply means that a section maps elements of B somewhere into their *own* fibers, and not into any other elements' fibers.

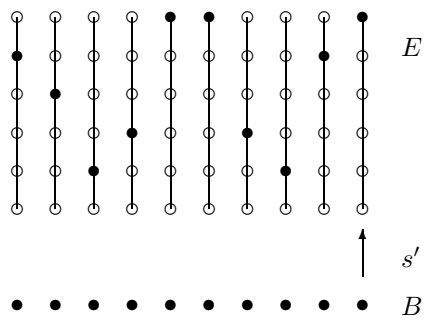
Here is a picture of sets E and B with a projection π. All the elements of E (black dots) are sent by π down to some element of B (also marked with black dots). For each b in B (for each dot on the bottom row) the elements of the fiber of b, i.e. $\pi^{-1}(b)$, are connected with a line.



Here is a picture of a section s for the same $\pi : E \to B$. Here, all the elements b of B are darkened, but only one element from each fiber of b is darkened. The dark dot in each fiber indicates $s(b)$. So, a section begins to look a lot like a **graph**, with B as the abscissa and fibers of B as the ordinate.



Here is another section of π:



For a given projection, there are in general many possible sections. This leads to the following definition:

Definition 4.44. Let $\pi : E \to B$ be a projection. Write $\Gamma(\pi, E, B)$ for the set of all sections, or **space of sections**, of π. (This is what Pickrell calls Ω^0; I choose to use a different symbol since Ω is used for other things. In fact, though, Γ is rather common in the literature. We can think of Γ for *graph*.)

Example 4.45. $\triangleright$ Let E be the ring of integers $\mathbb{Z}$ and let B be the integers mod n, written $\mathbb{Z}/n/Z$. Then π is the map which reduces mod n.

- We write the elements of $\mathbb{Z}/n\mathbb{Z}$ as $\{0 + \mathbb{Z}, 1 + \mathbb{Z}, 2 + \mathbb{Z}, \ldots, n - 1 + \mathbb{Z}\}$, or as $\{\overline{0}, \overline{1}, \overline{2}, \ldots, \overline{n-1}\}$.
- A section is a complete set of coset representatives. Usually, we take $\{0, 1, 2, \ldots, n-1\}$ to be that set. That is, $s(\overline{0}) = 0$, $s(\overline{1}) = 1$, etc. For example, with $n = 5$, a section of π is $\{0, 1, 2, 3, 4\}$. But we could just as well take another section to be $\{4, 8, 12, 16, 20\}$.
- The fiber above $\overline{k}$ is all numbers equivalent to k mod n. For example, with $n = 5$, the fiber above $\overline{2}$ is $\{\ldots, -8, -3, -2, 7, 12, 17, \ldots\}$.

$\triangleleft$

Example 4.46. $\triangleright$ Let E be the symmetric group $\mathcal{S}_n$ and let $B = \{\pm 1\}$. Let π be the sign function, which takes even permutations to 1 and odd permutations to -1.

- A section of π is any set containing one even permutation and one odd permutation.
- The fiber above -1 is the set of all odd permutations; the fiber above 1 is the set of all even permutations, which is the alternating group $\mathcal{A}_n$.

$\triangleleft$

Example 4.47. $\triangleright$ Let $\pi : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ be given by $\pi(x, y) = x$, i.e. π projects a point onto its x coordinate.

- A section is simply any real-valued function s: given x, $s(x)$ gives a point (x, y).
- The set of all sections of π is the set of all vector-to-scalar functions.
- If we further insist that $s(x)$ vary smoothly with x, then the set of all such sections is the set of all smooth vector-to-scalar functions.
- The fiber of $x_0 \in \mathbb{R}$ is the set of all (x, y) pairs with $x = x_0$, i.e. it is the vertical line through x_0.

$\triangleleft$

Example 4.48. $\triangleright$ Let $\pi : \mathbb{R}^3 \times \mathbb{R} \to \mathbb{R}$ be given by $\pi((x, y, z), u)) = u$. A section is simply any **vector-to-scalar function**: $f(x, y, z)$ gives a point $((x, y, z), u)$. The set of all sections of π is the set of all vector-to-scalar functions. Again, we may further insist that u vary smoothly with (x, y, z). $\triangleleft$

Example 4.49. $\triangleright$ Let $\pi : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3$ be given by $\pi((x, y, z), (u, v, w))) = (u, v, w)$. A section is any (smooth) **vector-to-vector function**: $\mathbf{F}(x, y, z)$ gives a point $((x, y, z), (u, v, w))$. $\triangleleft$

These examples recall the vector-calculus notions discussed in section 2.2.

xxx define quotient in two ways:

(1) If there is algebraic structure (e.g. groups, rings, vector spaces, etc.) then in terms of cosets. In particular, dimension of quotient is difference of quotients of originals.

(2) Else, A/B means contract B to a point and leave all else the same. Example: $\mathbb{RP}^n/\mathbb{RP}^m$. Present in terms of cell complexes. [Maybe this needs to be moved forward and xref'ed to. Or come up with a better example for here.] Here, the dimension-difference rule doesn't apply: e.g. $\mathbb{RP}^n/\mathbb{RP}^{n-1} = \mathbb{S}^n$ — not dimension 1.

4.5 Sequences, complexes, and homological algebra

In symbols one observes an advantage in discovery which is greatest when they express the exact nature of a thing briefly and, as it were, picture it; then indeed the labor of thought is wonderfully diminished.
— Gottfried Wilhelm Leibniz (1646-1716).

Here we have some fancy terminology for what (keep in mind!) are abstractions of down-to-earth concepts. So, when confused by the terminology, invoke your favorite example for concreteness. Several examples are presented in the following subsubsections.

4.5.1 Sequences

Throughout math we have maps from one object to another, e.g. a group homomorphism $G \to H$, a linear transformation $V \to W$, etc. But we can as well have a **sequence** of maps, e.g.

$$A \to B \to C$$

or

$$A_1 \to A_2 \to A_3 \to A_4$$

or perhaps an infinite sequence

$$\ldots \to A_1 \to A_2 \to A_3 \to A_4 \to \ldots$$

etc. We don't always name the maps, but we might, say, use the name $\phi_{i,i+1}$ for the map from the ith object to the $(i+1)$st object:

$$\ldots \to A_1 \stackrel{\phi_{12}}{\to} A_2 \stackrel{\phi_{23}}{\to} A_3 \to \ldots.$$

There's not much more we can say about such sequences of maps, as stated here. However, if we impose additional restrictions, we can obtain additional structure and additional results.

For any group homomorphism $\phi : G \to H$, the kernel and image of ϕ are always subgroups of G and H, respectively. Additionally, $\ker(\phi)$ is normal in G, but $\mathrm{im}(\phi)$ is not necessarily normal in H. Thus we can always form the quotient group $G/\ker(\phi)$, but we can't always form the quotient group $H/\mathrm{im}(\phi)$. If G and H are abelian, then *all* subgroups are normal, and we can always form these quotients. Note in particular that rings and vector spaces have abelian additive groups.

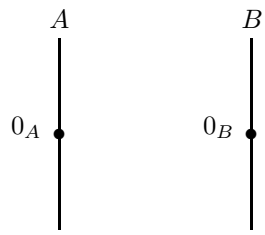
Suppose we have a sequence

$$A \stackrel{\phi}{\to} B \stackrel{\psi}{\to} C.$$

Then $\mathrm{im}(\phi)$ and $\ker(\psi)$ are both subobjects of B. We can ask about the relationship between the two, e.g. what additional results do we obtain if $\mathrm{im}(\phi) \subseteq \ker(\psi)$, $\mathrm{im}(\phi) = \ker(\psi)$, or $\mathrm{im}(\phi) \supseteq \ker(\psi)$? To help address this question, I will first define some notation.

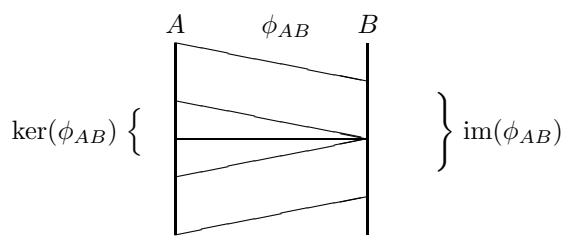
4.5.2 Trapezoid diagrams

Let A and B **objects** in a concrete category, with a homomorphism $\phi_{AB} : A \to B$. (For concreteness, perhaps think of them as being vector spaces.) We will draw a **trapezoid diagram** for this mapping. Each object has a zero element: 0_A and 0_B, respectively. Heavy vertical lines represent all the elements of the sets A and B, with the zero elements in the middle:



Suppose we have a **homomorphism** $\phi_{AB} : A \to B$. Since homomorphisms send 0 to 0, it makes sense to draw a line connecting the zero elements. Also, draw two nested trapezoids as follows:

- The outer trapezoid starts with all of A and ends inside of B. Its right-hand edge represents the **image** of ϕ_{AB}, which is at most all of B, but perhaps smaller.
- The inner trapezoid ends with 0_B. Its left-hand edge represents the **kernel** of ϕ_{AB}: all of the elements of A which map to 0_B. This is at least $\{0_A\}$, but perhaps bigger.
- By convention, I draw the outer-trapezoid lines parallel to the inner-trapezoid lines.



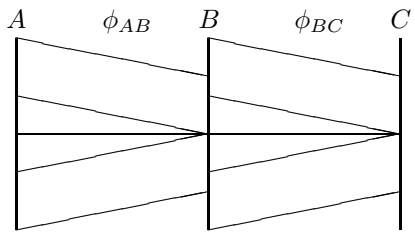
Consequences:

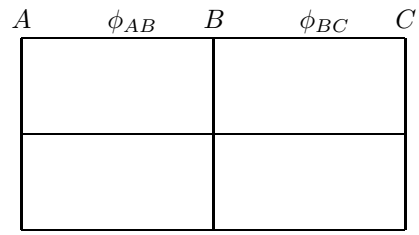
- If the inner trapezoid disappears into the center line, that means the map is injective (i.e. has zero kernel). Since I draw the outer-trapezoid lines parallel to the inner-trapezoid lines, one-to-one maps will have horizontal lines.
- If the outer and inner trapezoid both disappear into the center line, that means $A = \{0\}$.
- If outer trapezoid reaches all of B, then ϕ_{AB} is surjective.
- If the inner trapezoid disappears into the center line, and if the outer trapezoid reaches all of B, then ϕ_{AB} is an isomorphism.

4.5.3 Non-complex, non-exact sequences

Here are two sequences in which $\mathrm{im}(\phi_{AB}) \supseteq \ker(\phi_{BC})$. (For a specific example of the latter, suppose $A = B = C$ and $\phi_{AB} = \phi_{BC}$ are the identity map.)

[xxx include specific examples: use $\mathbb{Z}/m\mathbb{Z}$, and $\mathbb{R}^n$.]



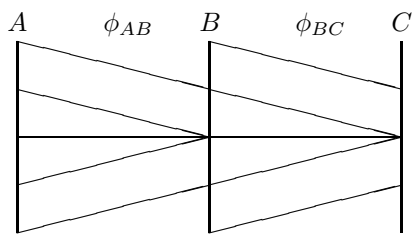


For the right-hand sequence, any non-zero element of A may be pushed through the sequence, all the way to the right, with the result still being non-zero. (This would still be true for a longer sequence of identity maps.) For the left-hand sequence, there are also some elements of A which map to non-zero elements of C.

We don't have a term for these kinds of sequences. However, do note that many common maps (including isomorphisms!) are of this form. Exact sequences and complexes (to be discussed below) deal with the cases $\mathrm{im}(\phi_{AB}) = \ker(\phi_{BC})$ and $\mathrm{im}(\phi_{AB}) \subseteq \ker(\phi_{BC})$, which are more special.

4.5.4 Exact sequences

Here is a sequence in which $\mathrm{im}(\phi_{AB}) = \ker(\phi_{BC})$:



Definition 4.50. Let $\phi_{AB} : A \to B$ and $\phi_{BC} : B \to C$ be homomorphisms of concrete categories. We say the sequence

$$A \xrightarrow{\phi_{AB}} B \xrightarrow{\phi_{BC}} C$$

is **exact at** B if

$$\mathrm{im}(\phi_{AB}) = \ker(\phi_{BC}).$$

xxx specific examples

now define exact sequences.

4.5.5 Complexes

Here is a sequence in which $\mathrm{im}(\phi_{AB}) \subseteq \ker(\phi_{BC})$:

4.5.6 [xxx merge w/ the above] Exact sequences

Definition 4.51. A sequence

$$\ldots \to A_1 \xrightarrow{\phi_{12}} A_2 \xrightarrow{\phi_{23}} A_3 \to \ldots$$

is said to be **exact** if $\mathrm{im}(\phi_{i-1,i}) = \ker(\phi_{i,i+1})$ for all i.

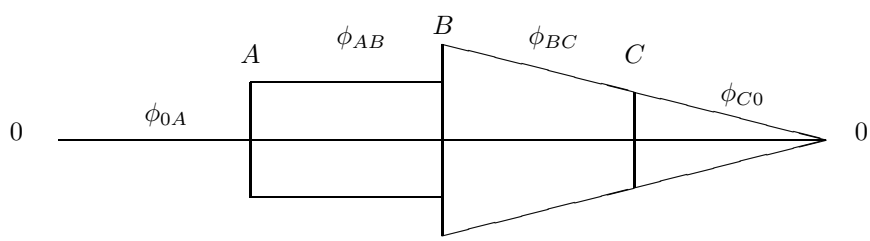
Exact sequences have an important special case:

Definition 4.52. A sequence

$$0 \to A \xrightarrow{\phi} B \xrightarrow{\psi} C \to 0$$

is said to be **short exact**.

Here is a graphical depiction of a short exact sequence:



In this diagram, the heavy vertical lines represent A, B, and C. At the middle of the vertical lines are the zero elements of A, B, and C, respectively. [xxx finish describing it.]

Remark 4.53. Since the image of the (unnamed) map from 0 to A is 0 (it can't be anything bigger!), and since that must be the same as $\ker(\phi)$, this means that ϕ is necessarily **injective**. Likewise, since the kernel of the (unnamed) map from C to 0 is all of C (it can't be anything smaller!), and since this is the same as the image of ψ, ψ is necessarily **surjective**. Thus we can also think of short exact sequences as follows: (1) ϕ is 1-1; (2) ψ is onto; (3) $\mathrm{im}(\phi) = \ker(\psi)$.

Example 4.54. $\triangleright$ Consider the sequence

$$0 \to \mathbb{Z} \xrightarrow{\phi} \mathbb{Z} \xrightarrow{\psi} \mathbb{Z}/2\mathbb{Z} \to 0$$

where ϕ is the multiplication-by-2 map, $a \mapsto 2a$, and ψ is the reduction-mod-2 map. The quotient $\mathbb{Z}/2\mathbb{Z}$ is just the two-element set of evens and odds. Multiplying by 2 is 1-1, and reducing mod 2 is surjective. Lastly, the image of ϕ is the even integers, which is precisely the kernel of ψ. (Also note that if we take all integers and double them, then reduce them mod 2, we get only evens, which are the zero element of $\mathbb{Z}/2\mathbb{Z}$.) $\triangleleft$

Remark 4.55. Given objects A and B where the quotient A/B is defined (for groups we need A to be a normal subgroup of B; for rings we need A to be an ideal of B; for modules we need A to be a submodule of B; etc.), we can always make an exact sequence

$$0 \to A \to B \to A/B \to 0$$

where the left arrow is the inclusion of A in B, and the right arrow is the reduction of B mod A.

Definition 4.56. A **long exact sequence** is one which is indexed by the natural numbers or the integers, i.e. one which is countably infinite in one or both directions. Note however that we can make any exact sequence countably infinite by putting zeroes on either side.

4.5.7 Techniques for exact sequences

Proposition 4.57. *We have*

$$0 \to A \to 0 \qquad \Longrightarrow \qquad A = 0.$$

Proof. Let $a \in A$. Then the right-hand map must send a to 0, i.e. a is in the kernel of the right-hand map. But the kernel of the right-hand map is equal to the image of the left-hand map, which is 0, so $a = 0$. □

Proposition 4.58. *We have*

$$0 \to A \xrightarrow{\phi} B \to 0 \qquad \Longrightarrow \qquad A \cong B.$$

Proof. The 0 on the left means that ϕ is 1-1, and the 0 on the right means that ϕ is onto. Therefore ϕ is a bijection. Since ϕ is also a homomorphism, ϕ is an isomorphism. □

Ways to **disconnect a long exact sequence**:

(i) Show that an object is zero. Then, its in-arrow must be surjective, and its out-arrow must be injective. For example, if we have

$$\ldots \to A \xrightarrow{\phi} B \xrightarrow{\psi} C \to \ldots$$

and if B is 0, then we obtain two long exact sequences:

$$\begin{array}{ll} \ldots \to A \xrightarrow{\phi} & 0 \\ & 0 \xrightarrow{\psi} C \to \ldots \end{array}$$

(ii) Show that an arrow is the zero map. For example, if we have

$$\ldots \longrightarrow A \xrightarrow{\phi_{AB}} B \xrightarrow{\phi_{BC}} C \xrightarrow{\phi_{CD}} D \xrightarrow{\phi_{DE}} E \longrightarrow \ldots,$$

and if ϕ_{BC} is the zero map, then its image is 0. Then we obtain two long exact sequences:

$$\begin{array}{ll} \ldots \longrightarrow A \xrightarrow{\phi_{AB}} B \xrightarrow{\phi_{BC}} & 0 \\ & 0 \xrightarrow{\phi_{CD}} D \xrightarrow{\phi_{DE}} E \longrightarrow \ldots \end{array}$$

(iii) Use the alternating-sum formula from p. 425 of [**Spivak2**].

4.5.8 [xxx merge w/ above] Complexes

xxx discuss: what if im $\neq$ ker?

Definition 4.59. complex: $\phi^2 = 0$. Show this pictorially via trapezoid diagrams.

note exact implies complex. prove with trapezoid diagrams.

note im $\supseteq$ ker implies non-complex. prove with trapezoid diagrams.

example with properly complex.

examples: $\mathbb{Z}$, times 2, mod 2. $\mathbb{Z}$, times 2, mod 10. $\mathbb{Z}$, times 10, mod 2. compute the homology when appropriate.

then, note we can take ker /im.

4.5.9 Zigzag lemma

xxx better to prove this with or without chains? If with, then move this section.

4.6 Linear algebra

Halmos and I share a philosophy about linear algebra: we think basis-free, we write basis-free, but when the chips are down we close the office door and compute with matrices like fury. — Irving Kaplansky.

Everything said in this section is applicable to finite-dimensional vector spaces over an arbitrary field. However, for this paper, the base field for all vector spaces is $\mathbb{R}$.

4.6.1 Notation

Given m vectors $\mathbf{v}_1, \ldots, \mathbf{v}_m$ in a vector space V, written in coordinates with respect to a basis, I will write

$$\left(\begin{array}{c} \mathbf{v}_1 \\ \hline \vdots \\ \hline \mathbf{v}_m \end{array}\right) \qquad \text{and} \qquad \left(\mathbf{v}_1 \,\middle|\, \cdots \,\middle|\, \mathbf{v}_m\right)$$

for the matrix of row vectors and the matrix of column vectors, respectively. In the case that $\dim V = m$, both of these matrices are square.

The standard basis for $\mathbb{R}^m$ is written

$$\{\mathbf{e}_1, \ldots, \mathbf{e}_m\}.$$

An arbitrary basis (not necessarily standard) will usually be written with the letter b, e.g.

$$\{\mathbf{b}_1, \ldots, \mathbf{b}_m\}.$$

4.6.2 To be filed

Cramer's rule:

$$\mathrm{Adj}(A)A = \det(A)I.$$

Rank-nullity theorem: Let $T : V \to W$ with finite dimensions. Then

$$\dim(V) = \dim(\ker(T)) + \dim(\mathrm{im}(T)).$$

4.6.3 Basis and dual space

Let V be a finite-dimensional vector space V over $\mathbb{R}$, say of dimension m. Then V is isomorphic with $\mathbb{R}^m$, which has many bases. In particular, it has a **standard basis**

$$\{\mathbf{e}_1, \ldots, \mathbf{e}_m\}$$

with elements of the form

$$\mathbf{e}_1 = (1, 0, 0, \ldots, 0), \mathbf{e}_2 = (0, 1, 0, \ldots, 0), \ldots, \mathbf{e}_m = (0, 0, 0, \ldots, 1).$$

Definition 4.60. The **dual space** of V is the set V^* of all linear transformations from V to $\mathbb{R}$.

Definition 4.61. An element of a dual space is called a **linear functional**.

Recall that V^* is a vector space over $\mathbb{R}$. It has the same dimension, namely m, as V. To prove this, we can choose a basis for V, do the following construction to obtain a basis for V^*, then use those bases to construct an isomorphism between V and V^*.

Definition 4.62. For any basis $\{\mathbf{b}_1, \ldots, \mathbf{b}_m\}$, there is a basis for the dual space (i.e. a **dual basis**)

$$\{\mathbf{b}_1^*, \ldots, \mathbf{b}_m^*\}$$

where the $\mathbf{b}_i^*$ functions are defined to have value

$$\mathbf{b}_i^*(\mathbf{b}_j) = \delta_{ij}$$

where the δ is Kronecker's, i.e. $\delta_{ij} = 1$ when $i = j$, 0 otherwise.

(See section 4.6.6 for a technique to compute the coefficients of the $\mathbf{b}_i^*$'s.)

In particular, the standard basis has a **standard dual basis**

$$\{\mathbf{e}_1^*, \ldots, \mathbf{e}_m^*\} \text{ with } \mathbf{e}_i^*(\mathbf{e}_j) = \delta_{ij}.$$

To construct an explicit isomorphism between V and V^*, we can map each $\mathbf{e}_i$ to the corresponding $\mathbf{e}_i^*$. The proof that this map is in fact an isomorphism is a (brief) standard exercise in linear algebra.

Example 4.63. $\triangleright$ Let $\mathbf{v} = (v_1, \ldots, v_m) \in \mathbb{R}^m$. Then $\mathbf{e}_2^*(\mathbf{v}) = v_2$, i.e. $\mathbf{e}_2^*$ extracts the second coordinate of $\mathbf{v}$. That is, the standard dual basis may be thought of as consisting of **coordinate-selector functions**. $\triangleleft$

Let $\lambda : V \to \mathbb{R}$ and $\mathbf{v} \in V$. Given the above basis, we may write

$$\lambda = \sum_{i=1}^{m} \ell_i \mathbf{e}_i^* \quad \text{and} \quad \mathbf{v} = \sum_{j=1}^{m} v_j \mathbf{e}_j.$$

Since the e_i^*'s are linear functions, and since $\mathbf{e}_i^*(\mathbf{e}_j) = \delta_{ij}$, we have

$$\lambda(\mathbf{v}) = \sum_{i=1}^{m} \ell_i \mathbf{e}_i^* \left(\sum_{j=1}^{m} v_j \mathbf{e}_j \right) = \sum_{i=1}^{m} \sum_{j=1}^{m} \ell_i v_j \mathbf{e}_i^*(\mathbf{e}_j) = \sum_{i=1}^{m} \ell_i v_i.$$

4.6.4 Geometric interpretation of dual

We can graph column vectors easily enough. How, though, do we graph functionals? We will see that functionals are row vectors, then define the notions of *spine* and *contour lines*.

First, a representation theorem: Every linear functional on a finite-dimensional vector space corresponds to a unique row vector. More precisely:

Proposition 4.64. *Let V be a finite-dimensional vector space. For all $\lambda \in V^*$ there is a unique $\mathbf{u} \in V$ such that for all $\mathbf{v} \in V$, $\lambda(\mathbf{v}) = \mathbf{u} \cdot \mathbf{v}$.*

Proof. The proof is by construction. Since V is finite-dimensional, it has a basis $\{\mathbf{b}_1, \ldots, \mathbf{b}_m\}$. Then $\mathbf{v}$ is written uniquely (which will make $\mathbf{u}$ unique below) as a linear combination of the $\mathbf{b}_i$'s. Since λ is a linear functional, we can move it through the sum, and since V is finite-dimensional there is no question as to the convergence of the sum:

$$\lambda(\mathbf{v}) \quad = \quad \lambda \left(\sum_{i=1}^{m} v_i \mathbf{b}_i \right) = \sum_{i=1}^{m} v_i \lambda(\mathbf{b}_i).$$

Simply put

$$u_i = \lambda(\mathbf{b}_i)$$

for each i from 1 to m. Then

$$\mathbf{u} \cdot \mathbf{v} \quad = \quad \sum_{i=1}^{m} u_i v_i = \sum_{i=1}^{m} v_i \lambda(\mathbf{b}_i) = \lambda(\mathbf{v}).$$

□

xxx xref to geom section; linear functionals.

xxx fishbone plots.

xxx define **spine**

xxx xref back to inner-product stuff in the very first few sections.

4.6.5 Dual and double dual

For finite-dimensional vector spaces V, we saw in section 4.6.3 that $V \cong V^*$. But since V^* is also a finite-dimensional vector space, we also have $V^* \cong V^{**}$. Since isomorphisms are transitive, $V \cong V^*$ and $V^* \cong V^{**}$ imply $V \cong V^{**}$. The space V^{**} is called the **double dual** of V.

However, more can be said. In section 4.6.3 we obtained a basis for V^* by first choosing a basis for V. To construct an isomorphism between V and V^{**}, though, we don't need to choose a basis. For each vector $\mathbf{v} \in V$ and each $\lambda \in V^*$, we can think of $\mathbf{v}$ operating on λ by

$$\mathbf{v}(\lambda) = \lambda(\mathbf{v}).$$

This makes sense because λ is in V^*, i.e. it is a linear function from V to $\mathbb{R}$. Thus $\lambda(\mathbf{v})$ is in $\mathbb{R}$, so we've taken $\lambda \in V^*$ and used $\mathbf{v}$ to obtain a real number, which is precisely what it means for something to be in V^{**}: v maps from V^* to $\mathbb{R}$. (The details of the proof of the isomorphism are a quick linear-algebra exercise. This leads to a **natural isomorphism** between V and V^{**}, in the sense that it works the same regardless of which basis we choose for V.)

The key point is that while we think of vectors V as **data**, they also can be thought of as **functions** as well, operating on the elements of the dual space.

4.6.6 Explicit computation of dual basis

A linear functional $\lambda \in V^*$ is a linear transformation from V to $\mathbb{R}$. With respect to a basis, say the standard basis, we can think of a general linear transformation from an m-dimensional space to an n-dimensional space as an $n \times m$ matrix. So, this means λ can be represented as a $1 \times m$ matrix. For example:

$$\lambda(\mathbf{v}) = \begin{pmatrix} a_1 a_2 \cdots a_m \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_m \end{pmatrix}.$$

Thus, a linear functional may be explicitly represented by a row vector. In section 4.6.3 we saw that each $\lambda \in V^*$ is a linear combination of the $\mathbf{b}_i^*$'s, the basis functionals corresponding to the basis vectors $\mathbf{b}_j$'s,

where $\mathbf{b}_i^*(\mathbf{b}_j) = \delta_{ij}$. So, we immediately ask, given a basis of $\mathbf{b}_j$'s, how do we compute the coefficients of the $\mathbf{b}_i^*$'s?

Recall the notation for row and column vectors shown in section 4.6.1. Use the fact that $\mathbf{b}_i^*(\mathbf{b}_j) = \delta_{ij}$. Now, $\mathbf{b}_i^*(\mathbf{b}_j)$ is the product of the $1 \times m$ array $\mathbf{b}_i^*$ and the $m \times 1$ array $\mathbf{b}_j$. Matrix multiplication is nothing more than a collection of such products: i.e. given any matrices A and B, the ijth cell of $C = AB$ is the dot product of the ith row of A with the j column of B. For example, in the product of two 4×4 matrices, the $2, 3$ cell of the product is the dot of the 2nd row of the first matrix with the 3rd column of the second matrix:

$$\begin{pmatrix} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \circ & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{pmatrix} = \begin{pmatrix} \cdot & \cdot & \cdot & \cdot \\ \circ & \circ & \circ & \circ \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{pmatrix} \begin{pmatrix} \cdot & \cdot & \circ & \cdot \\ \cdot & \cdot & \circ & \cdot \\ \cdot & \cdot & \circ & \cdot \\ \cdot & \cdot & \circ & \cdot \end{pmatrix}$$

Since the identity matrix I has δ_{ij} in its i, jth cell, this means that we can collect all the $\mathbf{b}_i^*(\mathbf{b}_j)$'s into a matrix product:

$$I = \begin{pmatrix} \mathbf{b}_1^* \\ \hline \vdots \\ \hline \mathbf{b}_m^* \end{pmatrix} \left(\mathbf{b}_1 \,\middle|\, \cdots \,\middle|\, \mathbf{b}_m\right) \quad \Longrightarrow \quad \begin{pmatrix} \mathbf{b}_1^* \\ \hline \vdots \\ \hline \mathbf{b}_m^* \end{pmatrix} = \left(\mathbf{b}_1 \,\middle|\, \cdots \,\middle|\, \mathbf{b}_m\right)^{-1}.$$

See section 4.6.7 for an example. Also see section 4.7.14, for change of coordinates for tensors.

[xxx perhaps note: If the basis is orthonormal then the primaries and their respective dual spines coincide. If it is merely orthogonal then they are collinear with reciprocal lengths. Else they are skewy.]

4.6.7 Change-of-basis matrices; covariance and contravariance

Here, by means of a carefully worked example, I relate the new (and important) concepts of **covariance** and **contravariance** to the familiar language of **change-of-basis matrices**. I also show the geometric significance of change of basis for dual vectors.

Let

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

as usual, and let

$$\mathbf{b}_1 = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \quad \text{and} \quad \mathbf{b}_2 = \begin{pmatrix} 1 \\ 3 \end{pmatrix}.$$

I want to see what happens to $\mathbf{b}_1^*$ and $\mathbf{b}_2^*$ when we change from $\{\mathbf{e}_1, \mathbf{e}_2\}$ coordinates to $\{\mathbf{b}_1, \mathbf{b}_2\}$ coordinates. Using the reasoning in section 4.6.6, I have

$$\begin{pmatrix} \mathbf{b}_1^* \\ \hline \mathbf{b}_2^* \end{pmatrix} = \left(\mathbf{b}_1 \,\middle|\, \mathbf{b}_2\right)^{-1} = \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix}^{-1} = \begin{pmatrix} 3/5 & -1/5 \\ -1/5 & 2/5 \end{pmatrix}$$

so

$$\mathbf{b}_1^* = \begin{pmatrix} 3/5 & -1/5 \end{pmatrix} \quad \text{and} \quad \mathbf{b}_2^* = \begin{pmatrix} -1/5 & 2/5 \end{pmatrix}.$$

(Note that these are not simply the transpose of $\mathbf{b}_1$ and $\mathbf{b}_2$. Hence my use of $\mathbf{b}_i^*$ rather than $\mathbf{b}_i^t$, as mentioned in section 1.1.)

Now I want to look at the change-of-basis matrix. Since $\mathbf{e}_1$ maps to $\mathbf{b}_1$ and $\mathbf{e}_2$ maps to $\mathbf{b}_2$, we write down the matrix

$$Q = \left(\mathbf{b}_1 \,\middle|\, \mathbf{b}_2\right) = \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix}.$$

This sends the original basis vectors to the new basis vectors:

$$Q\left(\mathbf{e}_1 \,\middle|\, \mathbf{e}_2\right) = QI = \left(\mathbf{b}_1 \,\middle|\, \mathbf{b}_2\right).$$

For the dual-basis vectors, we have

$$\left(\frac{\mathbf{b}_1^*}{\mathbf{b}_2^*}\right) = \left(\mathbf{b}_1 \,\middle|\, \mathbf{b}_2\right)^{-1} I = Q^{-1}\left(\frac{\mathbf{e}_1^*}{\mathbf{e}_2^*}\right)$$

or

$$\left(\frac{\mathbf{e}_1^*}{\mathbf{e}_2^*}\right) = Q\left(\frac{\mathbf{b}_1^*}{\mathbf{b}_2^*}\right).$$

In summary, we used the change-of-basis matrix Q as follows:

$$\begin{array}{ccc} \left(\mathbf{e}_1 \,\middle|\, \mathbf{e}_2\right) & \xrightarrow{Q} & \left(\mathbf{b}_1 \,\middle|\, \mathbf{b}_2\right) \\ \left(\frac{\mathbf{e}_1^*}{\mathbf{e}_2^*}\right) & \xleftarrow{Q} & \left(\frac{\mathbf{b}_1^*}{\mathbf{b}_2^*}\right). \end{array}$$

That is:

- The coefficients of a linear functional vary together with Q, i.e. **covariantly** with Q. (Now is as good a time as any to note that linear functionals are sometimes called **covectors**.)
- The coefficients of a vector vary opposite from Q, i.e. **contravariantly** with Q.

We will see more about this subject in sections 4.7.14 and [xxx write and xref — incl. $dx \neq dx$ section(s)].

* * *

So much for the symbolic manipulation. From a geometric point of view, why is it that the dual basis must transform in this way? Here is a plot of the basis vectors and the dual-basis vectors:

[xxx figure here]

I am using fishbone plots for the dual-basis vectors, as defined in section 4.6.4. Recall also the discussion in section 1.3.1. From the figure it's now clear that $\mathbf{b}_1^*$ has its spine pointing perpendicular to $\mathbf{b}_2$, and vice versa. Also, the magnitude of $\mathbf{b}_1^*$ is such that $\mathbf{b}_1^*(\mathbf{b}_1)$ is still 1.

4.6.8 Pullbacks

Now let $f : V \to W$ be a linear transformation, and let $\lambda : W \to \mathbb{R}$ be a linear functional, i.e. $\lambda \in W^*$. We can map from V to $\mathbb{R}$ by going through W first. That is, for $\mathbf{v} \in V$, $f(\mathbf{v})$ is in W, and so we can apply λ to it:

$$\begin{tikzcd} V \arrow[rr, "f"] \arrow[dr, "\lambda \circ f"'] & & W \arrow[dl, "\lambda"] \\ & \mathbb{R} & \end{tikzcd}$$

Formally, $\lambda \circ f : V \to \mathbb{R}$. Given $\lambda \in W^*$, we have obtained $\lambda \circ f \in V^*$. Post-composing λ by f is said to be a **pullback** of λ from W to V. Another way to look at this is that, given $f : V \to W$, we have a map from W^* to V^*:

$$\begin{aligned} f : &\quad V \to W \\ f^* : &\quad W^* \to V^* \\ f^*(\lambda) &= \lambda \circ f \\ (f^*\lambda)(\mathbf{v}) &= \lambda(f\mathbf{v}). \end{aligned}$$

In this way, f^* takes functionals to functionals:

$$\begin{tikzcd} V \arrow[r, "f"] & W \\ V^* & W^* \arrow[l, "f^*"'] \end{tikzcd}$$

There are two key points here:

(1) We have a **contravariant** functor from vector spaces to vector spaces. It is contravariant in the categorical sense, as in section 4.2.3. This is because the arrows are reversed. (See 4.2.4 or any of the cited algebra texts for more information; f^* is called a contravariant **hom functor**.)

(2) To compute the pullback $f^*\lambda$ of λ by f, just write $\lambda \circ f$.

4.6.9 Pairings

Definition 4.65. Let V and W be real vector spaces. A **pairing** is a bilinear function from $V \times W$ to $\mathbb{R}$. Given $\mathbf{v} \in V$ and $\mathbf{w} \in W$, we write

$$\langle \mathbf{v}, \mathbf{w} \rangle.$$

Example 4.66. $\triangleright$ The usual dot product is a pairing of V with itself:

$$\langle \mathbf{v}_1, \mathbf{v}_2 \rangle = \mathbf{v}_1 \cdot \mathbf{v}_2.$$

$\triangleleft$

Example 4.67. $\triangleright$ We can pair V and V^* via

$$\langle \lambda, \mathbf{v} \rangle = \lambda(\mathbf{v}).$$

$\triangleleft$

See section 7.2.2 for another example of a pairing.

Definition 4.68. Let V and W be paired vector spaces. For a linear transformation $T : V \to V$, a linear transformation $T^* : W \to W$ is the **adjoint** of T if

$$\langle T\mathbf{v}, \mathbf{w} \rangle = \langle v, T^*\mathbf{w} \rangle$$

for all $\mathbf{v} \in V$ and all $\mathbf{w} \in W$.

The adjoint exists and is unique [**FIS**].

Example 4.69. ▷ In particular, for the usual pairing of V and V^*, we have

$$\langle T\mathbf{v}, \lambda\rangle = \langle v, T^*\lambda\rangle = \langle \mathbf{v}, \lambda T^t\rangle.$$

[xxx elaborate.] ◁

4.6.10 Symmetric and skew-symmetric matrices

Definition 4.70. Let A be an $m \times m$ matrix. Then A is said to be **symmetric** if $a_{ij} = a_{ji}$ for all i, j from 1 to m. Likewise, A is said to be **skew-symmetric** if $a_{ij} = -a_{ji}$ for all i, j from 1 to m.

Remark 4.71. It is an easy exercise to show that an arbitrary $m \times m$ matrix A may be written as the sum of symmetric and skew-symmetric matrices. This is done explicitly by forming the **symmetrization** and **skew-symmetrization** of A, namely,

$$B = \frac{1}{2}\left(A + A^t\right), \quad C = \frac{1}{2}\left(A - A^t\right), \quad A = B + C.$$

These are appropriately named: we take the entries of A and make matrices which are, *by construction*, symmetric and skew-symmetric, respectively.

Example 4.72. ▷ Let

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}.$$

Then the symmetrization of A is

$$\frac{1}{2}\left[\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} + \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}\right] = \begin{pmatrix} 1 & 5/2 \\ 5/2 & 4 \end{pmatrix}$$

while the skew-symmetrization of A is

$$\frac{1}{2}\left[\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} - \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}\right] = \begin{pmatrix} 0 & -1/2 \\ 1/2 & 0 \end{pmatrix}.$$

◁

It is easy to show that the set of symmetric $m \times m$ matrices is a real vector space of dimension $m(m+1)/2$; the set of skew symmetric matrices is a real vector space of dimension $m(m-1)/2$; the set of all $m \times m$ matrices is a real vector space of dimension m^2, and is the direct sum of the former two spaces.

Let A be an $m \times m$ symmetric matrix with entries a_{ij}. Each a_{ij} must be the same as a_{ji}. Thus, to specify a symmetric matrix it suffices to specify the entries a_{ij} for $i \leq j$. There are $m(m+1)/2$ choices for values a_{ij} with $i \leq j$. Thus, for example, a basis for the vector space of symmetric 3×3 matrices is

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$
$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

Likewise, for the skew-symmetric matrices, all the a_{ii}'s must be 0; there are $m(m-1)/2$ choices for the remaining coefficients. We can specify the values of a skew-symmetric matrix by specifying the coefficients a_{ij} for $i < j$. A basis is

$$\begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}.$$

4.7 Tensors

Let us be patient! These severe afflictions
Not from the ground arise;
But oftentimes celestial benedictions
Assume this dark disguise.
— Henry Wadsworth Longfellow (1807-1882).

Tensors are encountered throughout math and physics. They are presented in at least four seemingly distinct guises:

(1) Tensor products from abstract algebra.

(2) Tensors as k-linear functions $T : V^k \to \mathbb{R}$.

(3) Tensors as k-dimensional arrays: a 0-tensor is related to a scalar, a 1-tensor is related to an array, a 2-tensor is related to a matrix, a 3-tensor is related to a 3-dimensional array, etc.

(4) Old-fashioned tensors from physics ("transform according to ...", e.g. [**PDM**]). (In this guise, tensors appear particularly foreign. You will see lots of superscripts, subscripts, Einstein summation, etc. Why bother? Well, this is the way tensors are usually viewed in applications, so this is the language that your scientific collaborators will be speaking.)

[xxx introduce a new section on the zeroth guise: geometry. Functional is $\mathbf{u}^*$. Bivector $\mathbf{u}\wedge\mathbf{v}$ is the equivalence class of all vector pairs coplanar with $\mathbf{u}$ and $\mathbf{v}$, with the same signed area. For the symmetric product, all pairs coplanar to $\mathbf{u}$ and $\mathbf{v}$ with the same inner product. Bifunctional $\mathbf{u}^* \wedge \mathbf{v}^*$ measures area spanned by two other vectors, projected onto the plane spanned by $\mathbf{u}$ and $\mathbf{v}$. Also equivalence classes using

$$(\lambda \wedge \mu)(\mathbf{u}, \mathbf{v}) = (\lambda \wedge \mu)(\mathbf{u} \wedge \mathbf{v}) = \det(\lambda|\mu) \det(\mathbf{u}|\mathbf{v}).$$

Generalizations to 3 and more dimensions. xref back frequently to the geometry section; xref forward to the axiomatic tensor definitions. Suggest that, although I don't know the history well enough, I *suspect* that it was precisely these geometric notions which led to the axioms we have today.]

These four guises are all equivalent, as we will see. Here, the first guise is taken to be the base definition. Note that [**Spivak2**] takes the second guise to be the base definition. Following [**Conlon**] and [**Abr**], I take the first guise to be the base definition; the second guise follows as a consequence as shown in section 4.7.4. There are two reasons for this: (1) we can draw on the abstract algebra that we possess as first-year graduate students in math; (2) we will be able to view vector fields as a special case of tensor fields. This will lead to a more elegant view of Lie derivatives later on.

4.7.1 Tensor products of vector spaces

The tensor product of modules (here, it suffices to consider tensor products of vector spaces) is discussed in any graduate algebra text, e.g. [**DF**], [**Grove**], [**Hungerford**], [**Lang**]. There is some amount of fuss about the existence and uniqueness of the tensor product; one loses the forest for the trees. The essence of the tensor product is very simple, as follows.

(1) Let V and W be vector spaces. We can form the **Cartesian product** $V \oplus W$. Elements of this product are of the form $(\mathbf{v}, \mathbf{w})$ for $\mathbf{v} \in V$ and $\mathbf{w} \in W$.

(2) Given any set we can form a **free abelian group** on that set, as discussed in section 4.3.5. That is, for the free abelian group on $V \oplus W$, all of the $(\mathbf{v}, \mathbf{w})$ pairs are distinct.

(3) The **tensor product** of V and W, written

$$V \otimes W,$$

with elements written

$$\mathbf{v} \otimes \mathbf{w}$$

for $\mathbf{v} \in V$ and $\mathbf{w} \in W$, is the free abelian group **modded out** by the following **relations** (see definition 4.25):

(a) For any $r \in \mathbb{R}$,

$$r\mathbf{v} \otimes \mathbf{w} = \mathbf{v} \otimes r\mathbf{w}.$$

(b) We can distribute:

$$\begin{aligned} (\mathbf{v}_1 + \mathbf{v}_2) \otimes \mathbf{w} &= \mathbf{v}_1 \otimes \mathbf{w} + \mathbf{v}_2 \otimes \mathbf{w} \\ \mathbf{v} \otimes (\mathbf{w}_1 + \mathbf{w}_2) &= \mathbf{v} \otimes \mathbf{w}_1 + \mathbf{v} \otimes \mathbf{w}_2. \end{aligned}$$

That is, all $(\mathbf{v}, \mathbf{w})$ pairs are distinct, *except* when we can do some of the above three operations to make them equal. (This is to say that all elements of a tensor product are equivalence classes, with equivalence given by the above relations.) In property (a), we permit ourselves to move *real* numbers back and forth. For this reason, in an algebra text, $V \otimes W$ would be written $V \otimes_{\mathbb{R}} W$.

Example 4.73. $\triangleright$ Let $V = W = \mathbb{R}^2$. Then

$$(2, 4) \otimes (3, 5) = (1, 2) \otimes (6, 10)$$

since we can move the 2 from the left-hand component to the right-hand component. $\triangleleft$

4.7.2 Explicit representation of the tensor product; tensor product as vector space

It is proved in the above-cited algebra texts (see [**DF**] section 10.4 for a particularly lucid discussion) that for vector spaces,

$$\mathbb{R}^m \otimes \mathbb{R}^n \cong \mathbb{R}^{mn}$$

as an isomorphism of vector spaces. What this means for us is that, given

$$\begin{aligned} \dim(V) = m, &\qquad \dim(W) = n \\ \mathbf{v} = (v_1, \ldots, v_m), &\qquad \mathbf{w} = (w_1, \ldots, w_n) \end{aligned}$$

where coefficients are taken with respect to the standard basis, we can think of $\mathbf{v} \otimes \mathbf{w}$ as the **outer product**

$$\begin{pmatrix} u_1 v_1 & \cdots & u_1 v_n \\ \vdots & & \vdots \\ u_m v_1 & \cdots & u_m v_n \end{pmatrix}$$

which is to say that $\mathbf{v} \otimes \mathbf{w}$ is an $m \times n$ **array** with i, jth entry

$$u_i v_j.$$

Example 4.74. ▷ Outer products may be done with pencil and paper as follows. Write $\mathbf{u}$ down a column and $\mathbf{v}$ across a row, then fill in products of $u_i v_j$. For example, let $\mathbf{u} = (1, 2)$ and $\mathbf{v} = (3, 4)$. Then we have

	3	4
1	3	4
2	6	8

◁

These are not matrices in the regular sense: such things are added componentwise just as for matrices, but we'll see in section 4.7.9 that they are multiplied using a different rule.

Of course, we can take the tensor product of several vector spaces. Sample elements of $U \otimes V \otimes W$ would be of the form $\mathbf{u} \otimes \mathbf{v} \otimes \mathbf{w}$, with i, jth entry $u_i v_j w_k$. For this course we are interested in V tensored with itself an arbitrary number of times. In fact, we have the following definitions:

Definition 4.75. Take V tensored with itself r times. This is called the **r-fold tensor product** of V.

Remark 4.76. Let $m = \dim(V)$. By the above isomorphism, $\dim(V \otimes V) = m^2$, and in general the r-fold tensor product of V with itself has dimension m^r over $\mathbb{R}$.

Notation will be introduced in section 4.7.6. We will write the r-fold tensor product of V with itself as $\mathcal{T}_0^r(V)$, where the subscript 0 will be explained in section 4.7.4.

Given the above isomorphism, we can think of $V \otimes W$ as a **vector space** over $\mathbb{R}$. Let $m = \dim V$, $n = \dim W$, and let

$$\{\mathbf{e}_1, \ldots, \mathbf{e}_m\}, \{\mathbf{f}_1, \ldots, \mathbf{f}_n\}$$

be the standard bases for V and W, respectively. Basis elements for $V \otimes W$ are of the form

$$\mathbf{e}_i \otimes \mathbf{f}_j$$

where $1 \leq i \leq m$ and $1 \leq j \leq n$.

Example 4.77. ▷ Let $V = \mathbb{R}^2$ and $W = \mathbb{R}^3$. Then we have bases

$$\{\mathbf{e}_1 = (1, 0), \mathbf{e}_2 = (0, 1)\}, \qquad \{\mathbf{f}_1 = (1, 0, 0), \mathbf{f}_2 = (0, 1, 0), \mathbf{f}_3 = (0, 0, 1)\}$$

for $\mathbb{R}^2$ and $\mathbb{R}^3$. Computing as in example 4.74, a basis for $\mathbb{R}^2 \otimes \mathbb{R}^3$ is

$$\mathbf{e}_1 \otimes \mathbf{f}_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \mathbf{e}_1 \otimes \mathbf{f}_2 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \mathbf{e}_1 \otimes \mathbf{f}_3 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix},$$
$$\mathbf{e}_2 \otimes \mathbf{f}_1 = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \mathbf{e}_2 \otimes \mathbf{f}_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \mathbf{e}_2 \otimes \mathbf{f}_3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

◁

Example 4.78. ▷ Let $V = W = \mathbb{R}^2$. In example 4.73 we saw that from the definition of the tensor product we could set

$$(2, 4) \otimes (3, 5) = (1, 2) \otimes (6, 10)$$

by moving the scalar 2 from one side to the other. What happens when we think of these as explicit outer products? The left-hand side is

$$(2, 4) \otimes (3, 5) = \begin{pmatrix} 2 \cdot 3 & 2 \cdot 5 \\ 4 \cdot 3 & 4 \cdot 5 \end{pmatrix} = \begin{pmatrix} 6 & 10 \\ 12 & 20 \end{pmatrix}$$

whereas the right-hand side is

$$(1,2)\otimes(6,10)=\begin{pmatrix}1\cdot 6 & 1\cdot 10\\ 2\cdot 6 & 2\cdot 10\end{pmatrix}=\begin{pmatrix}6 & 10\\ 12 & 20\end{pmatrix}.$$

This example drives home the fact that while the tensor product is associative, it is certainly not commutative. Also note that, in contrast to the representation used in example 4.73, arrays give us *unique* representations for tensors. ◁

4.7.3 Decomposability of tensors

In example 4.78 we wrote an element of $V\otimes V$ as the outer product of two vectors $\mathbf{u}$ and $\mathbf{v}$ in V. Given the vector-space representation, it is clear that we can add a pair of two-tensors componentwise, e.g. if $\mathbf{u}_1,\mathbf{u}_2,\mathbf{v}_1,\mathbf{v}_2\in V$, then

$$\mathbf{u}_1\otimes\mathbf{v}_1+\mathbf{u}_2\otimes\mathbf{v}_2\in V\otimes V.$$

When we do that, will the result be representable as the outer product of two vectors? Probably not. To see this, note the following: consider $\mathbf{u}\otimes\mathbf{v}$. With respect to the standard basis, we wrote this as an $m\times m$ array (where $m=\dim V$), with coefficients u_iv_j. Now, the ith row of $\mathbf{u}\otimes\mathbf{v}$ is just the vector $u_i\mathbf{v}$. So, $\mathbf{u}\otimes\mathbf{v}$ has row rank 1 (or 0, if $\mathbf{u}$ or $\mathbf{v}$ is zero) when viewed as a 2-dimensional array: every row is a multiple of the same vector $\mathbf{v}$. (Likewise, the jth column of $\mathbf{u}\otimes\mathbf{v}$ is $\mathbf{u}v_j$.) So, if an element ψ of $V\otimes V$ is of the form $\mathbf{u}\otimes\mathbf{v}$, then it will have row rank at most 1, with respect to any basis we choose. Likewise, if an element of $V\otimes V$ has row rank greater than 1, it could not be of the form $\mathbf{u}\otimes\mathbf{v}$. Now, we can certainly write down lots of $m\times m$ arrays which have row rank greater than 1. This means most elements of $V\otimes V$ aren't expressible in the form $\mathbf{u}\otimes\mathbf{v}$. This motivates the following definition.

Definition 4.79. A tensor $\psi\in\mathcal{T}^r(V)$ is said to be **simple** or **decomposable** if it can be written as a tensor product of r vectors in V.

Example 4.80. ▷ Let $V=\mathbb{R}^2$. With respect to the standard basis,

$$\alpha=\begin{pmatrix}1 & 2\\ 3 & 6\end{pmatrix}$$

is an element of $V\otimes V$ which is decomposable into

$$\alpha=(1,3)\otimes(1,2).$$

On the other hand,

$$\beta=\begin{pmatrix}1 & 2\\ 3 & 4\end{pmatrix}$$

is indecomposable. ◁

Note that using the vector-space representation as shown in section 4.7.2, while we cannot decompose an arbitrary tensor as a tensor product of vectors, we can write an arbitrary tensor as the *sum* of simple tensors. (We can do this because of the isomorphism in section 4.7.2.)

Example 4.81. ▷ Take β from the previous example. Proceeding as in example 4.77, write β with respect to the basis $\{\mathbf{e}_i\otimes\mathbf{f}_j\}$:

$$\begin{aligned}\beta &= \begin{pmatrix}1 & 2\\ 3 & 4\end{pmatrix}=1\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}+2\begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}+3\begin{pmatrix}0 & 0\\ 1 & 0\end{pmatrix}+4\begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix}\\ &= \mathbf{e}_1\otimes\mathbf{f}_1+2\mathbf{e}_1\otimes\mathbf{f}_2+3\mathbf{e}_2\otimes\mathbf{f}_1+4\mathbf{e}_2\otimes\mathbf{f}_2.\end{aligned}$$

◁

4.7.4 Tensor products of dual spaces; tensors as k-linear functions

In section 4.7.1 we saw how to form the tensor product $V \otimes W$ of vector spaces V and W, or the r-fold tensor product $\mathcal{T}_0^r(V)$. However, given a vector space V, the dual V^* is also a vector space so it makes sense to tensor V^* with itself. Here, however, we get an additional feature: we can think of such tensors as *functions*. We could write $V^* \otimes V^*$ as $\mathcal{T}_0^2(V^*)$, but we will instead write $V^* \otimes V^*$ as $\mathcal{T}_2^0(V)$. (This will be explained in section 4.7.6.) In general, the r-fold tensor product of V^* with itself will be written $\mathcal{T}_r^0(V)$.

Given linear functionals $\lambda, \mu : V \to \mathbb{R}$, we can certainly write down the tensor product $\lambda \otimes \mu$. But what can this mean as a function? Given $\mathbf{v}, \mathbf{w} \in V$, we think of this as

$$(\lambda \otimes \mu)(\mathbf{v} \otimes \mathbf{w}) = \lambda(\mathbf{v})\mu(\mathbf{w})$$

where we simply multiply the outputs $\lambda(\mathbf{v})$ and $\mu(\mathbf{w})$ in $\mathbb{R}$. This means that $V^* \otimes V^*$ consists of functions from $V \oplus V$ to $\mathbb{R}$.

However, more can be said about such functions: they are **bilinear**, i.e. they are linear in each slot if we hold the other slot fixed.

Example 4.82. $\triangleright$ For example, let $\mathbf{u}_1, \mathbf{u}_2, \mathbf{v}_1, \mathbf{v}_2 \in V$ and let $a, b \in \mathbb{R}$. Given $\lambda \otimes \mu$ as above,

$$\begin{aligned}
(\lambda \otimes \mu)(\mathbf{u}_1 + \mathbf{u}_2, \mathbf{v}) &= \lambda(\mathbf{u}_1 + \mathbf{u}_2)\mu(\mathbf{v}) \\
&= [\lambda(\mathbf{u}_1) + \lambda(\mathbf{u}_2)]\mu(\mathbf{v}) \\
&= \lambda(\mathbf{u}_1)\mu(\mathbf{v}) + \lambda(\mathbf{u}_2)\mu(\mathbf{v}) \\
&= (\lambda \otimes \mu)(\mathbf{u}_1, \mathbf{v}) + (\lambda \otimes \mu)(\mathbf{u}_2, \mathbf{v}).
\end{aligned}$$

Likewise, we have

$$\begin{aligned}
(\lambda \otimes \mu)(\mathbf{u}, \mathbf{v}_1 + \mathbf{v}_2) &= (\lambda \otimes \mu)(\mathbf{u}, \mathbf{v}_1) + \lambda(\mathbf{u}, \mathbf{v}_2) \\
(\lambda \otimes \mu)(a\mathbf{u}, \mathbf{v}) &= a(\lambda \otimes \mu)(\mathbf{u}, \mathbf{v}) = (\lambda \otimes \mu)(\mathbf{u}, a\mathbf{v})
\end{aligned}$$

These things work precisely because λ and μ are linear. As a consequence:

$$\begin{aligned}
\lambda(a\mathbf{u}, b\mathbf{v}) &= ab\lambda(\mathbf{u}, \mathbf{v}) \\
\lambda(\mathbf{u}_1 + \mathbf{u}_2, \mathbf{v}_1 + \mathbf{v}_2) &= \lambda(\mathbf{u}_1, \mathbf{v}_1) + \lambda(\mathbf{u}_1, \mathbf{v}_2) + \lambda(\mathbf{u}_2, \mathbf{v}_1) + \lambda(\mathbf{u}_2, \mathbf{v}_2)
\end{aligned}$$

and so on. $\triangleleft$

Likewise, given r linear functionals $\lambda_1, \ldots, \lambda_r \in V^*$, and r vectors $\mathbf{v}_1, \ldots, \mathbf{v}_r \in V$, we can write

$$(\lambda_1 \otimes \cdots \otimes \lambda_r)(\mathbf{v}_1 \otimes \cdots \otimes \mathbf{v}_r) = \lambda_1(\mathbf{v}_1) \cdots \lambda_r(\mathbf{v}_r).$$

The following notation is handy:

Definition 4.83. Write V^r for the r-fold Cartesian product of V with itself.

Then we can say that $\mathcal{T}_r^0(V)$ consists of **r-linear** functions from V^r to $\mathbb{R}$. Note that elements of V^r are just r-tuples of vectors. For example, if V has dimension m, we might represent an element of V^r using mr scalars, with respect to some basis for V:

$$(\mathbf{v}_1, \ldots, \mathbf{v}_r) = ((v_{11}, \ldots, v_{1m}), (v_{21}, \ldots, v_{2m}), \ldots, (v_{r1}, \ldots, v_{rm})).$$

Note that, just as in section 4.7.3, an element of $\mathcal{T}_r^0(V)$ may or may not be decomposable into a tensor product of linear functionals. However, as discussed in section 4.7.3, we can decompose an element of $\mathcal{T}_r^0(V)$ into a sum of simple tensors, using a basis of V^*. An example appears in the next section.

4.7.5 Explicit representation of multilinear functions

In section 4.7.2 we saw that, given a basis, we can view elements of

$$\mathcal{T}_0^s(V) = \overbrace{V \otimes \cdots \otimes V}^{s \text{ times}}$$

as multidimensional arrays. The same is true for elements of

$$\mathcal{T}_r^0(V) = \underbrace{V^* \otimes \cdots \otimes V^*}_{r \text{ times}}.$$

Example 4.84. ▷ Let $\psi \in V^* \otimes V^* \otimes V^*$ where $m = \dim V$. (I do not need to assume that ψ is decomposable into the form $\lambda \otimes \mu \otimes \nu$ for linear functionals λ, μ, and ν.) Then for $u, v, w \in V$, using multilinearity (here, trilinearity) we have:

$$\begin{aligned}
\psi(\mathbf{u}, \mathbf{v}, \mathbf{w}) &= \psi(u_1\mathbf{e}_1 + u_2\mathbf{e}_2 + u_3\mathbf{e}_3, v_1\mathbf{e}_1 + v_2\mathbf{e}_2 + v_3\mathbf{e}_3, w_1\mathbf{e}_1 + w_2\mathbf{e}_2 + w_3\mathbf{e}_3) \\
&= \sum_{i=1}^m \sum_{j=1}^m \sum_{k=1}^m u_i v_j w_k \psi(\mathbf{e}_i, \mathbf{e}_j, \mathbf{e}_k) = \sum_{i=1}^m \sum_{j=1}^m \sum_{k=1}^m u_i v_j w_k a_{ijk}
\end{aligned}$$

where the a_{ijk}'s are real numbers which depend on ψ. All the information about what ψ does is contained in the 3-dimensional $m \times m \times m$ array with entries a_{ijk}. In particular, all the a_{ijk}'s are *constants*. ◁

Here is the key **application formula** for tensors. That is, we can think of an r-tensor as an r-dimensional array, while we can also think of it as an r-linear function operating on r vectors. How do we apply the former to the latter? This is illustrated by example.

Example 4.85. ▷ Let T be a 3-tensor on V; for this example, use the standard basis for V. As discussed in section 4.7.2, T is a linear combination of the m^3 basis tensors $\mathbf{e}_i^* \otimes \mathbf{e}_j^* \otimes \mathbf{e}_k^*$. Likewise, vectors $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ are also linear combinations of the basis vectors $\mathbf{e}_i$:

$$\begin{aligned}
T &= \sum_{ijk} T(\mathbf{e}_i, \mathbf{e}_j, \mathbf{e}_k)\mathbf{e}_i^* \otimes \mathbf{e}_j^* \otimes \mathbf{e}_k^* = \sum_{ijk} t_{ijk}\mathbf{e}_i^* \otimes \mathbf{e}_j^* \otimes \mathbf{e}_k^* \\
\mathbf{u} &= \sum_{i'} u_{i'}\mathbf{e}_{i'}, \qquad \mathbf{v} = \sum_{j'} v_{j'}\mathbf{e}_{j'}, \qquad \mathbf{w} = \sum_{k'} w_{k'}\mathbf{e}_{k'} \\
T(\mathbf{u}, \mathbf{v}, \mathbf{w}) &= \sum_{ijk} t_{ijk}\mathbf{e}_i^* \otimes \mathbf{e}_j^* \otimes \mathbf{e}_k^* \left(\sum_{i'} u_{i'}e_{i'}, \sum_{j'} v_{j'}e_{j'}, \sum_{k'} w_{k'}e_{k'} \right) = \sum_{ijk} t_{ijk}u_i v_j w_k.
\end{aligned}$$

◁

Remark 4.86. When doing computations by hand, another expression is useful. This is illustrated by a 2-tensor example, with $\dim(V) = 2$. Suppose that, with respect to the standard basis, we have the following 2-tensor ψ and vectors $\mathbf{u}, \mathbf{v}$. We might write

$$\psi(\mathbf{u}, \mathbf{v}) = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 5 \\ 6 \end{pmatrix} \begin{pmatrix} 7 \\ 8 \end{pmatrix} = 1 \cdot 5 \cdot 7 + 2 \cdot 5 \cdot 8 + 3 \cdot 6 \cdot 7 + 4 \cdot 6 \cdot 8 = 433.$$

However, it is easy to mistake whether the 2 goes with the 5 and 8, or with the 6 and 7. It is less error-prone to observe that in the expression

$$\psi(\mathbf{u}, \mathbf{v}) = \sum_{ij} t_{ij}u_i v_j$$

we have $u_i v_j$ which are precisely the components of $\mathbf{u} \otimes \mathbf{v}$. Define $\odot$ on multi-dimensional arrays to be *componentwise* multiplication, e.g. if A and B have entries a_{ij} and b_{ij}, then $A \odot B$ has entries $a_{ij} b_{ij}$. Then

$$\psi(\mathbf{u}, \mathbf{v}) = \psi \odot (\mathbf{u} \otimes \mathbf{v}).$$

In this example, multiplying out $\mathbf{u} \otimes \mathbf{v}$ as in example 4.74, we have

$$\psi(\mathbf{u}, \mathbf{v}) = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \odot \left[\begin{pmatrix} 5 \\ 6 \end{pmatrix} \otimes \begin{pmatrix} 7 \\ 8 \end{pmatrix} \right] = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \odot \begin{pmatrix} 35 & 40 \\ 42 & 48 \end{pmatrix} = 35 + 80 + 126 + 192 = 433.$$

In general, for an r-tensor, write out the tensor product of the argument vectors, then do a componentwise product with the tensor coefficients, then sum up.

Remark 4.87. In this paper, I do many tensor computations with respect to coordinates — usually the standard basis. However, change of coordinates is a key part of tensor algebra, as well as differential geometry. In fact, coordinates are just an expression of an abstract object with respect to specific coordinates.

A multidimensional array is to a tensor as a matrix is to a linear transformation.

Change of coordinates will be discussed further in sections [xxx write and xref].

4.7.6 Mixed tensors

Definition 4.88. We write

$$\mathcal{T}_r^s(V) = \overbrace{V \otimes \cdots \otimes V}^{s \text{ times}} \otimes \underbrace{V^* \otimes \cdots \otimes V^*}_{r \text{ times}}.$$

The integer s is called the **contravariant degree**; r is called the **covariant degree**. There are s **contravariant indices** and r **covariant indices**. An element of $\mathcal{T}_r^s(V)$ is said to be of **type** s, r. If $s > 0$ and $r > 0$, elements of $\mathcal{T}_r^s(V)$ are said to be of **mixed type**. When we refer to an r-tensor, we mean a tensor of type $0, r$, i.e. an r-linear function on V. An r-tensor is said to have **degree** r, or sometimes **order** r. We write

$$\operatorname{ord}(\omega)$$

for the order of an r-tensor ω.

Note the following in particular:

- $\mathcal{T}_0^0(V) = \mathbb{R}$.
- $\mathcal{T}_0^1(V) = V$; $\mathcal{T}_1^0(V) = V^*$.
- $\mathcal{T}_0^2(V) = V \otimes V$; $\mathcal{T}_1^1(V) = V \otimes V^*$; $\mathcal{T}_2^0(V) = V^* \otimes V^*$.
- A sample decomposable element of $\mathcal{T}_0^2(V)$ is $\mathbf{u} \otimes \mathbf{v}$ for $\mathbf{u}, \mathbf{v} \in V$.
- A sample decomposable element of $\mathcal{T}_2^0(V)$ is $\lambda \otimes \mu$ for $\lambda, \mu : V \to \mathbb{R}$ i.e. $\lambda, \mu \in V^*$.

Note that $\mathbb{R}$ is a subspace of V and V^*, and so on. So we have the following:

$$
\begin{tikzcd}
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
\mathcal{T}_0^2(V) = V \otimes V \arrow[ur] & & \mathcal{T}_1^1(V) = V \otimes V^* \arrow[ul] \arrow[ur] & & \mathcal{T}_2^0(V) = V^* \otimes V^* \arrow[ul] \arrow[ur] & \\
& \mathcal{T}_0^1(V) = V \arrow[ul] \arrow[ur] & & \mathcal{T}_1^0(V) = V^* \arrow[ul] \arrow[ur] & & \\
& \text{vectors} & & \text{covectors} & & \\
& & \mathcal{T}_0^0(V) = \mathbb{R} \arrow[ul] \arrow[ur] & & & \\
& & \text{scalars} & & &
\end{tikzcd}
$$

Remark 4.89. As noted in section 4.6.5, we can identity V with V^{**}. So, we can view an arbitrary mixed tensor as a multilinear function from $(V^*)^s \oplus V^r$ to $\mathbb{R}$:

$$
\begin{aligned}
\psi \in \mathcal{T}_r^s(V) &= \overbrace{V^{**} \otimes \cdots \otimes V^{**}}^{s \text{ times}} \otimes \underbrace{V^* \otimes \cdots \otimes V^*}_{r \text{ times}} \\
\psi : (V^*)^s \oplus V^r &\to \mathbb{R}.
\end{aligned}
$$

Remark 4.90. For this course we focus on tensors of type $0, r$, which we call **r-tensors**, and tensors of type $1, 0$, which we call **vectors**. Mixed tensors perhaps seem like too much generality. However, they have the following benefits:

- In sections 6.3.9 and 6.3.10 we will extend this notion, and obtain **forms** and **vector fields**, respectively, over **manifolds**. Mixed tensors unify these concepts.
- Fundamental concepts such as linear transformations (section 4.7.14) and curvature ([**Guggenheimer**], [**Lee3**]) turn out to be mixed tensors.

4.7.7 Examples of tensors

Example 4.91. $\triangleright$ Let $\lambda \in V^*$. As discussed in section 4.6.6, with reference to a given basis, λ is represented by a $1 \times m$ matrix, or **row vector**. The action $\lambda(\mathbf{v})$ is the product of λ's row vector with $\mathbf{v}$'s **column vector**, which is the dot product of λ and $\mathbf{v}$. For example:

$$
\lambda(v) = \begin{pmatrix} a_1 a_2 \cdots a_m \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_m \end{pmatrix} = a_1 v_1 + \ldots + a_m v_m.
$$

Another way to see this is that linearity of λ gives

$$
\lambda(v) = \lambda\left(\sum_{i=1}^m v_i \mathbf{e}_j\right) = \sum_{i=1}^m v_i \lambda(\mathbf{e}_j) = \sum_{i=1}^m v_i a_i
$$

The 1-dimensional array $\mathbf{a}$, of length m, gives the information needed to compute λ of any $\mathbf{v} \in V$. $\triangleleft$

Example 4.92. ▷ An example of a 2-tensor is the standard dot product. Let $\mathbf{u}, \mathbf{v} \in V$ with $m = \dim(V)$. Holding either $\mathbf{u}$ or $\mathbf{v}$ fixed and varying the other gives a linear transformation, so this makes sense. With respect to the standard basis, we can think of $\mathbf{u} \cdot \mathbf{v}$ as

$$\begin{aligned}
\mathbf{u} \cdot \mathbf{v} &= \left(\sum_{i=1}^{m} u_i \mathbf{e}_i \right) \cdot \mathbf{v} \\
&= \sum_{i=1}^{m} u_i \left(\mathbf{e}_i \cdot \mathbf{v} \right) \\
&= \sum_{i=1}^{m} u_i \mathbf{e}_i \cdot \left(\sum_{j=1}^{m} v_j \mathbf{e}_j \right) \\
&= \sum_{i=1}^{m} \sum_{j=1}^{m} u_i v_j (\mathbf{e}_i \cdot \mathbf{e}_j) \\
&= \sum_{i=1}^{m} \sum_{j=1}^{m} u_i v_j \delta_{ij}.
\end{aligned}$$

Here, the $m \times m$ array is the **identity matrix**, with entries δ_{ij}. These m^2 numbers indicate how to form the inner product with respect to the standard basis in Euclidean space. (Since the $m \times m$ identity matrix has rank m, if $m > 1$ then the dot-product tensor is **indecomposable** by the reasoning in section 4.7.3.) ◁

Remark 4.93. Using matrices which are not necessarily the identity leads to **bilinear forms**; see [**FIS**] for more information. These two-tensors generalize the inner product: in fact, an inner product is nothing more than a symmetric positive-definite bilinear form.

Example 4.94. ▷ Recall that the determinant on $m \times m$ matrices is an m-linear function, either on rows or on columns. Here, think of it as an m-linear function on rows. (More can be said about the determinant: see section 4.7.12.) Let $\mathbf{u} = (u_1, u_2), \mathbf{v} = (v_1, v_2) \in \mathbb{R}^2$. Because of the bilinearity of the determinant (xref backward after marking the equation), we have

$$\begin{aligned}
\det(\mathbf{u}, \mathbf{v}) &= \det\left(u_1 \mathbf{e}_1 + u_2 \mathbf{e}_2, v_1 \mathbf{e}_1 + v_2 \mathbf{e}_2\right) \\
&= \det\left(u_1 \mathbf{e}_1, v_1 \mathbf{e}_1\right) + \det\left(u_1 \mathbf{e}_1, v_2 \mathbf{e}_2\right) + \det\left(u_2 \mathbf{e}_2, v_1 \mathbf{e}_1\right) + \det\left(u_2 \mathbf{e}_2, v_2 \mathbf{e}_2\right) \\
&= u_1 v_1 \det\left(\mathbf{e}_1, \mathbf{e}_1\right) + u_1 v_2 \det\left(\mathbf{e}_1, \mathbf{e}_2\right) + u_2 v_1 \det\left(\mathbf{e}_2, \mathbf{e}_1\right) + u_2 v_2 \det\left(\mathbf{e}_2, \mathbf{e}_2\right).
\end{aligned}$$

It remains to find the values of $\det(\mathbf{e}_i, \mathbf{e}_j)$. But since we are thinking of matrices as lists of row vectors, these are

$$\det(\mathbf{e}_1, \mathbf{e}_1) = \det \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} = 0 \qquad \det(\mathbf{e}_1, \mathbf{e}_2) = \det \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = 1$$

$$\det(\mathbf{e}_2, \mathbf{e}_1) = \det \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = -1 \qquad \det(\mathbf{e}_2, \mathbf{e}_2) = \det \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} = 0$$

So, the determinant tensor on rows of 2×2 matrices may be thought of as

$$\det(\mathbf{u}, \mathbf{v}) = \sum_{i=1}^{2} \sum_{j=1}^{2} u_i v_j d_{ij}$$

where d_{ij} are the entries of the 2×2 array

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

◁

Example 4.95. ▷ Similarly, for 3×3 matrices:

$$\begin{aligned}
\det(\mathbf{u}, \mathbf{v}, \mathbf{w}) &= \det\left(\sum_{i=1}^{3} u_i \mathbf{e}_i, \sum_{j=1}^{3} v_j \mathbf{e}_j, \sum_{k=1}^{3} w_k \mathbf{e}_k\right). \\
&= \sum_{i=1}^{3}\sum_{j=1}^{3}\sum_{k=1}^{3} \det\left(u_i \mathbf{e}_i, v_j \mathbf{e}_j, w_k \mathbf{e}_k\right) \\
&= \sum_{i=1}^{3}\sum_{j=1}^{3}\sum_{k=1}^{3} u_i v_j w_k \det\left(\mathbf{e}_i, \mathbf{e}_j, \mathbf{e}_k\right).
\end{aligned}$$

It remains to find the values of $\det(\mathbf{e}_i, \mathbf{e}_j, \mathbf{e}_k)$. There are 27 of them, but only 6 with non-zero values. They are

$$\det(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3) = \det\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = 1, \qquad \det(\mathbf{e}_1, \mathbf{e}_3, \mathbf{e}_2) = \det\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} = -1,$$

$$\det(\mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_1) = \det\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} = 1, \qquad \det(\mathbf{e}_3, \mathbf{e}_2, \mathbf{e}_1) = \det\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} = -1,$$

$$\det(\mathbf{e}_3, \mathbf{e}_1, \mathbf{e}_2) = \det\begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} = 1, \qquad \det(\mathbf{e}_2, \mathbf{e}_1, \mathbf{e}_3) = \det\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} = -1.$$

The determinant tensor on rows of 3×3 matrices may be thought of as

$$\det(\mathbf{u}, \mathbf{v}, \mathbf{w}) = \sum_{i=1}^{3}\sum_{j=1}^{3}\sum_{k=1}^{3} u_i v_j w_k d_{ijk}$$

where d_{ijk} are the entries of the following $3 \times 3 \times 3$ array where i indexes the face (front, middle, back), j indexes rows, and k indexes columns:

$$\begin{array}{ccc} \text{front} & \text{middle} & \text{back} \\ \end{array}$$
$$\left(\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix} \quad \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \quad \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}\right).$$

◁

4.7.8 Pullbacks

This example is from [**Spivak1**], p. 77.

Let $f : V \to W$ be a linear transformation, and let $\lambda : W^k \to \mathbb{R}$ be a k-tensor, i.e. $\lambda \in \mathcal{T}^k(W)$. We can map from V^k to $\mathbb{R}$ by going through W^k first. That is, for $(\mathbf{v}_1, \ldots, \mathbf{v}_k) \in V^k$, $(f(\mathbf{v}_1), \ldots, f(\mathbf{v}_k))$ is in W^k, and so we can apply λ to it:

$$\begin{tikzcd} V^k \arrow[d, "f"'] \arrow[dr, "\lambda \circ f"] & \\ W^k \arrow[r, "\lambda"'] & \mathbb{R} \end{tikzcd}$$

Formally, $\lambda \circ f : V^k \to \mathbb{R}$. Given $\lambda \in \mathcal{T}^k(W)$, we have obtained $\lambda \circ f \in \mathcal{T}^k(V)$. Post-composing λ by f is said to be a **pullback** of λ from W^k to V^k. Another way to look at this is that, given $f : V \to W$, we have a map from $\mathcal{T}^k(W)$ to $\mathcal{T}^k(V)$:

$$\begin{aligned} f: & \quad V \to W \\ f^*: & \quad \mathcal{T}^k(W) \to \mathcal{T}^k(V) \\ f^*(\lambda) &= \lambda \circ f \\ (f^*\lambda)(\mathbf{v}_1, \ldots, \mathbf{v}_k) &= \lambda(f(\mathbf{v}_1), \ldots, f(\mathbf{v}_k)). \end{aligned}$$

In this way, f^* takes k-tensors to k-tensors:

$$\begin{tikzcd} V^k \arrow[d, "f"'] \arrow[dr, "\lambda \circ f"] & \\ W^k \arrow[r, "\lambda"'] & \mathbb{R} \end{tikzcd} \qquad \begin{tikzcd} \mathcal{T}^k(V) & \\ \mathcal{T}^k(W) \arrow[u, "f^*"] & \end{tikzcd}$$

4.7.9 Tensor algebras

In section 4.7.2 we saw that $\mathcal{T}^r_s(V)$ is a vector space over $\mathbb{R}$, with dimension $\dim(V)^{r+s}$. This permits addition of tensors. We can think of this in two ways, both of which are illustrated by example. First, since tensors are multilinear functions, we can pick two tensors and r vectors, then see what happens to the sum of tensors. Doing this uses the vectors as something external, though, so this is less than satisfying. Second, we can see what happens to the coefficients in the tensors themselves. The former turns into the latter as follows, for a pair of 2-tensors ϕ and ψ:

$$\begin{aligned} \phi &= \sum_{ij} \phi(\mathbf{e}_i, \mathbf{e}_j)\mathbf{e}_i^* \otimes \mathbf{e}_j^*, \quad \phi(\mathbf{e}_i, \mathbf{e}_j) = a_{ij} \\ \psi &= \sum_{ij} \psi(\mathbf{e}_i, \mathbf{e}_j)\mathbf{e}_i^* \otimes \mathbf{e}_j^*, \quad \psi(\mathbf{e}_i, \mathbf{e}_j) = b_{ij} \\ (\phi + \psi)(\mathbf{u}, \mathbf{v}) &= \phi(\mathbf{u}, \mathbf{v}) + \psi(\mathbf{u}, \mathbf{v}) \\ &= \sum_{ij} a_{ij} u_i v_j + \sum_{ij} b_{ij} u_i v_j \\ &= \sum_{ij} [a_{ij} + b_{ij}]\, u_i v_j. \end{aligned}$$

So, by applying our pair of tensors to an arbitrary pair of vectors, we see that we didn't need them after all: namely, the coefficients of $\phi + \psi$ are the sums of coefficients of ϕ and ψ. This is just vector-space addition, as we would expect.

Recall from section 4.1.8 that we can turn this vector space into an algebra if only we can multiply tensors. As with addition, we can do this **argumentwise** (with reference to vectors) or **coefficientwise** (with reference

only to the tensors themselves). That is, knowing that tensors are r-linear functions, we can think of what tensor multiplication would *have* to be. Again, this is illustrated by example. Here, let ϕ be a 2-tensor, and let ψ be a 1-tensor.

$$\begin{aligned}
\phi &= \sum_{ij} \phi(\mathbf{e}_i, \mathbf{e}_j)\mathbf{e}_i^* \otimes \mathbf{e}_j^*, \quad \phi(\mathbf{e}_i, \mathbf{e}_j) = a_{ij} \\
\psi &= \sum_k \psi(\mathbf{e}_k)\mathbf{e}_k^*, \quad \psi(\mathbf{e}_k) = b_k \\
(\phi \otimes \psi)(\mathbf{u}, \mathbf{v}, \mathbf{w}) &= \phi(\mathbf{u}, \mathbf{v})\psi(\mathbf{w}) \\
&= \left[\sum_{ij} a_{ij} u_i v_j\right]\left[\sum_k b_k w_k\right] \\
&= \sum_{ijk} a_{ij} b_k u_i v_j w_k.
\end{aligned}$$

As above, we get the coefficients of the product tensor $\phi \otimes \psi$, namely $a_{ij}b_k$, as an object of its own.

I chose to do a specific example to keep the notation from becoming frightening. But we can extend this example to mixed tensors in general.

Definition 4.96. Let ϕ be a tensor of type s_1, r_1; let ψ be a tensor of type s_2, r_2. Then the **tensor product** $\phi \otimes \psi$ is of type $s_1 + s_2, r_1 + r_2$. It is defined by the property

$$(\phi \otimes \psi)(\lambda_1, \ldots, \lambda_{s_1}, \mu_1, \ldots, \mu_{s_2}, \mathbf{u}_1, \ldots, \mathbf{u}_{r_1}, \mathbf{v}_1, \ldots, \mathbf{v}_{r_2}) \tag{4.1}$$

$$= \phi(\lambda_1, \ldots, \lambda_{s_1}, \mathbf{u}_1, \ldots, \mathbf{u}_{r_1})\psi(\mu_1, \ldots, \mu_{s_2}, \mathbf{v}_1, \ldots, \mathbf{v}_{r_2}). \tag{4.2}$$

The coefficients of $\phi \otimes \psi$ are obtained by taking the outer product of the coefficients of ϕ and ψ, as in section 4.7.2.

Note however that unless ϕ or ψ is of type 0,0 — that is, a scalar — then the vector space $\mathcal{T}_r^s(V)$ is not closed under multiplication: orders increase when tensors are multiplied.

Definition 4.97. Let V be a vector space. The **full tensor algebra**, written $\mathcal{T}(V)$, is

$$\mathcal{T}(V) = \bigoplus_{s,r=0}^{\infty} \mathcal{T}_r^s(V).$$

Remark 4.98. Using categorical terminology, $\mathcal{T}_s^r$ is a functor from vector spaces to vector spaces; $\mathcal{T}$ is a functor from vector spaces to graded rings.

4.7.10 Permutations acting on multilinear functions

It is natural to ask about the symmetries of a tensor. For example, if ψ is a 2-tensor, then we might ask if

$$\psi(\mathbf{u}, \mathbf{v}) = \psi(\mathbf{v}, \mathbf{u}) \qquad \text{(symmetry)}$$

or

$$\psi(\mathbf{u}, \mathbf{v}) = -\psi(\mathbf{v}, \mathbf{u}) \qquad \text{(skew symmetry)}$$

for all vectors $\mathbf{u}$ and $\mathbf{v}$. For example, in section 4.7.7 we saw that the inner product is symmetric, while the determinant is skew symmetric. In order to generalize this concept, and in order to define forms in section

4.7.11, we will need a notion of how permutations act on tensors. Proceeding as in section 4.7.9, we can first do this argumentwise, using test vectors, then find out what the effect is on the tensor itself.

This works for all tensors of non-mixed type, i.e. contravariant or covariant tensors. (One could define a notion of symmetry for mixed tensors, but this appears [**Lee2**] not to be useful in practice.) Here, I will use covariant tensors for my examples. Let ψ be an r-tensor; let $\mathbf{u}_1, \ldots, \mathbf{u}_r \in V$. Let $\sigma \in \mathcal{S}_r$. Then we want

$$\psi\left((\mathbf{u}_1, \ldots, \mathbf{u}_3)^\sigma\right) = \psi\left(\mathbf{u}_{\sigma^{-1}(1)}, \ldots, \mathbf{u}_{\sigma^{-1}(r)}\right).$$

This is completely reasonable. The only question is why we use the inverse of the permutation on the right-hand side. [xxx discuss]

As we did for tensor addition and tensor multiplication, let's see if we can get rid of the need for test vectors. Taking the specific permutation $\sigma = (1\ 2\ 3)$, we have

$$\begin{aligned}
\psi &= \sum_{ijk} \psi(\mathbf{e}_i, \mathbf{e}_j, \mathbf{e}_k)\mathbf{e}_i^* \otimes \mathbf{e}_j^* \otimes \mathbf{e}_k^*; \quad a_{ijk} = \psi(\mathbf{e}_i, \mathbf{e}_j, \mathbf{e}_k) \\
\psi(\mathbf{u}, \mathbf{v}, \mathbf{w}) &= \sum_{ijk} a_{ijk} u_i v_j w_k \\
\psi(\mathbf{w}, \mathbf{u}, \mathbf{v}) &= \sum_{ijk} a_{ijk} w_i u_j v_k = \sum_{ijk} a_{ijk} u_j v_k w_i = \sum_{ijk} a_{kij} u_i v_j w_k.
\end{aligned}$$

Generalizing from this example, we can define

$$\psi^\sigma(\mathbf{u}_1, \ldots, \mathbf{u}_r) = \psi\left((\mathbf{u}_1, \ldots, \mathbf{u}_r)^\sigma\right)$$

where on the right-hand side, we permute the *arguments* to ψ, and on the left-hand side we permute the *coefficients*.

This generalizes the concept of transpose discussed in section 4.6.10.

4.7.11 Symmetric and alternating tensor spaces

Above we said that a 2-tensor ψ is symmetric if $\psi(\mathbf{v}, \mathbf{u}) = \psi(\mathbf{u}, \mathbf{v})$ for all $\mathbf{u}, \mathbf{v}$, and likewise ψ is skew-symmetric or alternating if $\psi(\mathbf{v}, \mathbf{u}) = -\psi(\mathbf{u}, \mathbf{v})$ for all $\mathbf{u}, \mathbf{v}$. We can now generalize this as follows.

Definition 4.99. A non-mixed tensor ψ of order r is **symmetric** if

$$\psi\left((\mathbf{v}_1, \ldots, \mathbf{v}_r)^\tau\right) = \psi\left(\mathbf{v}_1, \ldots, \mathbf{v}_r\right)$$

for all transpositions $\tau \in \mathcal{S}_r$. Likewise, ψ of order r is **alternating** if

$$\psi\left((\mathbf{v}_1, \ldots, \mathbf{v}_r)^\tau\right) = -\psi\left(\mathbf{v}_1, \ldots, \mathbf{v}_r\right)$$

for all transpositions τ. An alternating tensor is also called a **form**.

Remark 4.100. The definition was done argumentwise, but following the discussion in section 4.7.10, we can also look at it coefficientwise. That is, for symmetric r-tensors ψ and transpositions $\tau \in \mathcal{S}_r$,

$$\psi^\tau = \psi.$$

Likewise, for alternating ψ,

$$\psi^\tau = -\psi.$$

Note in particular that the **symmetric** r-tensors are held fixed by the **symmetric group** $\mathcal{S}_r$, while the **alternating** r-tensors are held fixed only by the **alternating group** $\mathcal{A}_r$.

Remark 4.101. All tensors of order 0 or 1 are (trivially) symmetric and alternating.

Remark 4.102. Let ψ be an alternating 3-tensor. The alternation property forces $\psi(\mathbf{u}, \mathbf{u}, \mathbf{v}) = 0$. In general, for an arbitrary alternating tensor ψ, ψ has value zero when evaluated on r vectors any time two or more of those vectors are equal. As discussed in section 4.7.5, the coefficients a_{ijk} of a 3-tensor ψ with respect to a basis $\{\mathbf{b}_1, \ldots, \mathbf{b}_m\}$ are given by $a_{ijk} = \psi(\mathbf{b}_i, \mathbf{b}_j, \mathbf{b}_k)$ and so the a_{ijk} coefficients must be zero whenever $i = j$, $i = k$, or $j = k$. In general, the coefficients of an arbitrary alternating r-tensor are zero when any index is repeated. (This is reminiscent of the fact that the diagonal entries of a skew-symmetric matrix must be zero. The diagonal entries a_{ii} have equal first and second indices.)

Remark 4.103. Just as 1-forms have zero output when vector arguments are repeated, they have zero output when wedge components are repeated. Let λ be a linear functional, i.e. a 1-tensor, on V. By the definition of wedge product, for any $\mathbf{u}, \mathbf{v} \in V$, we have

$$\begin{aligned} \lambda \wedge \lambda(\mathbf{u}, \mathbf{v}) &= \lambda \otimes \lambda(\mathbf{u}, \mathbf{v}) - \lambda \otimes \lambda(\mathbf{v}, \mathbf{u}) \\ &= \lambda(\mathbf{u})\lambda(\mathbf{v}) - \lambda(\mathbf{v})\lambda(\mathbf{u}) = \lambda(\mathbf{u})\lambda(\mathbf{v}) - \lambda(\mathbf{u})\lambda(\mathbf{v}) = 0. \end{aligned}$$

Note that this remark applies to one-forms. See remark 4.109 for full details.

It is easy to check that the sets of symmetric and alternating r-tensors are closed under scalar multiplication and vector addition, and thus that they are vector spaces.

Definition 4.104. More or less following [**Lee2**] and Pickrell, I write

$$\Sigma^r(V^*), \qquad \Lambda^r(V^*)$$

for the spaces of symmetric and alternating r-tensors, respectively.

One naturally asks about the dimension of these vector spaces. Just as for matrices in section 4.6.10, we note that to specify the coefficients of a symmetric tensor, it suffices to consider those with non-decreasing indices. For example, for a 3-tensor ψ, if we know what the coefficient a_{112} is, then we know that a_{121} and a_{211} must be the same; for the any of the coefficients of a symmetric 3-tensor, it suffices to specify the coefficients a_{ijk} for $i \leq j \leq k$. For a symmetric tensor of arbitrary order r, it suffices to specify the coefficients $a_{i_1,\ldots,i_r}$ for $i_1 \leq \ldots \leq i_r$.

For an alternating 3-tensor ψ, we know by remark 4.102 that the coefficients a_{ijk} are zero for any repeated indices. For this reason, along with the reasons from the previous paragraph, it suffices to specify the values of a_{ijk} for $i < j < k$. For an alternating tensor of arbitrary order r, it suffices to specify the coefficients $a_{i_1,\ldots,i_r}$ for $i_1 < \ldots < i_r$.

To find the dimensions of the spaces of symmetric and alternating r-tensors on a vector space of dimension m, we only need to count how many ways we can choose coefficients as just described — that is, non-decreasing and increasing indices. A counting argument shows that

$$\dim(\Sigma^r(V^*)) = \binom{m+r-1}{r} \qquad \text{and} \qquad \dim(\Lambda^r(V^*)) = \binom{m}{r}.$$

Remark 4.105. Note that the above formulas may be written as

$$\binom{m+r-1}{r} = \frac{m(m+1)\cdots(m+r-2)(m+r-1)}{r(r-1)\cdots 2\cdot 1}, \qquad \binom{m}{r} = \frac{m(m-1)\cdots(m-r+2)(m-r+1)}{r(r-1)\cdots 2\cdot 1}.$$

For $m = 2$, we have

$$\frac{m(m+1)}{2} + \frac{m(m-1)}{2} = m^2$$

and so an arbitrary 2-tensor may be written as the sum of symmetric and alternating components. However, for $r \geq 3$, $\binom{(m+r-1)}{r} + \binom{m}{r} < m^r$ and so we cannot write an arbitrary r-tensor in terms of symmetric and alternating components.

4.7.12 Symmetric and alternating tensor algebras

The usual tensor product of two symmetric or alternating tensors need not be symmetric or alternating, respectively. (Since all 1-tensors are both alternating and symmetric, we can use example 4.74 to see this.) So, in order to turn the vector spaces of symmetric and alternating tensors into algebras, we need to define multiplications on them which preserve the symmetries.

First, much as in section 4.6.10, we define techniques for computing the symmetrization and the alternation of a given tensor.

Definition 4.106. Let Sym map from r-tensors to r-tensors by

$$\mathrm{Sym}(\psi) = \frac{1}{r!} \sum_{\sigma \in \mathcal{S}_r} \psi^\sigma.$$

Likewise, let Alt map from r-tensors to r-tensors by

$$\mathrm{Alt}(\psi) = \frac{1}{r!} \sum_{\sigma \in \mathcal{S}_r} \mathrm{sgn}(\sigma) \psi^\sigma.$$

where $\mathrm{sgn}(\sigma)$ is $+1$ for even permutations and -1 for odd permutations.

See any of the geometry texts in the bibliography for proofs that the symmetrization and alternation do in fact produce symmetric and alternating tensors, respectively. Moreover, these are both **projections** from $\mathcal{T}^0_r(V)$ onto $\Sigma^r(V^*)$ and $\Lambda^r(V^*)$, respectively, in the sense of section 4.4: they are surjective as well as idempotent.

Now we can define symmetry-preserving multiplications. Just as with the tensor product as described in section 4.7.9, though, the product of tensors won't have the same order as the multiplicands.

Definition 4.107. We write

$$\Sigma(V^*) = \bigoplus_{r=0}^{\infty} \Sigma^r(V^*) \qquad \text{and} \qquad \Lambda(V^*) = \bigoplus_{r=0}^{\infty} \Lambda^r(V^*).$$

Definition 4.108. Let $\phi \in \Sigma^k(V^*)$ and $\psi \in \Sigma^\ell(V^*)$. That is, suppose ϕ has order k and ψ has order ℓ. The **symmetric product** of ϕ and ψ, written using juxtaposition, is

$$\phi\psi = \mathrm{Sym}(\phi \otimes \psi).$$

Likewise, if $\phi \in \Lambda^k(V^*)$ and $\psi \in \Lambda^\ell(V^*)$, then the **alternating product**, or **wedge product**, of ϕ and ψ is

$$\begin{aligned}
\phi \wedge \psi &= \frac{(k+\ell)!}{k!\ell!} \mathrm{Alt}(\phi \otimes \psi) \\
&= \frac{(k+\ell)!}{k!\ell!} \frac{1}{(k+\ell)!} \sum_{\sigma \in \mathcal{S}_{k+\ell}} \mathrm{sgn}(\sigma)(\phi \otimes \psi)^\sigma \\
&= \frac{1}{k!\ell!} \sum_{\sigma \in \mathcal{S}_{k+\ell}} \mathrm{sgn}(\sigma)(\phi \otimes \psi)^\sigma.
\end{aligned}$$

Note that in both cases the product has order $k + \ell$. The factorial multiplier is included so that the wedge of m basis functionals, where $m = \dim(V)$, will be exactly the determinant. This is because the $1/k!\ell!$ in the rightmost expression above is 1 whenever we are wedging 1-tensors, which have orders $k = \ell = 1$. We will see more about this in example 4.118.

With these multiplications, $\Sigma(V^*)$ and $\Lambda(V^*)$ are both **algebras**. That is, we can add, subtract, and multiply tensors. In particular, all the usual arithmetic rules for (non-commutative) rings apply, such as distributivity. But what about commutativity?

Proposition 4.109. *Let ϕ and ψ be covariant symmetric tensors of orders k and ℓ, respectively; let ω and η be covariant alternating tensors of orders k and ℓ. Then it can be shown that*

$$\phi\psi = \psi\phi$$

and

$$\omega \wedge \eta = (-1)^{k\ell}\eta \wedge \omega.$$

Mnemonic 4.110. For the symmetric case, take

$$\phi = \lambda_1\lambda_2$$

and

$$\psi = \lambda_3\lambda_4\lambda_5$$

where each λ_i has degree 1. Then

$$\phi\psi = (\lambda_1\lambda_2)(\lambda_3\lambda_4\lambda_5) = \lambda_1(\lambda_3\lambda_4\lambda_5)\lambda_2 = (\lambda_3\lambda_4\lambda_5)\lambda_1\lambda_2 = \psi\phi$$

since we can transpose degree-1 symmetric terms. Likewise, for the alternating case, take

$$\omega = \theta_1 \wedge \theta_2$$

and

$$\eta = \theta_3 \wedge \theta_4 \wedge \theta_5$$

where each θ_i has degree 1. Then

$$\omega \wedge \eta = (\theta_1 \wedge \theta_2) \wedge (\theta_3 \wedge \theta_4 \wedge \theta_5) = -\theta_1 \wedge (\theta_3 \wedge \theta_4 \wedge \theta_5) \wedge \theta_2 = (\theta_3 \wedge \theta_4 \wedge \theta_5) \wedge \theta_1 \wedge \theta_2 = \eta \wedge \omega.$$

This is because we get three minus signs when we move θ_2 past θ_3, θ_4, and θ_5, then three more minus signs when we move θ_1 past θ_3, θ_4, and θ_5. If either k or ℓ is even, there is an even number of transpositions.

The proof simply generalizes this specific example from 2 to k and 3 to ℓ, along with ϕ, ψ, ω, η being linear combinations of elementary terms of the form shown here: use the distributive property of symmetric and alternating algebras to complete the proof.

Corollary 4.111. *Let ω be a k-form. If k is odd, then $\omega \wedge \omega = 0$.*

Example 4.112. $\triangleright$ Let ω and η both be 1-forms. Applying the definition of wedge product to test vectors $\mathbf{u}$ and $\mathbf{v}$ gives

$$\begin{aligned}
\omega \wedge \eta(\mathbf{u}, \mathbf{v}) &= \omega(\mathbf{u})\eta(\mathbf{v}) - \omega(\mathbf{v})\eta(\mathbf{u}) \\
\eta \wedge \omega(\mathbf{u}, \mathbf{v}) &= \eta(\mathbf{u})\omega(\mathbf{v}) - \eta(\mathbf{v})\omega(\mathbf{u}) \\
&= \omega(\mathbf{v})\eta(\mathbf{u}) - \omega(\mathbf{u})\eta(\mathbf{v}) \\
&= -\omega \wedge \eta(\mathbf{u}, \mathbf{v}).
\end{aligned}$$

Since the test vectors are arbitrary,

$$\omega \wedge \eta = -\eta \wedge \omega.$$

Now suppose ω is a 1-form and η is a 2-form. Again applying the definition of wedge product, recalling that η is alternating, and pulling the factorial coefficient over to the left, we have

$$\begin{aligned}
2\omega \wedge \eta(\mathbf{u},\mathbf{v},\mathbf{w}) &= \begin{array}{rrcl} & \omega \otimes \eta(\mathbf{u},\mathbf{v},\mathbf{w}) & - & \omega \otimes \eta(\mathbf{v},\mathbf{u},\mathbf{w}) \\ + & \omega \otimes \eta(\mathbf{v},\mathbf{w},\mathbf{u}) & - & \omega \otimes \eta(\mathbf{w},\mathbf{v},\mathbf{u}) \\ + & \omega \otimes \eta(\mathbf{w},\mathbf{u},\mathbf{v}) & - & \omega \otimes \eta(\mathbf{u},\mathbf{w},\mathbf{v}) \end{array} \\
&= \begin{array}{rrcl} & \omega(\mathbf{u})\eta(\mathbf{v},\mathbf{w}) & - & \omega(\mathbf{v})\eta(\mathbf{u},\mathbf{w}) \\ + & \omega(\mathbf{v})\eta(\mathbf{w},\mathbf{u}) & - & \omega(\mathbf{w})\eta(\mathbf{v},\mathbf{u}) \\ + & \omega(\mathbf{w})\eta(\mathbf{u},\mathbf{v}) & - & \omega(\mathbf{u})\eta(\mathbf{w},\mathbf{v}) \end{array} \\
&= \begin{array}{rl} & 2\omega(\mathbf{u})\eta(\mathbf{v},\mathbf{w}) \\ + & 2\omega(\mathbf{v})\eta(\mathbf{w},\mathbf{u}) \\ + & 2\omega(\mathbf{w})\eta(\mathbf{u},\mathbf{v}) \end{array} \\
2\eta \wedge \omega(\mathbf{u},\mathbf{v},\mathbf{w}) &= \begin{array}{rrcl} & \eta \otimes \omega(\mathbf{u},\mathbf{v},\mathbf{w}) & - & \eta \otimes \omega(\mathbf{v},\mathbf{u},\mathbf{w}) \\ + & \eta \otimes \omega(\mathbf{v},\mathbf{w},\mathbf{u}) & - & \eta \otimes \omega(\mathbf{w},\mathbf{v},\mathbf{u}) \\ + & \eta \otimes \omega(\mathbf{w},\mathbf{u},\mathbf{v}) & - & \eta \otimes \omega(\mathbf{u},\mathbf{w},\mathbf{v}) \end{array} \\
&= \begin{array}{rrcl} & \eta(\mathbf{u},\mathbf{v})\omega(\mathbf{w}) & - & \eta(\mathbf{v},\mathbf{u})\omega(\mathbf{w}) \\ + & \eta(\mathbf{v},\mathbf{w})\omega(\mathbf{u}) & - & \eta(\mathbf{w},\mathbf{v})\omega(\mathbf{u}) \\ + & \eta(\mathbf{w},\mathbf{u})\omega(\mathbf{v}) & - & \eta(\mathbf{u},\mathbf{w})\omega(\mathbf{v}) \end{array} \\
&= \begin{array}{rl} & 2\eta(\mathbf{u},\mathbf{v})\omega(\mathbf{w}) \\ + & 2\eta(\mathbf{v},\mathbf{w})\omega(\mathbf{u}) \\ + & 2\eta(\mathbf{w},\mathbf{u})\omega(\mathbf{v}) \end{array} \\
&= 2\omega \wedge \eta(\mathbf{u},\mathbf{v},\mathbf{w}).
\end{aligned}$$

Since the test vectors are arbitrary,

$$\omega \wedge \eta = \eta \wedge \omega.$$

◁

Definition 4.113. The alternating tensor algebra $\Lambda(V^*)$ is also called the **exterior algebra** or **Grassmann algebra** on V.

Remark 4.114. One may write down explicit bases for $\Sigma^r(V^*)$ and $\Lambda^r(V^*)$, just as for matrices in section 4.6.10. In fact, letting $\{\mathbf{b}_1^*, \ldots, \mathbf{b}_r^*\}$ be a basis for V^*, one obtains

$$\left\{\mathbf{b}_{i_1}^* \cdots \mathbf{b}_{i_r}^* : 1 \le i_1 \le \ldots \le i_r \le m\right\}$$

as a basis for $\Sigma^r(V^*)$ and

$$\left\{\mathbf{b}_{i_1}^* \wedge \ldots \wedge \mathbf{b}_{i_r}^* : 1 \le i_1 < \ldots < i_r \le m\right\}$$

as a basis for $\Lambda^r(V^*)$.

Example 4.115. ▷ Let $m = 3$. Then a basis for $\Sigma^2(V^*)$ is

$$(\mathbf{b}_1^*\mathbf{b}_1^*, \quad \mathbf{b}_1^*\mathbf{b}_2^*, \quad \mathbf{b}_1^*\mathbf{b}_3^*, \quad \mathbf{b}_2^*\mathbf{b}_2^*, \quad \mathbf{b}_2^*\mathbf{b}_3^*, \quad \mathbf{b}_3^*\mathbf{b}_3^*.\}$$

◁

Example 4.116. ▷ Let $m = 3$. Bases for $\Lambda^r(V^*)$, for $r = 0, 1, 2, 3$, respectively, are

$$\begin{aligned}
&\{1\}, \\
&\{\mathbf{b}_1^*, \quad \mathbf{b}_2^*, \quad \mathbf{b}_3^*\}, \\
&\{\mathbf{b}_1^* \wedge \mathbf{b}_2^*, \quad \mathbf{b}_1^* \wedge \mathbf{b}_3^*, \quad \mathbf{b}_2^* \wedge \mathbf{b}_3^*\}, \\
&\{\mathbf{b}_1^* \wedge \mathbf{b}_2^* \wedge \mathbf{b}_3^*\}.
\end{aligned}$$

◁

Remark 4.117. We saw that $\Lambda^r(V^*)$ has dimension $\binom{m}{r}$. This means that $\Lambda^m(V^*)$ has dimension $\binom{m}{m} = 1$. Now, the determinant is an alternating m-linear function on m vectors of length m, so it is necessarily a scalar multiple of the single basis vector for $\Lambda^m(V^*)$. In fact, that scalar multiple is 1, and the determinant may be characterized as the unique alternating m-linear function whose value on the standard basis is 1.

Example 4.118. $\triangleright$ Here we see what $\mathbf{e}_1^* \wedge \mathbf{e}_2^* \wedge \mathbf{e}_3^*$ does to three vectors on a 3-dimensional space.

$$\begin{aligned}
\mathbf{e}_1^* \wedge \mathbf{e}_2^* \wedge \mathbf{e}_3^*(\mathbf{u}, \mathbf{v}, \mathbf{w}) &= \sum_{\sigma \in \mathcal{S}_3} \mathbf{e}_1^* \otimes \mathbf{e}_2^* \otimes \mathbf{e}_3^*(\mathbf{u}, \mathbf{v}, \mathbf{w})^\sigma \\
&= \mathbf{e}_1^* \otimes \mathbf{e}_2^* \otimes \mathbf{e}_3^*(\mathbf{u}, \mathbf{v}, \mathbf{w}) + \mathbf{e}_1^* \otimes \mathbf{e}_2^* \otimes \mathbf{e}_3^*(\mathbf{v}, \mathbf{w}, \mathbf{u}) + \mathbf{e}_1^* \otimes \mathbf{e}_2^* \otimes \mathbf{e}_3^*(\mathbf{w}, \mathbf{u}, \mathbf{v}) \\
&- \mathbf{e}_1^* \otimes \mathbf{e}_2^* \otimes \mathbf{e}_3^*(\mathbf{u}, \mathbf{w}, \mathbf{v}) - \mathbf{e}_1^* \otimes \mathbf{e}_2^* \otimes \mathbf{e}_3^*(\mathbf{w}, \mathbf{v}, \mathbf{u}) - \mathbf{e}_1^* \otimes \mathbf{e}_2^* \otimes \mathbf{e}_3^*(\mathbf{v}, \mathbf{u}, \mathbf{w}) \\
&= u_1 v_2 w_3 + v_1 w_2 u_3 + w_1 u_2 v_3 - u_1 w_2 v_3 - w_1 v_2 u_3 - v_1 u_2 w_3 \\
&= u_1 v_2 w_3 - u_1 v_3 w_2 + u_2 v_3 w_1 - u_2 v_1 w_3 + u_3 v_1 w_2 - u_3 v_2 w_1 \\
&= \begin{vmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix} \\
&= \det(\mathbf{u}, \mathbf{v}, \mathbf{w}).
\end{aligned}$$

$\triangleleft$

Remark 4.119. While our most important example of an alternating tensor is the determinant, the key symmetric example is the dot product. Starting in section 6.2.1, we will use varying coefficients rather than constant coefficients. Then, in sections 6.3.10 and 6.3.9, the determinant and the dot product will become the **volume form** and the **metric tensor**, respectively.

4.7.13 Some explicit computations with tensors

In example 4.84 we decomposed a three-tensor into a linear combination of its value on basis functions. Here we will do something similar for a one-tensor. This kind of computation comes up very often in exams and qualifiers, where the tensors in question are forms on a manifold. [xxx xref forward to section xxx, once I write it.] See for example the problems in section 10.3.1 and 10.5.2.

Example 4.120. $\triangleright$ Let η be a one-tensor on V and let $m = \dim(V)$. Let V have basis $\{\mathbf{b}_1, \ldots, \mathbf{b}_m\}$ with corresponding dual basis $\{\mathbf{b}_1^*, \ldots, \mathbf{b}_m^*\}$. Then we know that η is a linear combination of the one-tensors, namely,

$$\eta = \sum_{i=1}^m n_i \mathbf{b}_i^*.$$

How do we find out what the coefficients n_i are? We can apply η to each of the basis elements $\mathbf{b}_j$. We get

$$\eta(\mathbf{b}_j) = \sum_{i=1}^m n_i \mathbf{b}_i^*(\mathbf{b}_j) = n_j$$

since $\mathbf{b}_i^*(\mathbf{b}_j) = \delta_{ij}$. So, we have the useful result

$$\boxed{\eta = \sum_{i=1}^m \eta(\mathbf{b}_i) \mathbf{b}_i^*.}$$

That is, to write our one-tensor with respect to a certain coordinate system, we just need to find its value on the basis vectors for that coordinate system. Note that a one-tensor is trivially alternating, so this formula also applies to one-forms as well. $\triangleleft$

Example 4.121. ▷ A similar computation shows that for an alternating two-tensor η, we have

$$\boxed{\eta = \sum_{i<j} \eta(\mathbf{b}_i, \mathbf{b}_j)\mathbf{b}_i^* \wedge \mathbf{b}_j^*.}$$

◁

Example 4.122. ▷ For a symmetric two-tensor ϕ, we have

$$\boxed{\phi = \sum_{i\leq j} \phi(\mathbf{b}_i, \mathbf{b}_j)\mathbf{b}_i^* \mathbf{b}_j^*.}$$

◁

Example 4.123. ▷ Likewise, for a general two-tensor ψ, we have

$$\boxed{\psi = \sum_{i,j} \psi(\mathbf{b}_i, \mathbf{b}_j)\mathbf{b}_i^* \otimes \mathbf{b}_j^*.}$$

◁

4.7.14 Change of basis; more about covariance and contravariance

Note: This section is not qualifier material.

In section 4.6.7 we saw how the coefficients of linear functionals and vectors transform when we change bases. We now know these to be tensors of type 0,1 and 1,0, respectively, so one may ask how the coefficients of general tensors transform with respect to change of basis.

First, some notation. Let

$$A = \{\mathbf{a}_1, \ldots, \mathbf{a}_m\} \quad \text{and} \quad B = \{\mathbf{b}_1, \ldots, \mathbf{b}_m\}$$

be ordered bases for an m-dimensional real vector space V, and for concreteness assume here that A is the standard basis. (As usual, my examples use $V = \mathbb{R}^m$.) As in section 4.6.7 as well as [**FIS**], let Q be the change-of-basis matrix which converts vectors from B coordinates to A coordinates. Recall that Q simply has jth column equal to $\mathbf{b}_j$ in A coordinates. Write q_{ij} for the elements of Q, and write p_{ij} for the elements of Q^{-1}.

We saw in section 4.6.7 that the coefficients of a covariant tensor λ and a contravariant tensor $\mathbf{v}$ transform according to

$$[\lambda]_B = [\lambda]_A Q \qquad \text{and} \qquad [\mathbf{v}]_B = Q^{-1}[\mathbf{v}]_A.$$

Written in terms of coefficients, these are

$$\begin{pmatrix} [\lambda_1]_B & \ldots & [\lambda_m]_B \end{pmatrix} = \begin{pmatrix} [\lambda_1]_A & \ldots & [\lambda_m]_A \end{pmatrix} \begin{pmatrix} q_{11} & \cdots & q_{1m} \\ \vdots & & \vdots \\ q_{m1} & \cdots & q_{mm} \end{pmatrix},$$

$$\begin{pmatrix} [v_1]_B \\ \vdots \\ [v_m]_B \end{pmatrix} = \begin{pmatrix} p_{11} & \cdots & p_{1m} \\ \vdots & & \vdots \\ p_{m1} & \cdots & p_{mm} \end{pmatrix} \begin{pmatrix} [v_1]_A \\ \vdots \\ [v_m]_A \end{pmatrix}$$

which is to say

$$[\lambda_j]_B = \sum_{j=1}^{m} q_{ij}[\lambda_i]_A \qquad \text{and} \qquad [v_i]_B = \sum_{j=1}^{m} p_{ij}[v_j]_A.$$

Note in particular that the covariant 1-tensor uses q_{ij}, and the entries of $[\lambda]_A$ sum along the first (row) index of Q, namely, i. Likewise, the contravariant 1-tensor uses p_{ij}, and the entries of $[\lambda]_A$ sum along the j, or column, index of Q^{-1}.

I will discuss two more examples, as motivation for the general principle.

The **dot product** is a bilinear function on V, i.e. a covariant 2-tensor. It will be convenient to name it, so write

$$\mu(\mathbf{u}, \mathbf{v}) = \mathbf{u} \cdot \mathbf{v}.$$

Given $\mathbf{u}, \mathbf{v}$ written in coordinates with respect to the standard basis, we are accustomed to writing

$$\mathbf{u} \cdot \mathbf{v} = \sum_{i=1}^{m} u_i v_i.$$

In the spirit of section 4.7.13, though, we can also write

$$\mu = \sum_{ij} \mu(\mathbf{e}_i, \mathbf{e}_j)\mathbf{e}_i^* \otimes \mathbf{e}_j^*.$$

Recalling that $\mathbf{e}_i^*$ is nothing more than the ith coordinate-selector function, we have

$$\mu(\mathbf{u}, \mathbf{v}) \quad = \quad \sum_{ij} \mu(\mathbf{e}_i, \mathbf{e}_j)\mathbf{e}_i^* \otimes \mathbf{e}_j^*(\mathbf{u}, \mathbf{v}) = \sum_{ij} (\mathbf{e}_i \cdot \mathbf{e}_j) u_i v_j = \sum_{ij} \delta_{ij} u_i v_j = \sum_i u_i v_i$$

which is what we expect. So, with respect to the standard basis, the coefficients of μ are δ_{ij}, which is the same as the identity matrix. These coefficients $\mathbf{e}_i \cdot \mathbf{e}_j$ of μ are equal to δ_{ij} precisely for orthonormal coordinates.

With respect to another basis $\{\mathbf{b}_1, \ldots, \mathbf{b}_m\}$, we have

$$\mu = \sum_{ij} \mu(\mathbf{b}_i, \mathbf{b}_j)\mathbf{b}_i^* \otimes \mathbf{b}_j^*. \tag{4.3}$$

Given the discussion in section 4.6.7, I want to state this in terms of the change-of-basis matrix Q. Since the covariant 1-tensor above had

$$[\lambda_j]_B = \sum_{j=1}^{m} q_{ij}[\lambda_i]_A,$$

and guided by my readings elsewhere, I want the coefficients of the covariant 2-tensor μ to transform by using Q twice, summing along the first index of Q in each case. That is, for the $m \times m$ array of coefficients $[\mu]_B$, I want

$$[\mu_{ij}]_B = \sum_{k,\ell=1}^{m} q_{ki} q_{\ell j} [\mu_{k\ell}]_A. \tag{4.4}$$

It remains to show that

$$\mu(\mathbf{b}_i, \mathbf{b}_j) = \sum_{k,\ell=1}^{m} q_{ki} q_{\ell j} [\mu_{k\ell}]_A.$$

But this is true since the ith and jth columns of Q are precisely $\mathbf{b}_i$ and $\mathbf{b}_j$ in A coordinates. That is, in A coordinates we would write

$$\mu(\mathbf{b}_i, \mathbf{b}_j) = \sum_{k,\ell=1}^{m} \mu(\mathbf{a}_k, \mathbf{a}_\ell)(\mathbf{a}_k^* \otimes \mathbf{a}_\ell^*)(\mathbf{b}_i, \mathbf{b}_j) = \sum_{k,\ell=1}^{m} [\mu_{k\ell}]_A q_{ki} q_{\ell j}.$$

* * *

The next example after the dot product is a tensor of mixed variance. It might seem that we don't know of any such thing. However, we do: in fact, a linear transformation is a mixed tensor. (This becomes guessable when one uses the Einstein summation convention, as in [**Lee2**] and [**Guggenheimer**], where we write $y^i = A^i_j x^j$ with superscripts being contravariant and subscripts being covariant.)

Let $T : V \to V$. Then for $\mathbf{y} = T\mathbf{x}$, taking coordinates with respect to a basis A, for the ith component of $\mathbf{y}$ we write

$$[y_i]_A = \sum_{j=1}^{m} [T_{ij}]_A [x_j]_A.$$

Writing all the components of $\mathbf{y}$ at once, we can say

$$[\mathbf{y}]_A = \sum_{ij} [T_{ij}]_A \mathbf{a}_i \otimes \mathbf{a}_j^* \left(\sum_k [x_k]_A \mathbf{a}_k \right)$$

Cavalierly manipulating symbols (I should justify this using tensor contraction, which I've not defined here), and recalling that $\mathbf{a}_j^*(\mathbf{a}_k) = \delta_{jk}$, this is

$$[\mathbf{y}]_A = \sum_{ijk} [T_{ij}]_A [x_k]_A \mathbf{a}_i \otimes \mathbf{a}_j^*(\mathbf{a}_k) = \sum_{ijk} [T_{ij}]_A [x_j]_A \mathbf{a}_i$$

which is what we would expect.

The question at hand is how the coefficients of T transform when we change bases. Let Q be the change-of-basis matrix as above. It is shown in [**FIS**] that

$$[T]_B = Q^{-1}[T]_A Q.$$

This is in fact easy to see. Above we saw that for vectors $\mathbf{v}$, Q converts from B coordinates to A coordinates. So, when we compute $(Q^{-1}[T]_A Q)([\mathbf{v}]_B)$, we are doing the following:

- Convert $\mathbf{v}$ from B coordinates to A coordinates.
- Apply the linear transformation in A coordinates.
- Convert the result back to B coordinates.
- So, the net effect is to apply the linear transformation in B coordinates.

As with the covariant 2-tensor example, trying to generalize the pattern I've been seeing so far, I would want to write

$$[T_{ij}]_B = \sum_{k\ell} p_{ik} q_{\ell j} [T_{k\ell}]_A$$

which looks a bit more familiar as

$$[T_{ij}]_B = \sum_{k\ell} p_{ik} [T_{k\ell}]_A q_{\ell j}.$$

This is precisely what the triple matrix product $Q^{-1}[T]_A Q$ is:

$$
\begin{array}{c}
\\ i \downarrow \\ \\
\end{array}
\begin{array}{c}
\overset{k}{\longrightarrow} \\
\begin{pmatrix} p_{11} & \cdots & p_{1m} \\ \vdots & & \vdots \\ p_{m1} & \cdots & p_{mm} \end{pmatrix}
\end{array}
\quad
\begin{array}{c}
\\ k \downarrow \\ \\
\end{array}
\begin{array}{c}
\overset{\ell}{\longrightarrow} \\
\begin{pmatrix} [T_{11}]_A & \cdots & [T_{1m}]_A \\ \vdots & & \vdots \\ [T_{m1}]_A & \cdots & [T_{mm}]_A \end{pmatrix}
\end{array}
\quad
\begin{array}{c}
\\ \ell \downarrow \\ \\
\end{array}
\begin{array}{c}
\overset{j}{\longrightarrow} \\
\begin{pmatrix} q_{11} & \cdots & q_{1m} \\ \vdots & & \vdots \\ q_{m1} & \cdots & q_{mm} \end{pmatrix}
\end{array}.
$$

* * *

Following this pattern, we have a general rule for change of coordinates. To recap, let T be a tensor of type s, r on an m-dimensional real vector space V. Let A and B be ordered bases for V. Let Q be the change-of-basis matrix with entries q_{ij}, and write p_{ij} for the entries of Q^{-1}. Then

$$[T^{i_1,\ldots,i_s}_{j_1,\ldots,j_r}]_B = \sum_{k_1,\ldots,k_s,\ell_1,\ldots,\ell_r} p_{i_1,k_1} \cdots p_{i_s,k_s} q_{\ell_1,j_1} \cdots q_{\ell_r,j_r} [T^{k_1,\ldots,k_s}_{\ell_1,\ldots,\ell_r}]_A.$$

In section 4.6.7 and this section I've proved this for tensors of type 1,0, type 0,1, type 1,1, and type 0,2. I will not prove this general formula, but the examples here should make it plausible and rememberable.

[xxx xref forward to varying-coefficients section.]

4.7.15 Contractions

Definition 4.124. Let V be an m-dimensional real vector space. Let ω be a k-form, i.e. $\omega \in \mathcal{T}^k(V^*)$. Let $\mathbf{u} \in V$. Define the **contraction** $\mathbf{u} \lrcorner \omega$ (also written $i_{\mathbf{u}}(\omega)$) to be

$$\mathbf{u} \lrcorner \omega(\mathbf{v}_1, \ldots, \mathbf{v}_{k-1}) = \omega(\mathbf{u}, \mathbf{v}_1, \ldots, \mathbf{v}_{k-1}).$$

Thus, if ω is a k-form, then $\mathbf{u} \lrcorner \omega$ is a $k-1$ form.

Mnemonic 4.125. We have an i in $i_{\mathbf{u}}$, and it inserts $\mathbf{u}$ into a form's argument list: i for *insert*.

Definition 4.126. Let ϕ be a function from the alternating tensor algebra $\mathcal{T}^k(V^*)$ to itself. Suppose that ϕ is linear, i.e. for all forms ω and η and for all $s \in \mathbb{R}$,

$$\phi(\omega + \eta) = \phi(\omega) + \phi(\eta) \qquad \text{and} \qquad \phi(s\omega) = s\phi(\omega).$$

Let k and ℓ be the degrees of ω and η. We say that ϕ is a **derivation** if in addition to linearity,

$$\phi(\omega \wedge \eta) = \phi(\omega) \wedge \eta + \omega \wedge \phi(\eta).$$

Likewise, we say that ϕ is an **antiderivation** if in addition to linearity,

$$\phi(\omega \wedge \eta) = \phi(\omega) \wedge \eta + (-1)^k \omega \wedge \phi(\eta).$$

Remark 4.127. These two properties are also called the **signed Leibniz rule** and the **unsigned Leibniz rule**, respectively, due to the analogy with the product rule in calculus. Also note that the *anti* in *antiderivation* has only to do with the alternating sign — it has nothing to do with antidifferentiation. (We will see in section [xxx xref] that the differentiation and antidifferentiation operators d and $\mathbf{u} \lrcorner$ are both antiderivations on tangent bundles.)

Proposition 4.128. *The contraction* $\mathbf{u} \lrcorner$ *is an antiderivation.*

Proof. ...goes here. □

xxx examples.

4.7.16 Tensor products over free modules

Everything said about linear algebra and tensor algebra in sections 4.6 and 4.7 was stated in terms of finite-dimensional vector spaces over $\mathbb{R}$. I did this for the sake of familiarity: we've all seen vector spaces before. Since this is our first year of graduate school and since we'll be getting modules in the core algebra course only in the second semester, I delayed generalizing the presentation. But everything said above for vector spaces is in fact true for **free modules** (see section 4.1.7) over commutative rings with identity as well. This will be discussed in section 6.2.1, after some more terminology is defined.

5 Topological manifolds

All art is at once surface and symbol. Those who go beneath the surface do so at their peril.
— Oscar Wilde (1854-1900).

Algebraic topology is an extensive subject. See [**Hatcher**], [**Massey**], and/or [**Lee1**] for more information. The minimum needed for our purposes is enough for the student to understand and use the following:

- classification of **surfaces**,
- the **lifting lemmas**,
- the **Seifert-van Kampen** theorem,
- the **Mayer-Vietoris** theorem, and
- cubes and chains which are needed for **integration** on manifolds (section 7).

The essential idea is that section 2.11 showed us that traditional vector calculus fails when domains have **holes** in them. Here, we simply want to understand what holes are.

5.1 CW complexes

CW complexes early on, following Hatcher? Is there any reason *not* to? It would make it possible to do classification of surfaces all at once, without having to come back.

5.2 Classification of surfaces

$\mathbb{S}^2$	$\mathbb{T}$	$\mathbb{T}^2 = \mathbb{T}\#\mathbb{T}$	$\mathbb{T}^3$	$\dots$
$P = \mathbb{P}^2$	$\mathrm{KB} = \mathbb{P}^2\#\mathbb{P}^2 = P^2$	P^3	P^4	$\dots$

[xxx alg section: quotients / relations: we do NOT use equiv classes in most actual work. rather, pick reps and use xfmn rules.]

give edge presentations (multiple ones!) and stress they have different uses. S2 with/without the edge. In each case, stress careful counting of ID'd vertices and edges. Mention free groups on strings, mod canonicalizing relations. Pound operator as concat op on edge strings.

[xxx have a revisit section at the end of topo — incl various info including χ's, univ covs ($\mathbb{S}^2$ over $\mathbb{S}^2$ and $\mathbb{P}^2$, $\mathbb{R}^2$ over everything else); fundamental groups; what else ....]

to be filed: genus has been soft-pedaled in this course ... $\chi = 2 - 2g$ ....

5.3 Homotopy and the fundamental group

5.4 Homologies

Cellular, simplicial, singular. Note I prefer singular since it gels nicely with the elementary proof of Stokes using the FTC, as in little Spivak.

5.5 Singular homology

Crucial example: latitudinal paths along $\mathbb{S}^2$: $c_1 - c_2 = \partial B$ where [insert figure with cancelling arrows] B is the image of the unit square as shown.

5.6 Cellular homology

Work some examples. This makes quick work out of $\mathbb{RP}^n$, $\mathbb{CP}^n$, and $\mathbb{T}^n$.

xxx note it's a homology theory for CW complexes so I need to have first said something about CW complexes.

5.7 Homotopy equivalence and homotopy invariance

what, and why (xxx xref fwd). Makes it easier to do SvK and MV.

Definition 5.1 (xxx cite wiki)**.** Topological spaces X and Y are **homotopy equivalent** (or of the same **homotopy type**) if there exist continuous $f : X \to Y$ and $g : Y \to X$ such that $g \circ f$ is homotopic to id_X and $f \circ g$ is homotopic to id_Y. The functions f and g are called **homotopy equivalences**.

Remark 5.2. A homeom is a h.e. but not v.v. find a ctrex.

homotopy invariance: If X and Y are homotopy equivalent, then:

- X path conn iff Y is
- X simply conn iff Y is
- sing homol and coho groups are isom
- π_k (incl fund grps) are isom as long as X, Y path-conn

5.8 Deformation retracts

what, and why (xxx xref fwd). Makes it easier to do SvK and MV.

intuition: ctsly shrink space into subspace.

5.9 Seifert-van Kampen

stt of thm (lee not'n)

torus example

claim: for surfaces, just read off the edge presentation.

5.10 Mayer-Vietoris for homology

xxx include diagrams from Notation TBD section. Connections with SvK. Emphasize importance of choice of U, V, and X — be sure their homologies are known. Xref bkwd to def retracts and homotopy invariance. Tabulate known homologies. Also note that we needn't always solve for X: sometimes it works out more nicely to solve for U. Include example from August 2006 qual.

5.11 Covering spaces

Definition 5.3. Let X and $\tilde{X}$ be topological spaces, with p a continuous surjection of $\tilde{X}$ onto X. Then the map p is said to be a **cover**, and the space $\tilde{X}$ is said to be a **covering space** of X, if each $x \in X$ has an open neighborhood U such that $p^{-1}(U)$ is the union is disjoint open sets in $\tilde{X}$, each of which is homemorphic to U. Furthermore, Pickrell requires that $\tilde{X}$ and X be connected.

Mnemonic 5.4. You can recover all of the items of the definition by the following example: send the real line $\tilde{X} = \mathbb{R}$ to $X = \mathbb{S}^1$ (viewed as the unit circle in $\mathbb{C}$) via $p(t) = e^{2\pi i t}$. The point 1 on the unit circle has a neighborhood with preimages $(n - \varepsilon, n + \varepsilon)$ for each integer n. For other points of $\mathbb{S}^1$, just shift around the circle.

Lift criterion: nice diagram here. Massey V. 5. Note the intuitive obviousness of it (at least if *both* uppers are covering spaces), given the monomorphism property of covering paths.

monomorphism property: make a nice picture of a spoked wheel over a single-dotted wheel.

Def **automorphism** in this context. Make intuitive sense of Massey's commutative diagram: homomorphisms must preserve something; what is it that a homomorphism of cov spcs needs to preserve.

no-fixed-points — state thm and find examples. Or, just as a consequence of group actions.

def univ cov spc.

action of $\pi_1(X, x)$ on $p^{-1}(x)$ — Massey is most thorough here.

Let $\phi \in \mathrm{Aut}(\tilde{X}, X)$ and $\alpha \in \pi_1(X, x_0)$. Let $\tilde{x} \in p^{-1}(x_0)$. Then

$$\phi(\tilde{x} \cdot \alpha) = (\phi\tilde{x}) \cdot \alpha$$

[xxx note this is a compatibility between left and right group actions, namely, the aut group and π_1 of the base both acting on fibers. xref to the beautiful Wikipedia article.]

Massey's cor. 7.3 — central.

$$\mathrm{Aut}(\tilde{X}, X) \cong \frac{N[p_*(\pi_1(\tilde{X}, \tilde{x}))]}{p_*(\pi_1(\tilde{X}, \tilde{x}))}.$$

natural isom? Also belabor the group-action lemma in Massey's appendix B. Phrase $p_*(\pi_1)$ as an *isotropy subgroup*.

normalizer notation. How to actually *compute* such things? Note that the normalizer of H in G is the *largest* normal subgroup of G in which H is normal. In particular, if H is already normal in G (e.g. abelian, in particular if cyclic) then $N(H)$ is all of G.

Find a good example space and covering for this normalizer business: draw up a nice picture (or set of pictures) for 8-to-1 covering of $\mathbb{S}^1$. "Stop sign". See handwritten AT final-review notes from 5-10-06.

Definition 5.5. A covering space $\tilde{X}$ of X is said to be a **regular covering** if

$$p_*(\pi_1(\tilde{X}, \tilde{x})) \lhd \pi_1(X, x).$$

(Massey 7.4) $\tilde{X}/X$ a regular covering implies that for all $x \in X$ and for all $\tilde{x} \in p^{-1}(x)$,

$$\mathrm{Aut}(\tilde{X}, X) \cong \frac{\pi_1(X, x)}{p_*(\pi_1(\tilde{X}, \tilde{x}))}.$$

since the normalizer is taken inside the parent group which is $\pi_1(X, x)$.

(Massey 7.5) Let U be a universal covering space of X. Then

$$\mathrm{Aut}(U, X) \cong \pi_1(X, x)$$

and

$$\#\pi_1(X, x) = \#p^{-1}(x)$$

(Massey 8.1) Aut group is **transitive on fibers** iff the covering is regular.

$\tilde{X}/X$ a regular covering implies X is homeom to $\tilde{X}/\mathrm{Aut}(\tilde{X}, X)$.

Definition 5.6. The action of a group G on a topological space X is **free** (we say G acts **freely** on X) if the stabilizer subgroup of each x in X is trivial, i.e. if all non-identity elements of G move all points of X.

Definition 5.7. The action of a group G on a topological space X is **properly discontinuous** if for all $x \in X$, there is a neighborhood U of x such that for all $g \in G$ other than the identity of G, $gU \cap U = \emptyset$.

5.12 Topo TBD

xxx in particular, *lift recognition*: many problems turn out to be an application of the lifting theorem; the task is to *see* them as such.

- Retract and deformation retract: emphasize recognition and computation.
- Seifert-van Kampen
- Universal covering
- $\pi_1(X)$ acting on the covering space $\tilde{X}$. (This is an application of the lifting theorem.)
- $\mathrm{Aut}(\tilde{X}/X)$ acting on the covering space $\tilde{X}$.
- Mayer-Vietoris

Also look at the same topics in this order:

- What options do we have for computing a fundamental group?
- What options do we have for computing the automorphism group of a covering space?

For MV: disconnect via (a) showing a *group* is 0, or (b) showing a *map* has zero *image*, or (c) showing a *map* has zero *kernel*.

5.13 Cubes and chains

Definition 5.8. k-cube ....

Definition 5.9. k-chain ....

5.14 The boundary operator ∂

Make sure to emphasize the sign which alternates with variable index.

5.15 Homology

6 Smooth manifolds

Life is a short affair; we should try to make it smooth, and free from strife.
— Euripides (c. 480-406 B.C.).

6.1 Manifolds

6.1.1 Manifolds

Definition 6.1. A **manifold** is a metric space with the property that for all $q \in M$, there is some neighborhood U of q and some non-negative integer $m(q)$ such that U is homeomorphic to $\mathbb{R}^m$. If $m(q)$ is the same for all $q \in M$, then we say that M has **dimension** m.

For this course we consider only connected manifolds. In particular, all our manifolds have a dimension. (For a counterexample, consider the plane $z = 1$ in $\mathbb{R}^3$ along with the line $y = z = 0$. This has two disconnected components: the plane with dimension 2 and the line with dimension 1.)

6.1.2 Coordinates and parameterizations

Definition 6.2. Let M be an m-dimensional manifold. A **coordinate chart** is a homeomorphism ϕ from an open subset U of M to $\mathbb{R}^m$. (Formally, the chart is the ordered pair U, ϕ.)

A manifold by definition has such homeomorphisms for every point in the manifold, so we are just naming them. Also, homeomorphisms are invertible by definition, so it also makes sense to name their inverses.

Definition 6.3. Let M be an m-dimensional manifold. A **parameterization** is the inverse of a coordinate chart.

Definition 6.4. A **transition function** on a manifold M is of the form $y \circ x^{-1}$ where x and y are coordinate charts on a manifold M.

Note that a transition function is a homeomorphism from $\mathbb{R}^m$ to $\mathbb{R}^m$.

Example 6.5. $\triangleright$ There are several different coordinate charts we can put on the unit sphere $\mathbb{S}^2$.

- We use **graph coordinates** when we solve for one rectangular coordinate in terms of the others. E.g. off the equator we can solve for z in terms of x and y. (Here we are implicitly using the implicit function theorem, theorem 3.5.) An example coordinate chart around the north pole is

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} x \\ y \end{pmatrix}$$

  with parameterization

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x \\ y \\ \sqrt{1 - x^2 - y^2} \end{pmatrix}.$$

- We use **spherical coordinates** which work everywhere except the north and south poles. (Why? Well, a homeomorphism must be one-to-one. Points with $\phi = 0$ and *any* value of θ all correspond to the north pole. Likewise for the south pole and $\phi = \pi$.) Spherical coordinates for $\mathbb{R}^3$ are

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \rho\cos\theta\sin\phi \\ \rho\sin\theta\sin\phi \\ \rho\cos\phi \end{pmatrix}.$$

  So, when we restrict to $\mathbb{S}^2$ (where $\rho = 1$), the coordinate chart is

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} \theta \\ \phi \end{pmatrix}$$

  with parameterization

$$\begin{pmatrix} \theta \\ \phi \end{pmatrix} \mapsto \begin{pmatrix} \cos\theta\sin\phi \\ \sin\theta\sin\phi \\ \cos\phi \end{pmatrix}.$$

  The domain of this coordinate chart is any open set as big as the open set consisting of all of $\mathbb{S}^2$ except the poles and some meridian. If we were to include a meridian, then the resulting open set would not be homeomorphic to a subset of $\mathbb{R}^m$.

- **Stereographic coordinates** are discussed in section [xxx homework problem early on]. (That section also gives an example of transition functions.) Here, we have two coordinate charts, each of which cover all but *one* point of $\mathbb{S}^2$.

$\triangleleft$

One would want the transition maps $y \circ x^{-1}$ to be smooth. The notions of **atlas**, **maximal atlas**, and **differentiable structure** are important foundationally, but do not arise in practice in problems given in this course. See any of the geometry texts in the bibliography for definitions.

6.1.3 Maps of manifolds

In section 3.2, we defined smoothness for functions from $\mathbb{R}^m$ to $\mathbb{R}^n$. To apply this notion to manifolds, we can use coordinate charts and parameterizations. Let M and N be manifolds of dimension m and n, respectively. Let x and y be coordinate charts on M and N, respectively. If f is a function from M to N, then we want to say that f is smooth in the manifold sense if $y \circ f \circ x^{-1} : \mathbb{R}^m \to \mathbb{R}^n$ is smooth in the Euclidean-space sense. Of course, we have to consider what happens if the range of $f \circ x^{-1}$ doesn't intersect the domain of y, and whether this manifold smoothness is independent of the choice of x and y. This can in fact be done.

Definition 6.6. Let $f : M \to N$ be a map of manifolds. A **coordinate expression** for f is $y \circ f \circ x^{-1}$, where x and y are coordinate functions on M and N, respectively.

Definition 6.7. A map of manifolds $f : M \to N$ is **differentiable** if for all coordinates (x, U) of M and (y, V) of N, $y \circ f \circ x^{-1}$ is a differentiable function from $\mathbb{R}^m$ to $\mathbb{R}^n$.

Definition 6.8. A map of manifolds $f : M \to N$ is **smooth** if for all coordinates (x, U) of M and (y, V) of N, $y \circ f \circ x^{-1}$ is a smooth function from $\mathbb{R}^m$ to $\mathbb{R}^n$.

Definition 6.9. A map of manifolds $f : M \to N$ is a **diffeomorphism** if it is a smooth map of manifolds which is also bijective.

Of course we could say that diffeomorphisms need only be bijective differentiable (i.e. C^1) maps, but for this course we use the same convention as in section 3.3.

6.1.4 Immersion and embedding

Definition 6.10. Let $f : M \to N$ be a map of manifolds. We say f is an **immersion** at $\mathbf{q} \in M$ if $Df_{\mathbf{q}}$ is 1-1 (i.e. has zero kernel). If f is an immersion at all points $\mathbf{q} \in M$, then we simply say that f is an immersion.

Definition 6.11. Let $f : M \to N$ be a map of manifolds. We say that f is an **embedding** if:

- It is an immersion, and
- it is a homemorphism onto its image.

Remark 6.12. An immersion is locally 1-1. An embedding is globally 1-1.

Example 6.13. $\triangleright \ldots$ $\triangleleft$

Example 6.14. $\triangleright \ldots$ $\triangleleft$

Definition 6.15. Let $M \subseteq N$ be manifolds. Then the inclusion map i is a map of manifolds. We say that M is an **immersed submanifold** of N if the inclusion map i is an immersion. Likewise, we say that M is an **embedded submanifold** of N if the inclusion map is an embedding.

6.1.5 Regular and critical values

Definition 6.16. Let $f : M \to N$ be a map of manifolds. Then $\mathbf{a} \in M$ is said to be a **regular point** of f if the local derivative of f at $\mathbf{a}$ is surjective. [xxx draw a picture: x is a coordinate function on M and y is a coordinate function on N.] That is, f is regular at $\mathbf{a}$ if $D(yfx^{-1})|_{x(\mathbf{a})}$ has rank n where $n = \dim(N)$. [xxx note that we need to prove (or cite) that this is independent of the choice of coordinates. here or elsewhere: soap-box about the fact that much of the fuss in differential geometry is showing that this or that is independent of choice of coordinates. For this paper, such fuss will at most be deferred to an appendix or a citation.]

We then define the following related terms:

- A point in M is a **critical point** of f if it is not regular.
- A point $\mathbf{c} \in N$ is a **critical value** of f if any of the points in $f^{-1}(\mathbf{c})$ are critical points of f, i.e. if $\mathbf{c}$ is the image of a critical point.
- A point $\mathbf{c} \in N$ is a **regular value** of f if all the points in $f^{-1}(\mathbf{c})$ are regular points of f.

Note that the *points* are in the source manifold M; the *values* are in the destination manifold N.

[xxx draw a nice picture here.]

Theorem 6.17 (regular value theorem)**.** *Let M and N be manifolds of dimension m and n, respectively, and let $f : M \to N$ be a map of manifolds. If $\mathbf{c} \in N$ is a regular value of f, then either $f^{-1}(\mathbf{c}) = \emptyset$ or $f^{-1}(\mathbf{c})$ has the structure of an embedded submanifold of M. In the latter case $f^{-1}(\mathbf{c})$ has dimension $m - n$.*

Remark 6.18. For the last sentence we can intuitively think of m as the number of variables and n as the number of constraints, leaving $m - n$ degrees of freedom.

Remark 6.19. Note that [**Lee2**] refers to this as the *regular level set theorem.*

Remark 6.20. See section 10.6.3 for an example which shows that the converse of the regular value theorem does not hold.

Remark 6.21. Manifolds are often presented as embedded submanifolds of Euclidean space (e.g. $\mathbb{S}^2$ inside $\mathbb{R}^3$). The D in definition 6.16 must be restricted to the manifold. Either (1) compute D in the parent space and restrict as in section 1.3.2, or (2) use Lagrange multipliers as described in section 2.7. For an example using both techniques, see section 10.7.2.

zzzz

[xxx needs an example.]

xxx define codimension.

xxx move this to another section further on:

Remark 6.22. At a regular point $\mathbf{q} \in M$ of f, f_* has full rank n. The linearization $Df|_{\mathbf{q}}$ (which is $f_*|_{\mathbf{q}}$) is a map from the m-dimensional space $TM|_{\mathbf{q}}$ to the n-dimensional space $TN|_{\mathbf{c}}$. Therefore, by the rank-nullity theorem, $Df|_{\mathbf{q}}$ has kernel dimension $m - n$. [xxx need to fix this b0rk3n statement.] Let $f : M \to N$ be a map of manifolds. Let $\mathbf{q} \in M$ and $\mathbf{c} \in N$ be such that $\mathbf{c} = f(\mathbf{q})$. Assume that $\mathbf{c}$ is a regular value of f but do not assume that f is 1-1.

$$0 \quad \longrightarrow \quad \ker(f_*) \cong Tf^{-1}(\mathbf{c}) \quad \longrightarrow \quad TM|_{\mathbf{q}} \quad \xrightarrow{f_*} \quad TN_{\mathbf{c}} \quad \longrightarrow \quad 0.$$

xxx move this to the linear-algebra section:

Proposition 6.23 (adjugate criterion)**.** *An $n \times m$ matrix (with $m > n$) has full rank n iff not all $n \times n$ determinants are zero.*

Proof. This is an exercise. See section 10.1.1. [This can't be quite right — the exercise is for square matrices; this is for non-square. How do we generalize it?] □

xxx note how cramer's rule is related to exterior algebras. The proof of Cramer's rule is elementary and easy *if* you already have the formula — but how would a person have thought of it? Appeal to exterior algebra?

Example 6.24. ▷ Insert an example here, or xref. ◁

6.2 Vector bundles

6.2.1 Varying coefficients

All the vectors, linear functionals, tensors, etc. discussed up until now have had **constant coefficients**, i.e. coefficients in $\mathbb{R}$. Now picture the Euclidean space $V = \mathbb{R}^m$ evolving with time t. (Or, for $V = \mathbb{R}^2$, one may think of a loaf of bread, where each slice of bread is a vector space and the parameter t indexes the slices of bread. Of course, the loaf of bread is infinitely long, with slices that are infinitely thin and infinitely wide.)

- In this collection of vector spaces, you could imagine a vector evolving in time. Using the other metaphor, this would be a collection of vectors, one vector in each slice, with **varying coefficients**, depending smoothly on t. We could label the vector in the tth slice $\mathbf{v}(t)$, or, looking ahead to the notation used in differential geometry, $\mathbf{v}_t$. For example, in $\mathbb{R}^2$ we might have $\mathbf{v}_t = (\cos t, \sin t)$. We could label the collection $\mathbf{v}$.
- If, in the sense of section 4.4, we form a **projection** π from $V \times \mathbb{R}$ to $\mathbb{R}$ by $\pi(\mathbf{v}, t) = t$, then we can think of $\mathbf{v}$ as a **section** of π. (We will be calling it a **vector field**).
- We can write this as $V \times \mathbb{R}$, but it is not the same as $\mathbb{R}^{m+1}$. It doesn't make sense to add one time's $\mathbf{u}_{t_1}$ to another time's $\mathbf{v}_{t_2}$, but it does make sense to add vectors in the same time slice.
- Here the coefficients are not just in $\mathbb{R}$; they're smooth functions, i.e. elements of $C^\infty(\mathbb{R})$. We might think of this $V \times \mathbb{R}$ as a vector space over $\mathbb{R}$, but it sure feels like it ought to be some sort of space over $C^\infty(\mathbb{R})$ as well. (We will be calling it a **vector bundle** over $\mathbb{R}$.)
- Non-zero real numbers are invertible, so $\mathbb{R}$ is a field. Not all non-zero elements of $C^\infty(\mathbb{R})$ can be reciprocated: for a function $f \in C^\infty(\mathbb{R})$ to be non-zero, it must have a non-zero value for *some* t; to be reciprocable, it would have to be non-zero for *all* t. However, $C^\infty(\mathbb{R})$ is a commutative ring with identity. Furthermore, it may be shown (see section 6.2.2) that this $V \times \mathbb{R}$ is a **free module** over $C^\infty(\mathbb{R})$. Remember from section 4.1.7 that every element of a free module M over a ring R is uniquely expressible as an R-linear combination of basis elements. So, we get everything we would want from a vector space, except division by scalars.
- We can also form mixed tensors here: the tensor product, which we examined for vector spaces, can also be formed on modules. (This kind of construction will be the subject of most of the rest of this paper.) Here, the coefficients of the tensors won't be real numbers; instead, they will be smooth functions of the parameter t.
- If we select a particular value of t, then we do have a plain old vector space over $\mathbb{R}$, where all the vectors, tensors, etc. have constant coefficients which are simply the coefficient functions evaluated at t.

In this picture, constructions on V (i.e. the space, its dual, higher-order tensors, etc.) have coefficients in $\mathbb{R}$, while constructions in $V \times \mathbb{R}$ have coefficients in $C^\infty(\mathbb{R})$:

$$\begin{array}{ccc} V & \qquad & V \times \mathbb{R} \\ | & & | \\ \mathbb{R} & & C^\infty(\mathbb{R}) \end{array}$$

Note also instead of parameterizing by a single variable t, we could attach a copy of V to every point (x, y) of $\mathbb{R}^2$. In that case, scalars (coefficients) would be of the form $f(x, y)$ for $f \in C^\infty(\mathbb{R}^2, \mathbb{R})$.

6.2.2 Definition of vector bundle

Motivated by the preceding $V \times \mathbb{R}$ example, what properties would we want a vector bundle to have?

- Every time t should have a copy of V attached to it.
- Even though we took V to be $\mathbb{R}^m$, our $V \times \mathbb{R}$ should not be the same as $\mathbb{R}^{m+1}$. It should only make sense to add vectors from the same time slice t.
- Coefficients on vector fields $\mathbf{v}$ should vary smoothly with time.
- The vector spaces V_t themselves should be connected to one another smoothly with time.

Here is the formal definition (following [**Lee2**]).

Definition 6.25. Let E and M be manifolds, and let $\pi : E \to M$ be a surjective smooth map of manifolds. (Note that π is a **projection** in the sense of section 4.4.) Then E, M, and π form a **vector bundle** of **rank** k if the following conditions are satisfied:

- For each $\mathbf{q} \in M$, the **fiber** (see definition 4.42) $\pi^{-1}(\mathbf{q})$ is a k-dimensional real vector space. Write $E_{\mathbf{q}}$ for the fiber $\pi^{-1}(\mathbf{q})$. Note that since π is surjective, every point $\mathbf{q}$ of M has a fiber over it. We call M the **base space**.
- For each $\mathbf{q} \in M$, there exist an open subset U of M and a diffeomorphism $\Phi : \pi^{-1}(U) \to U \times \mathbb{R}^k$. This means that the fibers vary smoothly as we move around open sets in the base space.
- For the $\mathbf{q}$, U, and Φ in the previous item, the following diagram commutes:

$$\begin{tikzcd} \pi^{-1}(U) \arrow[rr, "\Phi"] \arrow[dr, "\pi"'] & & U \times \mathbb{R}^k \arrow[dl, "\pi_1"] \\ & U & \end{tikzcd}$$

  where π_1 is the projection onto the first component. If $\mathbf{e}$ is a point of E lying above $\mathbf{q}$, then $\pi(\mathbf{e}) = \mathbf{q}$, whereas if we take $\Phi(\mathbf{e})$ and then project down to M, we have to get the same base point $\mathbf{q}$. So, all this means is that Φ respects which fiber lies over which base point.
- Furthermore, for each $\mathbf{q}' \in U$ as in the previous item, Φ restricted to the fiber $E_{\mathbf{q}'}$ is a vector-space isomorphism between $E_{\mathbf{q}'}$ and $\{\mathbf{q}'\} \times \mathbb{R}^k$, which in turn is isomorphic to $\mathbb{R}^k$. Sometimes we state this condition by saying that Φ is **linear on fibers**.

Some more terminology:

Recall from definition 4.43 that a **section** of π is a function $s : M \to E$ such that $\pi \circ s$ is the identity on M, which simply means that a section maps base points into their own fibers. As in section 4.43, whenever we have projections and sections, we want those maps to have whatever properties are appropriate for the category in question. Here, this means that projections and sections are taken to be **smooth**.

Definition 6.26. A **vector field** is a section of vector bundle.

That is, a section is something of the form described in section 6.2.1, i.e. a vector whose coefficients depend smoothly on the base point $\mathbf{q}$.

Definition 6.27. Given a vector field $\mathbf{v}$ on a manifold M, for each point $\mathbf{q}$ of M there is a vector $\mathbf{v_q}$. We say that $\mathbf{v_q}$ has its **foot** at $\mathbf{q}$.

Definition 6.28. Sometimes we write $\mathbf{v}|_{\mathbf{q}}$ instead of $\mathbf{v_q}$. We do this particularly when the thing being subscripted (or the subscript itself, or both) is a complicated expression, has other subscripts, etc. See also notation 6.33.

Recall from definition 4.44 that we write $\Gamma(E)$ for all the sections of E, or the **space of sections** of E. This is actually makes $\Gamma(E)$ a **module** over $C^\infty(M, \mathbb{R})$, as was alluded to in section 6.2.1:

- Given two sections, or vector fields, $\mathbf{v}_1, \mathbf{v}_2 : M \to E$, we have vector-plus-vector addition of the form $(\mathbf{v}_1 + \mathbf{v}_2)|_{\mathbf{q}} = \mathbf{v}_1|_{\mathbf{q}} + \mathbf{v}_2|_{\mathbf{q}}$.
- Given a "scalar" $f \in C^\infty(M, \mathbb{R})$ and $\mathbf{v} \in \Gamma(E)$, we have multiplication by the scalar of the form $(f\mathbf{v})|_{\mathbf{q}} = f|_{\mathbf{q}}\mathbf{v}|_{\mathbf{q}}$.

Thus we have the following analogy between vector spaces and vector bundles:

$$\begin{array}{ccc} V & \qquad & \Gamma(E) \\ | & & | \\ \mathbb{R} & & C^\infty(M, \mathbb{R}) \end{array}$$

6.2.3 Examples of vector bundles

Let M be a manifold.

Example 6.29. $\triangleright$ For any k, we can attach a copy of $\mathbb{R}^k$ at each point $\mathbf{q}$ of M. This is called a **product bundle**. $\triangleleft$

Example 6.30. $\triangleright$ For a manifold which is, say, the level set of some function from $\mathbb{R}^m$ to $\mathbb{R}$, we can define a **normal bundle**, which at each point $\mathbf{q}$ of M consists of all the vectors in the ambient space $\mathbb{R}^m$ which are normal to M at $\mathbf{q}$. For example, for the sphere $\mathbb{S}^2$, the normal bundle has rank 1, where the fiber above a point $\mathbf{q}$ of $\mathbb{S}^2$ is the line normal to the sphere at $\mathbf{q}$. $\triangleleft$

Example 6.31. $\triangleright$ Similarly, in the same context, we can also define a **tangent bundle**, where the vector space attached to each point $\mathbf{q}$ of M consists of the space perpendicular to the normal space. For example, on $\mathbb{S}^2$, this is of rank 2 where fibers are tangent planes. $\triangleleft$

Tangent bundles are central to the geometry course. In section 6.3, they will be defined in a way that is independent of an ambient space.

6.2.4 Trivializations of vector bundles

Definition 6.32. A diffeomorphism Φ, defined on an open subset U of M as in section 6.2.2, is called a **local trivialization** of the vector bundle. If there is such a diffeomorphism Φ defined on all of M, then it

is said to be a **global trivialization** of the vector bundle. If a vector bundle has a global trivialization, the bundle is said to be **trivial**.

Notation 6.33. Pickrell adopts the convention that $\mathbf{v_q}$ may be written as the ordered pair $(\mathbf{v}, \mathbf{q})$ only for trivial bundles.

There is a criterion to determine whether a vector bundle is trivial. It requires a few definitions.

Definition 6.34. Let (E, M, π) be a rank-k vector bundle. Given an open subset U of M, a **local frame** is an ordered set $\{\mathbf{b}_1, \ldots, \mathbf{b}_k\}$ of k (smooth) sections of the vector bundle such that for all $\mathbf{q} \in U$, the set $\{\mathbf{b}_1|_{\mathbf{q}}, \ldots, \mathbf{b}_k|_{\mathbf{q}}\}$ forms a basis for the fiber $E_{\mathbf{q}}$. If there is such an ordered set defined on all of M, then it is called a **global frame**.

Lee proves the following in chapter 5 of [**Lee2**]. Pickrell uses this criterion implicitly: when he asks you to compute an explicit trivialization, what he wants is a global frame.

Proposition 6.35 (Global frame criterion)**.** *A vector bundle is trivial iff it admits a global frame.*

For an example, see section 10.2.1. [xxx that problem uses terminology from sections downstream from here.]

Note that the product bundle (example 6.29) is certainly trivial: take the standard basis for each $\mathbb{R}^k$ to be the global frame.

6.3 The tensor bundle

In this section we define tangent vectors, then build a tensor algebra on top of that. This will encapsulate many concepts, including derivatives, functions, functionals, area forms, and dot products.

6.3.1 Notions of tangent vectors

There are many different definitions of tangent space given in [**Spivak2**], [**Conlon**], and [**Abr**]. If I count correctly, [**Spivak2**] alone presents no less than 5 definitions, and proves them to be equivalent. This overabundance of definitions is a great source of confusion for the learner.

For Pickrell's course, two notions are actually used reguarly in computations, with a third notion appearing briefly early on. (Namely, this third notion is tangent vectors as equivalence classes of curves. This is discussed in section 6.3.3). As well, these are precisely the two notions that [**Lee2**] emphasizes. I will follow that approach. Namely:

- If we have a surface explicitly embedded in $\mathbb{R}^k$ for some k, then we can think of the tangent space in the usual way from vector calculus. Namely, we compute the normal vector to the surface at each point, then find the space of vectors perpendicular to the normal. These are called **geometric tangent vectors**. Such tangent vectors are **extrinsic** in that they stick off the manifold: for example, consider tangent planes on a sphere. This works fine when our manifold is already embedded into a higher-dimensional Euclidean space in some obvious way.
- The standard example of why geometric tangent vectors are insufficient is the spacetime of the universe itself: regardless of whether it *could* be visualized as embedded in some higher-dimensional space, we might not *want* to do so, particularly when our manifold is already three-dimensional. (Trying to draw three-dimensional pictures on a chalkboard is hard enough!) The more general way to think of the tangent space is as the space of **directional derivatives**. These are **intrinsic** in that we don't need to visualize anything sticking off the manifold.

6.3.2 Geometric tangent vectors

Geometric tangent vectors are just what you would expect from vector calculus ([**Anton**], [**HHGM**]). In the case that our manifold M is the level set of a single equation, we can get the tangent space as a two-step process:

(1) Compute the normal space to the surface. If the surface is the level set of a function F, then the normal vector to a point $\mathbf{q} \in M$ is $\nabla(F)|_{\mathbf{q}}$.

(2) Do linear algebra to compute the perpendicular space, which is the tangent space.

This technique is time-consuming; a more efficient technique is presented in section 6.3.4. An example comparing both methods appears in section 10.6.5.

6.3.3 Tangent vectors

Here I want to accomplish two things: make some sense of Pickrell's choice of assigned problems, and intuitively connect geometric tangent vectors with tangent vectors as directional derivatives. The latter is made quite clear in chapter 3 of [**Lee2**], so I will confine myself to presenting motivation and examples.

Preliminary #1. Suppose M is the Euclidean space $\mathbb{R}^m$ and let $f : \mathbb{R}^m \to \mathbb{R}$. Then, given a point $\mathbf{q}$ of M, it is reasonable to equate a vector $\mathbf{v_q}$ footed at $\mathbf{q}$ with the **directional derivative**, taken on functions from $\mathbb{R}^m$ to $\mathbb{R}$, in the direction of $\mathbf{v}$. This directional derivative is written

$$Df|_{\mathbf{q}}(\mathbf{v}) = \frac{d}{dt}\bigg|_{t=0} f(\mathbf{q} + t\mathbf{v})$$

The left-hand side is the derivative of f at $\mathbf{q}$ in the direction of $\mathbf{v}$. One can show that there is in fact an isomorphism identifying each $\mathbf{v_q}$ with the directional derivative at $\mathbf{q}$ in the direction of $\mathbf{v}$.

Preliminary #2. Let $\mathbf{x} = \mathbf{q} + t\mathbf{v}$, i.e. for $i = 1, \ldots, m$, $x_i = q_i + tv_i$, and let $f \in C^\infty(\mathbb{R}^m, \mathbb{R})$. Then $f(\mathbf{x})$ is a function of t. By the chain rule, we have

$$\frac{d}{dt}\bigg|_{t=0} f(\mathbf{q} + t\mathbf{v}) = \sum_{i=1}^{m} \frac{\partial f}{\partial x_i}\frac{dx_i}{dt} = \sum_{i=1}^{m} \frac{\partial f}{\partial x_i} v_i.$$

In particular, take $\mathbf{v}$ to be a standard basis vector $\mathbf{e}_j$. Then the above gives

$$\frac{d}{dt}\bigg|_{t=0} f(\mathbf{q} + t\mathbf{e}_j) = \sum_{i=1}^{m} \frac{\partial f}{\partial x_i}\frac{dx_i}{dt} = \frac{\partial f}{\partial x_j}.$$

Preliminary #3. Now let M be a manifold not necessarily $\mathbb{R}^m$. Note that in the previous step, any other curve through $\mathbf{q}$ in the direction of $\mathbf{v}$ which is equal to $\mathbf{q} + t\mathbf{v}$ up to first order would have produced the same result. The intuitive example is $M = \mathbb{S}^2$. Then tangent lines of the form $\mathbf{q} + t\mathbf{v}$ stick off $\mathbb{S}^2$, but we can have a curve through $\mathbf{q}$, heading off in the direction of $\mathbf{v}$, but staying on the sphere. (This notion makes an appearance in some problems early on in the course. See for example section 10.1.1.) We could use this intuition to motivate a definition of tangent vectors as equivalence classes of curves which agree up to first order. However, we do not in fact define them that way.

Definition 6.36. We now define **tangent vectors** as follows.

- Since we have an m-dimensional manifold M, at each point $\mathbf{q}$ of M there is an open set U with a coordinate chart ϕ which is a homeomorphism from M to $\mathbb{R}^m$.
- From step 1 we know that in $\mathbb{R}^m$ there is an isomorphism between $\mathbf{v_q}$'s and directional derivatives at $\mathbf{q}$ in the direction of $\mathbf{v}$.
- In particular, using that isomorphism we can identify each $\mathbf{e}_j$ in $\mathbb{R}^m$ with the directional derivative $\partial/\partial x_j$.
- We use that isomorphism followed by the parameterization ϕ^{-1} (the inverse of the coordinate chart ϕ) to map the standard basis for $\mathbb{R}^m$ back to an open subset U of M. (See Lee's proposition 3.6 for a technical detail.)

In summary, the tangent space at a point $\mathbf{q}$ of M has a basis, in coordinates, which is the set

$$\left\{ \frac{\partial}{\partial x_1}\bigg|_{\mathbf{q}}, \ldots, \frac{\partial}{\partial x_m}\bigg|_{\mathbf{q}} \right\}$$

Definition 6.37. Let M be a manifold, and let $\mathbf{q} \in M$. The **tangent space** of M at $\mathbf{q}$, written $T_{\mathbf{q}}M$ or $TM|_{\mathbf{q}}$, consists of all **directional derivatives**, applied to $C^\infty(M) = \mathcal{T}_0^0(M)$.

Definition 6.38. The **tangent bundle** of M is the disjoint union

$$TM = \coprod_{\mathbf{q} \in M} T_{\mathbf{q}} M.$$

Intuition 6.39. Disjointification is a way to make otherwise equal things unequal. For example, take two copies of the integers, one colored red and one colored blue. Then red 3 is unequal to blue 3. Two copies of $\mathbb{Z}$ have been disjointified by the color label. Here, we have many tangent spaces, which are disjointified by the foot label.

6.3.4 Explicit computations of tangent vectors

For problems appearing in Pickrell's course, we start with the directional-derivative notion of tangent vector and end up with a geometric tangent vector. An example will help illustrate.

Example 6.40. $\triangleright$ Let $M = \mathbb{S}^2$. A point $\mathbf{q}$ of $\mathbb{S}^2$ is (in rectangular and spherical coordinates, respectively)

$$\mathbf{q} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \cos\theta \sin\phi \\ \sin\theta \sin\phi \\ \cos\phi \end{pmatrix}.$$

The tangent space is spanned by the two (orthogonal) basis vectors $\partial/\partial\theta$ and $\partial/\partial\phi$. Now, we make the key observation that $\partial/\partial\theta$ and $\partial/\partial\phi$ are directional derivative operators, acting on $C^\infty(\mathbb{S}^2)$, while the x, y, and z coordinates are in fact smooth functions of the point $\mathbf{q}$. So, we *apply* $\partial/\partial\theta$ and $\partial/\partial\phi$ to the coordinate functions x, y, and z:

$$\begin{aligned} \frac{\partial}{\partial\theta}\begin{pmatrix} x \\ y \\ z \end{pmatrix} &= \frac{\partial}{\partial\theta}\begin{pmatrix} \cos\theta \sin\phi \\ \sin\theta \sin\phi \\ \cos\phi \end{pmatrix} = \begin{pmatrix} -\sin\theta \sin\phi \\ \cos\theta \sin\phi \\ 0 \end{pmatrix}; \\ \frac{\partial}{\partial\phi}\begin{pmatrix} x \\ y \\ z \end{pmatrix} &= \frac{\partial}{\partial\phi}\begin{pmatrix} \cos\theta \sin\phi \\ \sin\theta \sin\phi \\ \cos\phi \end{pmatrix} = \begin{pmatrix} \cos\theta \cos\phi \\ \sin\theta \cos\phi \\ -\sin\phi \end{pmatrix}. \end{aligned}$$

Observe that when we do this, the results no longer look like a directional-derivative operators: they are just geometric tangent vectors. The isomorphism of section 6.3.3 goes unspoken; we write

$$\frac{\partial}{\partial\theta} = \begin{pmatrix} -\sin\theta \sin\phi \\ \cos\theta \sin\phi \\ 0 \end{pmatrix} \qquad \text{and} \qquad \frac{\partial}{\partial\phi} = \begin{pmatrix} \cos\theta \cos\phi \\ \sin\theta \cos\phi \\ -\sin\phi \end{pmatrix}.$$

$\triangleleft$

Here is another example.

Example 6.41. $\triangleright$ [Surface of revolution.] Let $g(y) > 0$ be a smooth function. Let M be the surface of revolution obtained by revolving $g(y)$ about the y axis. Then a point $\mathbf{q}$ of M has rectangular coordinates

$$\mathbf{q} = \begin{pmatrix} g(y)\cos\theta \\ y \\ g(y)\sin\theta \end{pmatrix}.$$

Now, we can parameterize this surface using y and θ, where θ is the angle from the x axis to $\mathbf{q}$. Thus, the tangent space will be spanned by $\partial/\partial y$ and $\partial/\partial\theta$. Proceeding as in example 6.40, we apply $\frac{\partial}{\partial y}$ and $\frac{\partial}{\partial\theta}$ to

the x, y, and z coordinates of $\mathbf{q}$ to obtain

$$\frac{\partial}{\partial y} = \begin{pmatrix} g'(y)\cos\theta \\ 1 \\ g'(y)\sin\theta \end{pmatrix}, \quad \frac{\partial}{\partial \theta} = \begin{pmatrix} -g(y)\sin\theta \\ 0 \\ g(y)\cos\theta \end{pmatrix}.$$

$\lhd$

6.3.5 F_* and DF

Let $F : M \to N$ be a map of manifolds of dimensions m and n, respectively. Then we would like to define a **linearization** of F at $\mathbf{q} \in M$ as a function on the respective tangent spaces, i.e. a linear approximation to F at $\mathbf{q}$. How can we do this? Well, we saw in section 6.3.3 that the tangent space $T_{\mathbf{q}}M$ is spanned by the $\partial/\partial x_i$'s. An arbitrary element X of the tangent space is just a linear combination of those. A map on the tangent spaces is defined by what it does to the X's. And what do the X's do? They operate on functions from a manifold to $\mathbb{R}$. So ([**Lee2**], chapter 3) we make the following definition.

Definition 6.42. Let $F : M \to N$ be a map of manifolds of dimensions m and n, respectively. Let $\mathbf{q} \in M$. Let $X \in T_{\mathbf{q}}M$ and $f : N \to \mathbb{R}$. Recall that the pullback F^*f is $f \circ F$. Then

$$F_* : T_{\mathbf{q}}M \to T_{F(\mathbf{q})}N$$

is defined by

$$\boxed{(F_*X)(f) = X(F^*f).}$$

Note the following diagram of the situation:

$$\begin{tikzcd}
X & T_{\mathbf{q}}M \arrow[rr, "F_*"] & & T_{F(\mathbf{q})}N & F_*X \\
\mathbf{q} & M \arrow[rr, "F"] \arrow[dr, "F^*f"'] & & N \arrow[dl, "f"] & F(\mathbf{q}) \\
& & \mathbb{R} & &
\end{tikzcd}$$

There are several points:

- A map $f : N \to \mathbb{R}$ has been **pulled back**, via F, from N to M.
- A tangent vector X has been **pushed forward**, via F_*, from $T_{\mathbf{q}}M$ to $T_{F(\mathbf{q})}N$.
- This is a **coordinate-free** definition of F_*.
- As shown in [**Lee2**], whenever we do use coordinates, this map F_* is represented by the Jacobian matrix
$$DF = \left(\frac{\partial F_i}{\partial x_j}\right)_{i,j=1,\ldots,m}.$$

6.3.6 The cotangent bundle

Definition 6.43. Let M be a manifold, and let $\mathbf{q} \in M$. The **cotangent space** of M at $\mathbf{q}$, written $T^*_{\mathbf{q}}M$ or $T^*M|_{\mathbf{q}}$, is the dual space of $T_{\mathbf{q}}M$. It consists of all linear functionals, or **covectors**, on $T_{\mathbf{q}}M$. The **cotangent bundle** of M is the disjoint union

$$T^*M = \coprod_{\mathbf{q}\in M} T^*_{\mathbf{q}}M.$$

Just as in section 4.6.3, given a basis we can form a **dual basis**. We saw in section 6.3.3 that a basis for $T_{\mathbf{q}}M$, in coordinates, is given by

$$\left\{ \frac{\partial}{\partial x_1}\bigg|_{\mathbf{q}}, \ldots, \frac{\partial}{\partial x_m}\bigg|_{\mathbf{q}} \right\}.$$

The dual basis elements are

$$\left\{ d\,x_1\bigg|_{\mathbf{q}}, \ldots, d\,x_m\bigg|_{\mathbf{q}} \right\}$$

and they are defined, just as in section 4.6.3, by

$$dx_i\left(\frac{\partial}{\partial x_j}\right) = \delta_{ij}.$$

That is, these are nothing more than coordinate-selector functions, in the sense of section 4.6.3, on the tangent space.

Just as for frames (see definition 6.34), we can have **coframes** of the form

$$\{d\,x_1, \ldots, d\,x_m\}.$$

6.3.7 The differential df

Definition 6.44. Let $f : M \to \mathbb{R}$. Let $\mathbf{q} \in M$ and $X_{\mathbf{q}} \in T_{\mathbf{q}}M$. Then df, the **differential** of f, is defined (in a **coordinate-free** way) as

$$df|_{\mathbf{q}}(X_{\mathbf{q}}) = X_{\mathbf{q}}f$$

Note that df is a linear functional on each tangent space: it takes a tangent vector and produces a scalar, and it is linear.

Remark 6.45. In coordinates ([**Lee2**]), df is written

$$df = \sum_{i=1}^{m} \frac{\partial f}{\partial x_i}\, dx_i.$$

This looks a lot like the gradient, but written as a covector rather than a vector. Why? Well, the Jacobian matrix for a vector-to-scalar function is a $1 \times m$ matrix. This looks like a **row vector**, and we've been equating row vectors with linear functionals (see section 4.6.6). Recall from the discussion in section 4.6.7 that if we were to treat it as a **column vector** (plain old vector) rather than as a row vector (linear functional or covector) then it wouldn't transform correctly on change of basis.

Remark 6.46. In addition to asking what kind of map df is (we saw that it's a 1-form), we can also ask what kind of map d is. Here, we started with f, which is a scalar, or 0-form. Then df is a 1-form. One might guess that d will in general send a k-form to a $(k+1)$-form. That is indeed the case, as is discussed in section 6.4.

6.3.8 The tensor bundle

Given a manifold M, we now have a tangent bundle TM and a cotangent bundle T^*M. We also have the respective spaces of smooth sections, $\Gamma(TM)$ and $\Gamma(T^*M)$.

Definition 6.47. Elements of $\Gamma(TM)$ and $\Gamma(T^*M)$ are called **vector fields** and **covector fields**, respectively.

As discussed in section 6.2.2, $\Gamma(TM)$ and $\Gamma(T^*M)$ are free modules over $C^\infty(M,\mathbb{R})$. We can appropriate the discussion of tensors in section 4.7 wholesale, replacing constant coefficients with coefficients that vary smoothly over the manifold. This includes forming the full tensor algebra as well as defining symmetric and alternating tensors.

Notation 6.48. We write $\mathcal{T}^s_r(M)$ in place of $\mathcal{T}^s_r(TM)$.

Just as in section 4.7.6, $C^\infty(M,\mathbb{R})$ is a subspace of $\Gamma(TM)$ and so on. So we have:

$$\begin{tikzcd}
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
 & \mathcal{T}^2_0(M) \arrow[ul] \arrow[ur] & & \mathcal{T}^1_1(M) \arrow[ul] \arrow[ur] & & \mathcal{T}^0_2(M) \arrow[ul] \arrow[ur] \\
 & & \mathcal{T}^1_0(M) = TM \arrow[ul] \arrow[ur] & & \mathcal{T}^0_1(M) = T^*M \arrow[ul] \arrow[ur] & \\
 & & \text{tangent space} & & \text{cotangent space} & \\
 & & & \mathcal{T}^0_0(M) = C^\infty(M,\mathbb{R}) \arrow[ul] \arrow[ur] & & \\
 & & & \text{scalars} & &
\end{tikzcd}$$

as well as:

$$\begin{tikzcd}
 & & & & & \textbf{higher-order covariant} \\
\cdots & \cdots & \cdots & \cdots & \cdots & \textbf{tensor fields} \\
 & \Gamma(\mathcal{T}^2_0(M)) \arrow[ul] \arrow[ur] & & \Gamma(\mathcal{T}^1_1(M)) \arrow[ul] \arrow[ur] & & \Gamma(\mathcal{T}^0_2(M)) \arrow[ul] \arrow[ur] \\
 & \text{contravariant 2-tensor} & & \text{mixed 2-tensor} & & \textbf{covariant 2-tensor} \\
 & \text{fields} & & \text{fields} & & \textbf{fields} \\
 & & \Gamma(\mathcal{T}^1_0(M)) = \Gamma(TM) \arrow[ul] \arrow[ur] & & \Gamma(\mathcal{T}^0_1(M)) = \Gamma(T^*M) \arrow[ul] \arrow[ur] & \\
 & & \textbf{vector fields} & & \textbf{covector fields} & \\
 & & & \Gamma(\mathcal{T}^0_0(M)) = \Gamma(C^\infty(M,\mathbb{R})) \arrow[ul] \arrow[ur] & & \\
 & & & \textbf{scalar fields} & &
\end{tikzcd}$$

In the second diagram, I've boldfaced the parts of the full tensor bundle which are of interest for the geometry course: namely, covariant tensor fields (the right-hand linear branch of the tree), as well as contravariant 1-tensor fields, i.e. vector fields (the left-hand spot).

6.3.9 The metric tensor

Definition 6.49. As foreshadowed in remark 4.119, the **metric tensor** is nothing more than the dot product in tangent spaces. That is, the metric tensor $g \in \Gamma(\mathcal{T}_2^0(M))$ is defined by

$$g(\mathbf{v}_q, \mathbf{w}_q) = \mathbf{v} \cdot \mathbf{w}.$$

Explicit computations are much as in section 4.7.13, except that the basis elements $\mathbf{b}_1, \ldots, \mathbf{b}_m$ are replaced by $\partial/\partial x_1, \ldots, \partial/\partial x_m$, and likewise the dual basis elements $\mathbf{b}_1^*, \ldots, \mathbf{b}_m^*$ are replaced by $dx_1, \ldots, dx_m$.

Remark 6.50. In Euclidean coordinates,

$$\mathbf{v} \cdot \mathbf{w} = \sum_{i=1}^m v_i w_i.$$

Recall that dx_i is the ith coordinate-selector function, so we can write this in tensor lingo as

$$g(\mathbf{v}_q, \mathbf{w}_q) = \sum_{i=1}^m (dx_i \otimes dx_i)(\mathbf{v}, \mathbf{w}).$$

Treating g as a 2-tensor, we can write

$$g = \sum_{i=1}^m dx_i \otimes dx_i.$$

In particular, for Euclidean 3-space, we have

$$g = dx \otimes dx + dy \otimes dy + dz \otimes dz$$

which we sometimes abbreviate as

$$g = dx^2 + dy^2 + dz^2.$$

See section 10.3.1 for an example computation.

6.3.10 Forms

The human mind has first to construct forms, independently, before we can find them in things.
— Albert Einstein (1879-1955).

xxx to do.

somewhere (perhaps not here): **pullback of forms**. xxx incl. a diagram too?

$$\begin{aligned} f : z &\mapsto z^n \\ \omega &= dz \\ f^*(\omega) &= \omega \circ f \\ f^*(dz) &= d(z^n) = nz^{n-1}dz. \end{aligned}$$

incl/xref to i^* example too.

xxx diagram with coords/parameterizations and dp vs. $\partial/\partial p$ — which goes which way.

6.4 The exterior derivative

Mathematicians are like Frenchmen: whatever you say to them they translate into their own language and forthwith it is something entirely different. — Johann Wolfgang von Goethe (1749-1832).

6.4.1 Definition of d

And what are these fluxions? The velocities of evanescent increments. And what are these same evanescent increments? They are neither finite quantities, nor quantities infinitely small, nor yet nothing. May we not call them the ghosts of departed quantities...?
— Bishop Berkeley, in *The Analyst: A Discourse Addressed to an Infidel Mathematician* (1734).

Both [**Lee2**] and [**Spivak2**] define the exterior derivative in coordinates. Namely, if we have an r-form

$$\omega = f(x_1, \ldots, x_m) dx_{i_1} \wedge \ldots \wedge dx_{i_r}$$

then we write

$$d\omega = df(x_1, \ldots, x_m) dx_{i_1} \wedge \ldots \wedge dx_{i_r}.$$

where the df is taken on coefficient functions in the sense of section 6.3.7. The d operator takes r-forms to $(r+1)$-forms. It can be shown that the following properties hold:

- The d operator is unique, and is independent of change of coordinates.
- Given forms ω and η, $d(\omega + \eta) = d\omega + d\eta$.
- $d(s\omega) = sd\omega$ for $s \in \mathbb{R}$.
- $d(\omega \wedge \eta) = d\omega \wedge \eta + (-1)^{\mathrm{ord}(\omega)} \omega \wedge d\eta$. (This is the **signed Leibniz rule**.)
- (Note that the previous three items satisfy the definition of an **antiderivation** on the alternating tensor algebra.)
- $d^2 = 0$. [xxx xref to complex section.]

The following computational rules apply:

- We simply write out the differentials, then simplify.
- From corollary 4.111, if ω is an odd-degree form then $\omega \wedge \omega = 0$. This applies in particular to the 1-forms dx et al.: $dx \wedge dx = 0$.
- We can replace $dy \wedge dx$ with $-dx \wedge dy$.
- Note that $f\ dx$ is shorthand for $f \wedge dx$.
- We are working in the algebra of alternating tensors, with the wedge as multiplication. So, the usual rules of arithmetic apply, e.g. the distributive property. However, see immediately below for the commutativity rule.

Example 6.51. $\triangleright$

$$\begin{aligned}
\omega &= f(x,y,z)dy \wedge dz + g(x,y,z)dz \wedge dx + h(x,y,z)dx \wedge dy \\
d\omega &= df \wedge dy \wedge dz + dg \wedge dz \wedge dx + dh \wedge dx \wedge dy \\
&= (f_x dx + f_y dy + f_z dz) \wedge dy \wedge dz \\
&+ (g_x dx + g_y dy + g_z dz) \wedge dz \wedge dx \\
&+ (h_x dx + h_y dy + h_z dz) \wedge dx \wedge dy \\
&= f_x dx \wedge dy \wedge dz + g_y dy \wedge dz \wedge dx + h_z dz \wedge dx \wedge dy \\
&= (f_x + g_y + h_z) dx \wedge dy \wedge dz.
\end{aligned}$$

Furthermore, if

$$\mathbf{F} = (f, g, h)$$

then we can write

$$(f_x + g_y + h_z) dx \wedge dy \wedge dz = (\nabla \cdot \mathbf{F}) dV.$$

$\triangleleft$

[xxx: major lack — geometric view of the d operator. Without this it's hard to come up with a geometric notion of when two forms are cohomologous. Perhaps it suffices to work locally. Can I claim that then all k forms are linear combinations of $dx_1 \wedge \cdots \wedge dx_k$? Then does it reduce to obtaining a geometric view of df? If so:

$$df = \frac{\partial f}{\partial x}\, dx + \frac{\partial f}{\partial y}\, dy.$$

Use dx, dy as spines. This is just the gradient (covariantly of course) . How to extend to higher-level forms?

$$\begin{aligned}
\omega &= f(x,y)\, dx + g(x,y)\, dy \\
d\omega &= \left(\frac{\partial f}{\partial x}\, dx + \frac{\partial f}{\partial y}\, dy \right) \wedge dx \; + \; \left(\frac{\partial g}{\partial x}\, dx + \frac{\partial g}{\partial y}\, dy \right) \wedge dy \\
&= \left(\frac{\partial g}{\partial y} - \frac{\partial f}{\partial x} \right) dx \wedge dy.
\end{aligned}$$

First, plot a contour of f. Then, superimpose gradient arrows. Second, draw a quiver of arrows, with f and g components in a pair of adjacent dotted triangles. Then, what does $\partial g/\partial y - \partial f/\partial x$ look like? ...]

6.4.2 Forms in 3-space TBD

$$
\begin{tikzcd}
 & & \begin{array}{c}\mathbf{F} \lrcorner\, dV = \\ f\, dy \wedge dz + \\ g\, dz \wedge dx + \\ h\, dx \wedge dy\end{array} & \\
f & \begin{array}{c}\mathbf{F} \cdot d\mathbf{x} = \\ f dx + g dy + h dz\end{array} & & f dV \\
\Omega^0(\mathbb{R}^3) \arrow[r, "d"] & \Omega^1(\mathbb{R}^3) \arrow[r, "d"] & \Omega^2(\mathbb{R}^3) \arrow[r, "d"] & \Omega^3(\mathbb{R}^3) \\
C^\infty(\mathbb{R}^3, \mathbb{R}) \arrow[u, "\mathrm{Id}"] \arrow[r, "\mathrm{grad}=\nabla"] & \mathrm{vect}(\mathbb{R}^3) \arrow[u, "\mathbf{F} \cdot d\mathbf{x}"] \arrow[r, "\mathrm{curl} = \nabla\times"] & \mathrm{vect}(\mathbb{R}^3) \arrow[u, "\mathbf{F} \lrcorner\, dV"] \arrow[r, "\mathrm{div} = \nabla\cdot"] & C^\infty(\mathbb{R}^3, \mathbb{R}) \arrow[u, "f\, dV"] \\
f & \mathbf{F} = (f, g, h) & \mathbf{F} = (f, g, h) & f
\end{tikzcd}
$$

Include the cor. 12 stuff here. The property $d^2 = 0$ means:

- curl of grad is 0
- div of curl is 0

incl. $X \lrcorner$ in the diagram.

$\hat{\mathbf{n}} \lrcorner\, dV = dA$.

need to show $\mathbf{F} \lrcorner\, dV = (\mathbf{F} \cdot \hat{\mathbf{n}}) dS$. Lee lemma 13.25.

xxx here or somewhere else: closed, exact. dx is exact; $d\theta$ is not; dA is exact iff M is non-compact. Explain more .....

6.5 Computations with tangent vectors and forms

6.5.1 When $dx \neq dx$: charts, embeddings, projections, and variance

also graph coordinates on, say, upper sheet of hyperboloid of two sheets. different meanings for $\partial/\partial x$ etc.

xxx crucial distiction between *enough coordinates* (i.e. a chart) and *too many coordinates* (i.e. an embedding). Neither is "wrong"; both have their uses. We simply need to be careful about what the symbols mean.

xxx peril: illustrate by example. On $\mathbb{S}^2$, the function $f = x^2 + y^2 + z^2 - 1$ is 0 — obviously. Less obviously, $y\, dy \wedge dz - x\, dz \wedge dx$ is also zero. Show this using spherical coordinates.

* * *

In the plane, we convert *from* polar coordinates *to* rectangular coordinates as follows:

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} r\cos\theta \\ r\sin\theta \end{pmatrix}.$$

Then [xxx point out directions] the following two pop out easily:

$$\frac{\partial}{\partial r} = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix} = \begin{pmatrix} x/r \\ y/r \end{pmatrix} = \begin{pmatrix} \frac{x}{\sqrt{x^2+y^2}} \\ \frac{y}{\sqrt{x^2+y^2}} \end{pmatrix} \tag{6.1}$$

$$\frac{\partial}{\partial \theta} = \begin{pmatrix} -r\sin\theta \\ r\cos\theta \end{pmatrix} = \begin{pmatrix} -y \\ x \end{pmatrix} \tag{6.2}$$

[xxx footing caveats] and

$$dx = \frac{dx}{dr}\,dr + \frac{dx}{d\theta}\,d\theta = \cos\theta\,dr - r\sin\theta\,d\theta \tag{6.3}$$

$$dy = \frac{dy}{dr}\,dr + \frac{dy}{d\theta}\,d\theta = \sin\theta\,dr + r\cos\theta\,d\theta. \tag{6.4}$$

For the other four, it takes a bit more manipulation. [xxx disclaim about the arctan nonsense.] First, dr and $d\theta$.

$$\begin{aligned}
\begin{pmatrix} dx \\ dy \end{pmatrix} &= \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix} \begin{pmatrix} dr \\ d\theta \end{pmatrix} = \begin{pmatrix} \frac{x}{\sqrt{x^2+y^2}} & -y \\ \frac{y}{\sqrt{x^2+y^2}} & x \end{pmatrix} \begin{pmatrix} dr \\ d\theta \end{pmatrix} \\
\begin{pmatrix} dr \\ d\theta \end{pmatrix} &= \begin{pmatrix} \frac{x}{\sqrt{x^2+y^2}} & -y \\ \frac{y}{\sqrt{x^2+y^2}} & x \end{pmatrix}^{-1} \begin{pmatrix} dx \\ dy \end{pmatrix} \\
&= \begin{pmatrix} x & y \\ -\frac{y}{\sqrt{x^2+y^2}} & \frac{x}{\sqrt{x^2+y^2}} \end{pmatrix} \left(\frac{1}{\sqrt{x^2+y^2}}\right) \begin{pmatrix} dx \\ dy \end{pmatrix} \\
&= \begin{pmatrix} \frac{x}{\sqrt{x^2+y^2}} & \frac{y}{\sqrt{x^2+y^2}} \\ -\frac{y}{x^2+y^2} & \frac{x}{x^2+y^2} \end{pmatrix} \begin{pmatrix} dx \\ dy \end{pmatrix}.
\end{aligned}$$

Second, $\partial/\partial x$ and $\partial/\partial y$:

$$\begin{aligned}
\begin{pmatrix} \partial/\partial r \\ \partial/\partial \theta \end{pmatrix} &= \begin{pmatrix} \frac{x}{\sqrt{x^2+y^2}} & \frac{y}{\sqrt{x^2+y^2}} \\ -y & x \end{pmatrix} \begin{pmatrix} \partial/\partial x \\ \partial/\partial y \end{pmatrix} \\
\begin{pmatrix} \partial/\partial x \\ \partial/\partial y \end{pmatrix} &= \begin{pmatrix} x & \frac{-y}{\sqrt{x^2+y^2}} \\ y & \frac{x}{\sqrt{x^2+y^2}} \end{pmatrix} \left(\frac{1}{\sqrt{x^2+y^2}}\right) \begin{pmatrix} \partial/\partial r \\ \partial/\partial \theta \end{pmatrix} \\
&= \begin{pmatrix} \frac{x}{\sqrt{x^2+y^2}} & \frac{-y}{x^2+y^2} \\ \frac{y}{\sqrt{x^2+y^2}} & \frac{x}{x^2+y^2} \end{pmatrix} \begin{pmatrix} \partial/\partial r \\ \partial/\partial \theta \end{pmatrix} = \begin{pmatrix} \frac{r\cos\theta}{r} & \frac{-r\sin\theta}{r^2} \\ \frac{r\sin\theta}{r} & \frac{r\cos\theta}{r^2} \end{pmatrix} \begin{pmatrix} \partial/\partial r \\ \partial/\partial \theta \end{pmatrix} \\
&= \begin{pmatrix} \cos\theta & \frac{-\sin\theta}{r} \\ \sin\theta & \frac{\cos\theta}{r} \end{pmatrix} \begin{pmatrix} \partial/\partial r \\ \partial/\partial \theta \end{pmatrix}.
\end{aligned}$$

So, as a final result,

$$\begin{aligned}
\frac{\partial}{\partial x} &= \cos\theta\,\frac{\partial}{\partial r} - \frac{\sin\theta}{r}\,\frac{\partial}{\partial \theta} & dx &= \cos\theta\,dr - r\sin\theta\,d\theta \\
\frac{\partial}{\partial y} &= \sin\theta\,\frac{\partial}{\partial r} + \frac{\cos\theta}{r}\,\frac{\partial}{\partial \theta} & dy &= \sin\theta\,dr + r\cos\theta\,d\theta
\end{aligned}$$

and

$$dr = \frac{x\,dx \;+\; y\,dy}{\sqrt{x^2+y^2}} \qquad\qquad \frac{\partial}{\partial r} = \frac{x}{\sqrt{x^2+y^2}}\,\frac{\partial}{\partial x} \;+\; \frac{y}{\sqrt{x^2+y^2}}\,\frac{\partial}{\partial y}$$

$$d\theta = \frac{-y\,dx \;+\; x\,dy}{x^2+y^2} \qquad\qquad \frac{\partial}{\partial \theta} = -y\,\frac{\partial}{\partial x} \;+\; x\,\frac{\partial}{\partial y}.$$

[xxx compute norms and show, and plot: length of $\partial/\partial r$ is 1, length of $\partial/\partial\theta$ is r; spine length of dr is 1, spine length of $d\theta$ is $1/r$. Thus it's no surprise that $d\theta$ is undefined at zero.]

* * *

When we restrict to $\mathbb{S}^1$, i.e. set $x^2+y^2=1$ and project onto the tangents, then we have

$$\frac{\partial}{\partial x} = -\sin\theta\,\frac{\partial}{\partial\theta} \qquad\qquad dx = -\sin\theta\,d\theta$$

$$\frac{\partial}{\partial y} = \cos\theta\,\frac{\partial}{\partial\theta} \qquad\qquad dy = \cos\theta\,d\theta$$

and

$$d\theta = -y\,dx \;+\; x\,dy \qquad\qquad \frac{\partial}{\partial\theta} = -y\,\frac{\partial}{\partial x} \;+\; x\,\frac{\partial}{\partial y}$$

[xxx then, use this to drive home the point about when $dx \neq dx$. Spine plots will make this quite clear.]

* * *

Note that [xxx $\mathbf{e}_1$ with $\partial/\partial x$ and $\mathbf{e}_2$ with $\partial/\partial y$] equations 6.1 and 6.2 are the same as

$$\frac{\partial}{\partial r} = \frac{\partial x}{\partial r}\,\frac{\partial}{\partial x} + \frac{\partial y}{\partial r}\,\frac{\partial}{\partial y}$$

$$\frac{\partial}{\partial \theta} = \frac{\partial x}{\partial \theta}\,\frac{\partial}{\partial x} + \frac{\partial y}{\partial \theta}\,\frac{\partial}{\partial y}.$$

Compare to equations 6.3 and 6.4:

$$dx = \frac{dx}{dr}\,dr + \frac{dx}{d\theta}\,d\theta \tag{6.5}$$

$$dy = \frac{dy}{dr}\,dr + \frac{dy}{d\theta}\,d\theta. \tag{6.6}$$

Graphically, we see

$$\begin{pmatrix} r \\ \theta \end{pmatrix} \xrightarrow{\quad F \quad} \begin{pmatrix} x \\ y \end{pmatrix} = F(r,\theta) = \begin{pmatrix} r\cos\theta \\ r\sin\theta \end{pmatrix}$$

$$F^*(dx) = d(r\cos\theta) = \cos\theta\, dr - r\sin\theta\, d\theta \xleftarrow{\quad F^* \quad} dx$$

$$F^*(dy) = d(r\sin\theta) = \sin\theta\, dr + r\cos\theta\, d\theta \xleftarrow{\quad F^* \quad} dy$$

$$\tfrac{\partial}{\partial r} \xmapsto{\quad F_* \quad} F_*\left(\tfrac{\partial}{\partial r}\right) = \tfrac{\partial x}{\partial r}\tfrac{\partial}{\partial x} + \tfrac{\partial y}{\partial r}\tfrac{\partial}{\partial y} = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}$$

$$\tfrac{\partial}{\partial \theta} \xmapsto{\quad F_* \quad} F_*\left(\tfrac{\partial}{\partial \theta}\right) = \tfrac{\partial x}{\partial \theta}\tfrac{\partial}{\partial x} + \tfrac{\partial y}{\partial \theta}\tfrac{\partial}{\partial y} = \begin{pmatrix} -r\sin\theta \\ r\cos\theta \end{pmatrix}.$$

Two points:

- This echoes the discussion in section 4.6.7, wherein I noted that column vectors (e.g. $\partial/\partial x$), which are contravariant, and row vectors (e.g. dx), which are covariant, change coordinates differently.
- This justifies our use of $\partial/\partial x$ for vectors and dx for functionals: namely, the ∂'s and d's cancel correctly in the numerators and denominators in the above formulas, in accordance with their respective contravariance and covariance.

6.5.2 Stereographic coordinates on $\mathbb{S}^1$

xxx dp, dq, etc. In terms of (known) $d\theta$.

6.5.3 A gallery of curves and surfaces, part 2

xref back to gallery1, and have it xref here. what coordinates to use. computing partials. Make sure to do upper sheet of two-sheeted hyperboloid; compare and contrast $\partial/\partial x$ with $\partial/\partial x$.

6.5.4 asdfasdfasdf

Formalize the following mnemonic: with $i : M \to \mathbb{R}^m$: $M = \ker F$, $TM = \ker DF$.

xxx xref to the SL problem from the final review.

xxx Include here some references to vector calculus. Common surfaces such as the sphere are a level set of some $F : \mathbb{R}^3 \to \mathbb{R}$. The gradient points in the direction of greatest change; it is the field of vectors normal to the surface. Perpendicular to that, F is (to linear approximation) constant. Now, $\mathbf{v} \in \ker DF$ means

$DF \cdot \mathbf{v} = 0$ which simply means that the dot of $\mathbf{v}$ with DF is zero, i.e. $\mathbf{v}$ is perpendicular to the normal. This is exactly what we would expect.

6.6 Cohomologies

xxx cellular, simplicial, singular, de Rham w/ xref back to homologies sectoin. xxx note prefer singular for the same reason as there.

xxx put the de Rham thm right up front. This is a representation theorem, equating singular cohomology with de Rham cohomology. So, it suffices to work with the latter. Note that qual questions only ever ask about de Rham cohomology.

xxx def or xref to closed, exact.

xxx in VS section, have already done sequences of mappings including exact and short exact ones. Then, note d as a sequence of mappings. Then, closed mod exact as ker mod im.

6.6.1 de Rham cohomology

Definition 6.52. de Rham cohomology ...

Theorem 6.53 (Poincaré lemma)**.** *A contractible (technically,* star-shaped*) open subset U of $\mathbb{R}^n$ has trivial cohomology $H^k_{dR}(U)$ for $k > 0$.*

Remark 6.54. This means that *if* a form is closed on a contractible open set, then it is exact. If a form is not closed, then the Poincaré lemma does not apply.

long exact seqs, complexes, homology and cohomology. where to put this in the right order vs. d and ∂.

6.7 Section title goes here

This section is not qualifier material.

6.7.1 Old-fashioned tensors

From [**PDM**]:

- Let $\mathbf{x}$ and $\mathbf{y}$ be coordinate functions (xxx for what). Thus, a point q (of what) has coordinates $(x^1(q), \ldots, x^m(q))$ in the first coordinate system, and coordinates $(y^1(q), \ldots, y^m(q))$ in the second coordinate system.
- A set of m components, written A^i, that are functions of the first set of coordinates will become a set of m components, denoted B^i, after transformation to the second set of coordinates.
- A set of m^2 components, written A^{ij}, that are functions of the first set of coordinates will become a set of m^2 components, denoted B^{ij}, after transformation to the second set of coordinates.
- A set of m^3 components, written A^{ijk}, that are functions of the first set of coordinates will become a set of m^3 components, denoted B^{ijk}, after transformation to the second set of coordinates.
- A tensor is a set of components that obeys some transformation law. The number of suffixes indicates the **rank** of the tensor; their position indicates the type (covariant, contravariant, or mixed) of the tensor.
- A covariant 1-tensor is a set (?) A_i satisfying, for each i,
$$B_i = \sum_{r=1}^{n} \frac{\partial x^r}{\partial y^i} A_r.$$
- A contravariant 1-tensor is a set (?) A^i satisfying, for each i,
$$B^i = \sum_{r=1}^{n} \frac{\partial y^i}{\partial x^r} A^r.$$
- A covariant 2-tensor is a set (?) A_{ij} satisfying, for each i and j,
$$B_{ij} = \sum_{r=1}^{n} \sum_{s=1}^{n} \frac{\partial x^r}{\partial y^i} \frac{\partial x^s}{\partial y^j} A_{rs}.$$
- A contravariant 2-tensor is a set (?) A^{ij} satisfying, for each i and j,
$$B^{ij} = \sum_{r=1}^{n} \sum_{s=1}^{n} \frac{\partial y^i}{\partial x^r} \frac{\partial y^j}{\partial x^s} A^{rs}.$$
- A mixed 2-tensor is a set (?) A^i_j satisfying, for each i and j,
$$B^i_j = \sum_{r=1}^{n} \sum_{s=1}^{n} \frac{\partial y^i}{\partial x^r} \frac{\partial x^s}{\partial y^j} A^r_s.$$

xxx several examples here. First, a simple rotation in $\mathbb{R}^2$. Show how linear functional, dot, and determinant transform.

Try to express this in terms of concepts used in the course.

Is this just using the COB matrix?? Try a couple specific examples.

$$
\begin{aligned}
\lambda(\mathbf{u}, \mathbf{v}, \mathbf{w}) &= \lambda(u_1\mathbf{e}_1 + u_2\mathbf{e}_2 + u_3\mathbf{e}_3, v_1\mathbf{e}_1 + v_2\mathbf{e}_2 + v_3\mathbf{e}_3, w_1\mathbf{e}_1 + w_2\mathbf{e}_2 + w_3\mathbf{e}_3) \\
&= \sum_i \sum_j \sum_k u_i v_j w_k \lambda(\mathbf{e}_i, \mathbf{e}_j, \mathbf{e}_k) \\
&= \sum_i \sum_j \sum_k u_i v_j w_k L_{ijk}
\end{aligned}
$$

xxx xref back to section 4.6.7: here we are generalizing that concept, which was done there for vectors and functionals (i.e. tensors of type 1,0 and type 0,1).

7 Duality

The union of the mathematician with the poet, fervor with measure, passion with correctness, this surely is the ideal. — William James (1842-1910), *Collected Essays.*

xxx quack about chains and forms; homology and cohomology.

7.1 Integration

xxx xref to chapter 10 of [**Rudin**].

Definition 7.1. Integral over a k-chain ....

Proposition 7.2. *Well-definedness with respect to choice of pullback ....*

7.2 Stokes' theorem

Here we discuss a generalized Stokes' theorem. This modern formulation includes several classical integral theorems as special cases.

7.2.1 Stokes theorem for chains

steps (xxx make a picture):

- c is a single k-cube and w is a simple $(k-1)$-form. Moreover c is the inclusion of $[0,1]^k$ into $\mathbb{R}^k$. Write out using definitions of ∂c, dw, Fubini, and the FTC on one variable.
- Show independence of parameterization on $[0,1]^k$ (source).
- Let c be an arbitrary k-cube (destination).
- Sums of c_i's and ω_j's.

7.2.2 Stokes theorem for manifolds

xxx need to define:

Definition 7.3. manifold with boundary

Definition 7.4. boundary of a manifold.

This is written ∂M. If M has dimension m,then ∂M has dimension $m-1$.

Definition 7.5. integral.

Definition 7.6. induced orientation.

Let M be a compact, oriented manifold. Suppose $\omega \in \Omega^{m-1}(M)$. Then

$$\int_M d\omega = \int_{\partial M} \omega.$$

xxx express in terms of the pairing of forms and chains: ∂ is the adjoint of d:

$$\langle d\omega, c\rangle = \langle \omega, \partial c\rangle.$$

7.2.3 Proof of Stokes

[xxx first use POUs to say why it suffices to work within a single chart.]

[xxx next use chart invariance (cancelling dets) to say why it now remains to work in $\mathbb{R}^n$.]

Sketch (needs a picture w/ transition function):

$$\int_U (y^{-1})^*\omega = \int_U (yx^{-1})^*(y^{-1})^*\omega = \int_U (x^{-1})^* y^*(y^{-1})^*\omega = \int_U (x^{-1})^*\omega.$$

Let f be a 0-form; let c be a 1-chain. That is, c is a smooth map from $[0,1]$ to X.

$$\begin{aligned}
\int_c df &= \int_0^1 c^*(df) \quad \text{(pull } df \text{ back to } [0,1]) \\
&= \int_0^1 dc^*(f) \quad (d \text{ commutes with pullback}) \\
&= (c^*f)(1) - (c^*f)(0) \quad \text{(fundamental theorem of calculus)} \\
&= f(c(1)) - f(c(0)) \quad \text{(definition of pullback)}
\end{aligned}$$

Meanwhile,

$$\begin{aligned}
\partial c &= 1 \cdot c(1) - 1 \cdot c(0) \\
\int_{\partial c} f &= f(c(1)) - f(c(0)) \\
&= \int_c df.
\end{aligned}$$

See section 10.6.12 for the two-dimensional case. [xxx merge that in here, or eliminate.]

Next, we will look at how the classical integration theorems of section 2.6 are special cases of Stokes' theorem. We will pay particular attention to how the classical notation translates into the generalized notation.

7.2.4 The fundamental theorem of calculus

Let M be the closed interval $[a,b]$ for some $a,b \in \mathbb{R}$. Then $\partial M = \{a,b\}$. Let $f(x)$ be a a differentiable function on $[a,b]$, i.e. $f \in \Omega^0(M)$. For an antiderivative $F(x)$ of f, i.e. $dF/dx = f$, we are accustomed to thinking of the fundamental theorem of calculus as

$$\int_a^b f(x)dx = F(b) - F(a).$$

Using our current terminology, we can observe that the integrand fdx is a 1-form. This is d of something, namely, $fdx = dF$. Then we can write the left-hand side of the Stokes equation as

$$\int_a^b fdx = \int_a^b \frac{dF}{dx}dx = \int_a^b dF,$$

and we can write the right-hand side as

$$\int_{\partial[a,b]} F = F(b) - F(a).$$

7.2.5 Line integrals independent of path

xxx xref to the figure in section 6.4.

Let M be a curve (i.e. a manifold of dimension 1) of $\mathbb{R}^2$ between points $\mathbf{p}$ and $\mathbf{q}$. Then $\partial M = \{\mathbf{p}, \mathbf{q}\}$. Given $f, g : \mathbb{R}^2 \to \mathbb{R}$, we write the line integral

$$\int_M f(x,y)dx + g(x,y)dy.$$

The integrand is a 1-form, i.e.

$$f(x,y)dx + g(x,y)dy \in \Omega^1(M).$$

To use generalized Stokes, we need to write the integrand as d of something. If f and g are smooth (which we assume in this course) then in fact ([**Anton**], section xxx) there is a vector-to-scalar function $H : \mathbb{R}^2 \to \mathbb{R}$ such that

$$\frac{\partial H}{\partial x} = f, \frac{\partial H}{\partial y} = g.$$

Then

$$f dx + g dy = \frac{\partial H}{\partial x} dx + \frac{\partial H}{\partial y} dy = \nabla H \cdot d\mathbf{x}$$

where $d\mathbf{x}$ is the vector (dx, dy). Note that we can also write

$$dH = \frac{\partial H}{\partial x} dx + \frac{\partial H}{\partial y} dy$$

so

$$\int_M f dx + g dy = \int_M dH.$$

Now we can use Stokes' theorem to say

$$\int_M dH = \int_{\partial M} H = H(\mathbf{q}) - H(\mathbf{p}).$$

Hence the classical result that

$$\int_M f dx + g dy = H(\mathbf{q}) - H(\mathbf{p}).$$

Note that the right-hand side is independent of the curve M, as we would expect from xxx cite.

7.2.6 Green's theorem

Let M be a connected blob, i.e. a bounded manifold of dimension 2, in $\mathbb{R}^2$. (One might think of M as a pancake, with $\mathbb{R}^2$ as the griddle.) Let C be a closed curve around the perimeter of M. Use the positive induced orientation (xxx put in a picture here) with tangent vector $\hat{\mathbf{n}}$. Let

$$\omega = f(x,y)dx + g(x,y)dy \in \Omega^1(M).$$

Then we want to compute

$$\oint_C f(x,y)dx + g(x,y)dy.$$

Note that we have $C = \partial M$. In the previous examples, we had to obtain the integrand as d of something. Here, we are given the other side of a Stokes equation: we are given ω and all we need to do is to compute $d\omega$. We have (xxx xref to above section)

$$\begin{aligned}
d\omega &= d(f dx + g dy) = d(f dx) + d(g dy) \\
&= \left(\frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy\right) \wedge dx + \left(\frac{\partial g}{\partial x} dx + \frac{\partial g}{\partial y} dy\right) \wedge dy \\
&= \frac{\partial f}{\partial x} dx \wedge dx + \frac{\partial f}{\partial y} dy \wedge dx + \frac{\partial g}{\partial x} dx \wedge dy + \frac{\partial g}{\partial y} dy \wedge dy \\
&= \frac{\partial f}{\partial y} dy \wedge dx + \frac{\partial g}{\partial x} dx \wedge dy \\
&= \left(\frac{\partial g}{\partial x} - \frac{\partial f}{\partial y}\right) dx \wedge dy.
\end{aligned}$$

Thus

$$\int_M d\omega = \int_M \left(\frac{\partial g}{\partial x} - \frac{\partial f}{\partial y}\right) dx \wedge dy = \int_M \left(\frac{\partial g}{\partial x} - \frac{\partial f}{\partial y}\right) dA$$

and we obtain the classical result

$$\int_M \left(\frac{\partial g}{\partial x} - \frac{\partial f}{\partial y}\right) dA = \oint_C f dx + g dy.$$

xxx need to write this in terms of the normal vector as well.

7.2.7 Classical Stokes

Let M be a 2-dimensional submanifold of $\mathbb{R}^3$. That is, M is a surface, where the canonical example is a potato chip. Let $\hat{\mathbf{n}}$ be an outward-pointing normal vector. Start with a 1-form

$$\omega = \mathbf{B} \cdot d\mathbf{x} = B_1(x, y, z)dx + B_2(x, y, z)dy + B_3(x, y, z)dz.$$

xxx incomplete.

xxx think of: the work done by a magnetic field around a wire equals flux of field through interpolating surface?

7.2.8 Divergence theorem

[xxx elaborate on this sketch]

$$\int_M (\nabla \cdot \mathbf{F}) dV = \int_M d(\mathbf{F} \lrcorner dV) = \int_{\partial M} \mathbf{F} \lrcorner dV = \int_{\partial M} \mathbf{F} \cdot \hat{\mathbf{n}}\, dA.$$

xxx more:

$$\begin{aligned}
L_{\mathbf{F}}(dV) &= d(\mathbf{F} \lrcorner dV) + \mathbf{F} \lrcorner (d(dV)) \\
L_{\mathbf{F}}(dx \wedge dy \wedge dz) &= (L_{\mathbf{F}}(dx)) \wedge dy \wedge dz + dx \wedge (L_{\mathbf{F}}(dx)) \wedge dz + dx \wedge dy \wedge (L_{\mathbf{F}}(dx)) \\
&= d(L_{\mathbf{F}}(x)) \wedge dy \wedge dz + dx \wedge d(L_{\mathbf{F}}(x)) \wedge dz + dx \wedge dy \wedge d(L_{\mathbf{F}}(x)) \\
L_{\mathbf{F}}(x) &= (F_1 \partial/\partial x + F_2 \partial/\partial y + F_3 \partial/\partial z)x = F_1 \\
& \quad \text{etc.} \\
L_{\mathbf{F}}(dx \wedge dy \wedge dz) &= dF_1 \wedge dy \wedge dz + dx \wedge dF_2 \wedge dz + dx \wedge dy \wedge dF_3 \\
&= (\nabla \cdot \mathbf{F}) dV.
\end{aligned}$$

7.2.9 Cauchy's theorem

7.3 Duality theorems

Pairing (see section 4.6.9) of homology and cohomology (cycles and forms). This:

$$\Omega^k(M) \times C_k(M) \to R$$

via

$$(\omega, c) \mapsto \int_c \omega$$

Induces a pairing of homology and cohomology:

$$H^k_{dR}(M) \times H_k(M) \to R$$

via

$$([\omega], [c]) \mapsto \int_c \omega$$

which is well-defined. Type up 4-13: pull back to I^m and Riemann integrals.

Proposition 7.7 (Poincaré duality)**.** *If M is an orientable smooth manifold, then*

$$H^{m-k}_{dR}(M) \cong H^k_{dR}(M)^*$$

via

$$[\omega^{(m-k)}] \quad \longmapsto \quad \left([\eta^{(k)}] \mapsto \int_M \omega \wedge \eta\right).$$

Remark 7.8. We can remember that this theorem has orientability as a hypothesis — else the integral of a top form over all of M would be zero.

Remark 7.9. Also (via what map? "cap product" (?!?) with "fundamental class" $[M]$):

$$H^{m-k}(M) \cong H_k(M).$$

Proposition 7.10 (de Rham Theorem)**.** *Let M be a smooth manifold. Then*

$$H^k_{dR}(M) \cong H_k(M)^*$$

via

$$[\omega^{(k)}] \quad \longmapsto \quad \left([c^{(k)}] \mapsto \int_c \omega\right).$$

Remark 7.11. Note that orientability is not needed here — the integrals are over chains, which are oriented.

Proof. See [**Lee2**] for the full proof, which is non-trivial. Here, though, I will note why it is that the above map is well-defined on cohomology (the ω's) and homology (the c's).

First suppose $[\omega] = [\eta]$. Recall that cohomology classes are closed forms mod exact forms, and homology classes are cycles mod boundaries. This means $\omega - \eta$ is exact: say $\omega - \eta = d\alpha$. Then

$$\int_c (\omega - \eta) = \int_c d\alpha = \int_{\partial c} \alpha = 0$$

by Stokes theorem, and because c is a cycle so $\partial c = 0$. Thus

$$\int_c \omega = \int_c \eta.$$

Second suppose $[c] = [b]$. Then $c - b$ is the boundary of a chain: say $c - b = \partial a$. Then

$$\int_{c-b} \omega = \int_{\partial a} \omega = \int_a d\omega = 0$$

by Stokes theorem, and because ω is closed so $d\omega = 0$. Thus

$$\int_c \omega = \int_b \omega.$$

□

Remark 7.12. One direction is obvious: given any k form ω we can certainly turn it into a functional as shown above. The payload is the other direction. This is really a **representation theorem**, much like proposition 4.64. Namely, *all* functionals on (real) homology come from forms in this way.

xxx *nature* of the explicit maps for M-V on cohomology: $j_U^* \oplus j_V^*$ and $i_U^* - i_V^*$, with $j_U^*(\omega)$ being just the *restriction* of ω) to U, etc. Make a nice picture. What about the connecting map δ?

7.4 Notation TBD

By relieving the brain of all unnecessary work, a good notation sets it free to concentrate on more advanced problems, and, in effect, increases the mental power of the race. — Alfred North Whitehead (1861-1947).

[xxx needs elaboration]

- $C_k(M)$: k-chains.
- $Z_k(M)$: k-cycles ($\partial c = 0$).
- $B_k(M)$: k-boundaries ($c = \partial b$).
- $H_k(M)$: $Z_k(M)/B_k(M)$ (kth homology is k-cycles mod k-boundaries).

- $\Omega^k(M)$: k-forms.
- $Z^k(M)$: closed k-forms ($d\omega = 0$).
- $B^k(M)$: exact k-forms ($\omega = d\eta$).
- $H^k(M)$: $Z^k(M)/B^k(M)$. (kth cohomology is closed k forms mod exact k-forms).

and quotients (since $\partial^2 = 0$ and $d^2 = 0$):

$$\begin{array}{ll} C_k(M) \supseteq Z_k(M) \supseteq B_k(M); & H_k(M) = Z_k(M)/B_k(M) \\ \Omega^k(M) \supseteq Z^k(M) \supseteq B^k(M); & H^k(M) = Z^k(M)/B^k(M) \end{array}$$

Here are the inclusions and their induced maps on chains and homology (used, with varying connectivity hypotheses, in Seifert-van Kampen and Mayer-Vietoris). Note that the subscript # and $*$ are covariant functors from manifolds to chain complexes and homology groups: they preserve the directions of the arrows (see notation 4.18).

$$\begin{tikzcd} & M & \\ U \arrow[ur, "j_U"] & & V \arrow[ul, "j_V"'] \\ & U \cap V \arrow[ul, "i_U"] \arrow[ur, "i_V"'] & \end{tikzcd} \qquad \begin{tikzcd} & C_k(M) & \\ C_k(U) \arrow[ur, "j_{U\#}"] & & C_k(V) \arrow[ul, "j_{V\#}"'] \\ & C_k(U \cap V) \arrow[ul, "i_{U\#}"] \arrow[ur, "i_{V\#}"'] & \end{tikzcd} \qquad \begin{tikzcd} & H_k(M) & \\ H_k(U) \arrow[ur, "j_{U*}"] & & H_k(V) \arrow[ul, "j_{V*}"'] \\ & H_k(U \cap V) \arrow[ul, "i_{U*}"] \arrow[ur, "i_{V*}"'] & \end{tikzcd}$$

Here are the inclusions and their induced maps on forms and cohomology. Note that the superscript # and $*$ are contravariant functors from manifolds to cochain complexes and cohomology groups: they reverse the directions of the arrows (see notation 4.18).

$$\begin{tikzcd} & M & \\ U \arrow[ur, "j_U"] & & V \arrow[ul, "j_V"'] \\ & U \cap V \arrow[ul, "i_U"] \arrow[ur, "i_V"'] & \end{tikzcd} \qquad \begin{tikzcd} & \Omega^k(M) \arrow[dl, "j_U^\#"'] \arrow[dr, "j_V^\#"] & \\ \Omega^k(U) \arrow[dr, "i_U^\#"'] & & \Omega^k(V) \arrow[dl, "i_V^\#"] \\ & \Omega^k(U \cap V) & \end{tikzcd} \qquad \begin{tikzcd} & H^k(M) \arrow[dl, "j_U^*"'] \arrow[dr, "j_V^*"] & \\ H^k(U) \arrow[dr, "i_U^*"'] & & H^k(V) \arrow[dl, "i_V^*"] \\ & H^k(U \cap V) & \end{tikzcd}$$

Behavior of induced maps:

- Induced maps on chains are *inclusions*. E.g. a chain c in $U \cap V$ is a chain in U by inclusion: $i_{U\#}(c) = c$.
- Induced maps on forms are *restrictions*. E.g. a form ω on M is a form on U by restriction: $j_U^{\#}(\omega) = \omega|_U$.

Short exact sequences of chain maps:

$$\begin{tikzcd}
 & \vdots & \vdots & \vdots & \\
0 \arrow[r] & C_k(U \cap V) \arrow[r, "i_{U\#} \oplus i_{V\#}"] \arrow[d, "\partial"] & C_k(U) \oplus C_k(V) \arrow[r, "j_{U\#} - j_{V\#}"] \arrow[d, "\partial"] & C_k(M) \arrow[r] \arrow[d, "\partial"] & 0 \\
0 \arrow[r] & C_{k-1}(U \cap V) \arrow[r, "i_{U\#} \oplus i_{V\#}"] & C_{k-1}(U) \oplus C_{k-1}(V) \arrow[r, "j_{U\#} - j_{V\#}"] & C_{k-1}(M) \arrow[r] & 0 \\
 & \vdots & \vdots & \vdots &
\end{tikzcd}$$

Long exact sequence of homology:

$$\begin{tikzcd}
\vdots & \vdots & \vdots & \\
H_k(U \cap V) \arrow[r, "i_{U*} \oplus i_{V*}"] & H_k(U) \oplus H_k(V) \arrow[r, "j_{U*} - j_{V*}"] & H_k(M) \arrow[dll, rounded corners, "\partial_*"] & \\
H_{k-1}(U \cap V) \arrow[r, "i_{U*} \oplus i_{V*}"] & H_{k-1}(U) \oplus H_{k-1}(V) \arrow[r, "j_{U*} - j_{V*}"] & H_{k-1}(M) & \\
\vdots & \vdots & \vdots &
\end{tikzcd}$$

Short exact sequences of form maps:

$$\begin{tikzcd}
 & \vdots & \vdots & \vdots & \\
0 & \Omega^k(U \cap V) \arrow[l] & \Omega^k(U) \oplus \Omega^k(V) \arrow[l, "i_U^{\#} - i_V^{\#}"'] & \Omega^k(M) \arrow[l, "j_U^{\#} \oplus j_V^{\#}"'] & 0 \arrow[l] \\
0 & \Omega^{k-1}(U \cap V) \arrow[l] \arrow[u, "d"] & \Omega^{k-1}(U) \oplus \Omega^{k-1}(V) \arrow[l, "i_U^{\#} - i_V^{\#}"'] \arrow[u, "d"] & \Omega^{k-1}(M) \arrow[l, "j_U^{\#} \oplus j_V^{\#}"'] \arrow[u, "d"] & 0 \arrow[l] \\
 & \vdots & \vdots & \vdots &
\end{tikzcd}$$

Long exact sequence of cohomology:

$$\begin{tikzcd}
\vdots & \vdots & \vdots & \\
H^k(U \cap V) & H^k(U) \oplus H^k(V) \arrow[l, "i_U^* - i_V^*"'] & H^k(M) \arrow[l, "j_U^* \oplus j_V^*"'] & \\
H^{k-1}(U \cap V) \arrow[u, rounded corners, "\delta"] & H^{k-1}(U) \oplus H^{k-1}(V) \arrow[l, "i_U^* - i_V^*"'] & H^{k-1}(M) \arrow[l, "j_U^* \oplus j_V^*"'] & \\
\vdots & \vdots & \vdots &
\end{tikzcd}$$

8 Flows

Flow with whatever may happen and let your mind be free. Stay centered by accepting whatever you are doing. This is the ultimate. — Chuang Tzu (c. 370-301 B.C.).

xxx define in terms of diffeomorphisms

xxx define **flow**, **complete flow**, **one-parameter family**.

xxx make sure to xref back to the ODE section.

1-parameter group of diffeomorphisms:

$$\phi_s \circ \phi_t = \phi_{s+t}.$$

To convert a diffeomorphism into a vector field, we have the following formula:

$$\mathbf{v}\Big|_{\left(\begin{smallmatrix} x \\ y \\ z \end{smallmatrix}\right)} = \frac{\partial \phi_t}{\partial t}\Big|_{t=0} \left(\begin{pmatrix} x \\ y \\ z \end{pmatrix} \right).$$

[xxx How to convert a vector field into a diffeomorphism? Solve ODE? Include an example or two. xref to Lee2 thms about existence.]

If

$$X = f(x,y,z)\,\frac{\partial}{\partial x} + g(x,y,z)\,\frac{\partial}{\partial y} + h(x,y,z)\,\frac{\partial}{\partial z}$$

then solve

$$\begin{aligned} \dot{x} &= f(x,y,z) \\ \dot{y} &= g(x,y,z) \\ \dot{z} &= h(x,y,z). \end{aligned}$$

Example 8.1. ▷ (Fall 2001 qualifying exam, #4.) For each $t \in \mathbb{R}$, let ϕ_t denote the map of $\mathbb{S}^2$ into itself which is defined by

$$\phi_t : \mathbb{S}^2 \to \mathbb{S}^2 : \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} \cos(t)x - \sin(t)y \\ \sin(t)x + \cos(t)y \\ z \end{pmatrix}.$$

Show that ϕ_t is a one-parameter group of diffeomorphisms, and compute and graph the vector field $\mathbf{v}$ on $\mathbb{S}^2$ for which ϕ_t is the corresponding flow.

Remark: In defining the vector field $\mathbf{v}$, please specify how you are viewing the tangent bundle of $\mathbb{S}^2$.

Solution. To show that ϕ_t is a one-parameter group of diffeomorphisms, we need to show:

- Each ϕ_t is bijective on $\mathbb{S}^2$.
- Each ϕ_t is smooth with smooth inverse.
- For each s and t we have $\phi_s \circ \phi_t = \phi_{s+t}$.

First note that the map ϕ_t may be written in matrix form as

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \overset{\phi_t}{\to} \begin{pmatrix} \cos(t) & -\sin(t) & 0 \\ \sin(t) & \cos(t) & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

The matrix has determinant $\cos^2(t) + \sin^2(t) = 1 \neq 0$ so it is invertible, hence bijective. In particular its inverse is

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \overset{\phi_t^{-1}}{\to} \begin{pmatrix} \cos(t) & \sin(t) & 0 \\ -\sin(t) & \cos(t) & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

This much shows that ϕ_t is bijective on $\mathbb{R}^3$. To show bijectivity on $\mathbb{S}^2$, we need to show that $\phi_t(\mathbf{q}) \in \mathbb{S}^2$ for all $q \in \mathbb{S}^2$, and likewise for ϕ_t^{-1}. Now, we've already shown that ϕ_t^{-1} is rotation by $-t$ (since $\cos(-t) = \cos(t)$ and $\sin(-t) = \sin(t)$) so it suffices to show this for ϕ_t only. Namely, the coordinates of $(u, v, w) = \phi_t(x, y, z)$ should satisfy $u^2 + v^2 + w^2 = 1$. Check:

$$\begin{aligned} u^2 + v^2 + w^2 &= (\cos(t)x - \sin(t)y)^2 + (\sin(t)x + \cos(t)y)^2 + z^2 \\ &= \cos^2(t)x - \cos(t)\sin(t)xy + \sin^2(t)y^2 \\ &+ \sin^2(t)x^2 + \cos(t)\sin(t)xy + \cos^2(t)y^2 \\ &+ z^2 \\ &= x^2 + y^2 + z^2 = 1. \end{aligned}$$

The transformation is linear, hence it is its own derivative and thus eminently differentiable; likewise for the inverse. (Remember we're checking smoothness for fixed t. Smoothness with respect to t is a different question.) For the composition property, we compute (using the sum formulas for sine and cosine)

$$\begin{aligned} \phi_{s+t} \begin{pmatrix} x \\ y \\ z \end{pmatrix} &= \begin{pmatrix} \cos(s+t) & -\sin(s+t) & 0 \\ \sin(s+t) & \cos(s+t) & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \\ &= \begin{pmatrix} \cos(s)\cos(t) - \sin(s)\sin(t), & -\sin(s)\cos(t) - \cos(s)\sin(t), & 0 \\ \sin(s)\cos(t) + \cos(s)\sin(t), & \cos(s)\cos(t) - \sin(s)\sin(t), & 0 \\ 0, & 0, & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}. \end{aligned}$$

On the other hand,

$$\begin{aligned} \phi_s \circ \phi_t \begin{pmatrix} x \\ y \\ z \end{pmatrix} &= \begin{pmatrix} \cos(s) & -\sin(s) & 0 \\ \sin(s) & \cos(s) & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos(t) & -\sin(t) & 0 \\ \sin(t) & \cos(t) & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \\ &= \begin{pmatrix} \cos(s)\cos(t) - \sin(s)\sin(t), & -\cos(s)\sin(t) - \sin(s)\cos(t), & 0 \\ \sin(s)\cos(t) + \cos(s)\sin(t), & -\sin(s)\sin(t) + \cos(s)\cos(t), & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \end{aligned}$$

which is the same as desired. This was the last item needed to prove that the ϕ_t's form a one-parameter group of diffeomorphisms.

Recalling the necessary formula for the corresponding vector field, we have

$$
\begin{aligned}
\mathbf{v}\Big|_{\left(\begin{smallmatrix} x \\ y \\ z \end{smallmatrix}\right)} &= \frac{\partial \phi_t}{\partial t}\Big|_{t=0} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \\
&= \frac{\partial}{\partial t}\Big|_{t=0} \begin{pmatrix} \cos(t)x - \sin(t)y \\ \sin(t)x + \cos(t)y \\ z \end{pmatrix} \\
&= \begin{pmatrix} -\sin(t)x - \cos(t)y \\ \cos(t)x - \sin(t)y \\ 0 \end{pmatrix}\Bigg|_{t=0} \\
&= \begin{pmatrix} -y \\ x \\ 0 \end{pmatrix}\Bigg|_{\left(\begin{smallmatrix} x \\ y \\ z \end{smallmatrix}\right)} \\
&= -y\,\partial/\partial x + x\,\partial/\partial y.
\end{aligned}
$$

This is of course rotation about the z-axis. [xxx insert figure here.] Here, we are viewing the tangent bundle as directional derivatives.

This is a vector field on $\mathbb{R}^3$. [xxx to do: convert to graph coordinates, or spherical.] $\lhd$

Example 8.2. $\rhd$ Here is a variation of the previous example. Let $\omega = x\,dy \wedge dz + y\,dz \wedge dx + z\,dx \wedge dy$. Compute $\phi_t^*(\omega)$. (In example 9.16 we will revisit this problem using the Lie derivative.)

Solution. Let $c = \cos(t)$ and $s = \sin(t)$. Then, replacing x with $cx - sy$ and y with $sx + cy$, we have

$$
\begin{aligned}
\phi_t^*(\omega) &= (cx - sy)\,d(sx + cy) \wedge dz \\
&+ (sx + cy)\,dz \wedge d(cx - sy) \\
&+ z\,d(cx - sy) \wedge d(sx + cy) \\
\\
&= (cx - sy)\,(sdx \wedge dz + cdy \wedge dz) \\
&+ (sx + cy)\,(cdz \wedge dx - sdz \wedge dy) \\
&+ z\,(cdx - sdy) \wedge (sdx + cdy) \\
\\
&= (cx - sy)\,(-sdz \wedge dx + cdy \wedge dz) \\
&+ (sx + cy)\,(cdz \wedge dx + sdy \wedge dz) \\
&+ z\,(c^2 + s^2) dx \wedge dy \\
\\
&= -cs\,x\,dz \wedge dx + c^2\,x\,dy \wedge dz + s^2\,y\,dz \wedge dx - cs\,y\,dy \wedge dz \\
&+ cs\,x\,dz \wedge dx + s^2\,x\,dy \wedge dz + c^2\,y\,dz \wedge dx + cs\,y\,dy \wedge dz \\
&+ z\,dx \wedge dy \\
\\
&= x\,dy \wedge dz + y\,dz \wedge dx + z\,dx \wedge dy \\
&= \omega.
\end{aligned}
$$

Note that this is a bit messy, only because I chose a form with three terms in it. $\lhd$

9 Lie derivatives and the Cartan calculus

If you do not change the direction in which you are going, you will end up where you are headed.
— Confucius (551-479 B.C.).

9.1 The $X \lrcorner$ operator

xxx note we are revisiting this for tensor fields; xref back to section 4.7.15.

xxx xref back to d section as well.

xxx collect rules here.

$$\begin{aligned} X \lrcorner f &= 0 \\ X \lrcorner \omega^{(1)} &= \omega(X) \\ X \lrcorner (\omega \wedge \eta) &= (X \lrcorner \omega) \wedge \eta + (-1)^{\mathrm{ord}(\omega)} \omega \wedge (X \lrcorner \eta). \end{aligned}$$

Rules for $X \lrcorner \omega$:

$$\begin{aligned} (f\partial/\partial x) \lrcorner (g\, dx \wedge dy \wedge dz) &= fg\, dy \wedge dz \\ (f\partial/\partial y) \lrcorner (g\, dx \wedge dy \wedge dz) &= -fg\, dx \wedge dz \\ &= g\, dz \wedge dx \\ (f\partial/\partial z) \lrcorner (g\, dx \wedge dy \wedge dz) &= fg\, dx \wedge dy. \end{aligned}$$

In general,

$$f\partial/\partial x_k \lrcorner (g\, dx_1 \wedge \cdots \wedge dx_k \wedge \cdots \wedge dx_m) = (-1)^{k-1} fg\, dx_1 \wedge \cdots \wedge \hat{dx}_k \wedge \cdots \wedge dx_m)$$

where the overhat indicates that that term is omitted.

Mnemonic 9.1. Put your left forefinger over the $\partial/\partial x_k$ and your right forefinger over the dx_k. Write down what remains, toggling the minus sign by the number of d's between your fingers.

9.2 Lie derivatives

Let X be a vector field on a manifold M, and T be a tensor field on M. The **Lie derivative** $L_X(T)$ is, intuitively, the **rate of change** of T in the direction of X. We will see that L_X takes tensors to tensors of the same rank. This is in contrast to d and $X \lrcorner$ which are rank-raising and rank-lowering operators.

9.2.1 Lie derivatives on general tensor fields

Definition 9.2. Let X be a vector field on a manifold M, with flow ϕ_t^X. Let T be a tensor on M. Then

$$L_X(T) = \frac{d}{dt}\bigg|_{t=0} (\phi_t^X)^* T.$$

Meaning: How T changes under the flow in the direction of X: (1) Pull back (w/r/t flow — need to have defined pullback in terms of the diffeomorphism map), then (2) differentiate with respect to time.

xxx make a better diagram (and xref to pullback section(s)):

$$\begin{array}{ccc} M & \xrightarrow{\phi_t} & M \\ T \circ \phi_t \downarrow & & T \downarrow \end{array}$$

9.2.2 Lie derivatives on 0-forms

Proposition 9.3. *Let $f : M \to \mathbb{R}$, and let X be a vector field on M. Then*

$$L_X(f) = Xf$$

where the action of X on f is the directional derivative as usual.

Proof. [**Lee2**], proposition 18.9. □

Note that f and Xf are both functions from M to $\mathbb{R}$, i.e. 0-forms on M.

9.2.3 Lie derivatives on vector fields

Definition 9.4. For smooth vector fields X and Y, the **Lie bracket** of X and Y is

$$[X, Y] = XY - YX$$

via

$$[X, Y]f = XYf - YXf$$

for a smooth function f.

Since Xf and Yf are 0-forms, just like f, it makes sense to compute YXf and XYf.

Vector fields on M become a **Lie algebra**: Addition is as usual and multiplication is $[X, Y]$.

Proposition 9.5. $L_X(Y) = [X, Y]$.

Proof. See [**Lee2**], proposition 18.9. □

9.2.4 Properties of L_X and relationships between $X \lrcorner$, L_X, and d

Definition 9.2 is unwieldy. Propositions 9.3 and 9.4 (the proofs of which also use definition 9.2) give us better ways to compute Lie derivatives of 0-forms and vector fields. For forms in general, using a couple more propositions along with the properties we've seen, we can derive some more handy properties that allow us to manipulate L_X's, $X \lrcorner$'s, and d's directly.

[xxx xref:]

Proposition 9.6.

$$\begin{aligned} d(\omega \wedge \eta) &= d(\omega) \wedge \eta + (-1)^{\mathrm{ord}(\omega)}\, \omega \wedge d\eta \\ X \lrcorner (\omega \wedge \eta) &= X \lrcorner (\omega) \wedge \eta + (-1)^{\mathrm{ord}(\omega)}\, \omega \wedge X \lrcorner (\eta) \end{aligned}$$

Proposition 9.7. $L_X(\omega \wedge \eta) = L_X(\omega) \wedge \eta + \omega \wedge L_X(\eta)$.

Proof. This is [**Lee2**], proposition 18.9. □

Remark 9.8. That is, while $X \lrcorner$ and d are **antiderivations**, the Lie derivative is a **derivation**. (See definition 4.126.)

Proposition 9.9. $L_X(Y \lrcorner (\omega)) = L_X Y \lrcorner (\omega) + Y \lrcorner (L_X \omega)$.

Proposition 9.10. *For any smooth vector field X and any smooth differential form ω,*

$$L_X(d\omega) = d(L_X \omega).$$

Proof. This is [**Lee2**], corollaries 18.11 (for 1-forms) and 18.14 (for k-forms). □

What does it mean to contract a one-form?

Proposition 9.11. *For a smooth function f and a vector field X on a manifold M,*

$$X \lrcorner (df) = df(X) = Xf.$$

Proof. The first equality is by definition of the $X \lrcorner$ operator (definition 4.124): we insert X into the argument list of the form df. Since df is a 1-form, we need no more arguments:

$$X \lrcorner (df) = df(X).$$

For the second equality, the only way I know how to do this is using coordinates. The vector field X is a linear combination of the basis vectors $\partial/\partial x_i$, with coefficients being smooth functions, say, v_i. Then, since the dx_i's are the dual basis to the $\partial/\partial x_j$'s,

$$\begin{aligned} df(X) &= \sum_{i=1}^{m} \frac{\partial f}{\partial x_i} dx_i \left(\sum_{j=1}^{m} v_j \frac{\partial}{\partial x_j} \right) = \sum_{i=1}^{m} \frac{\partial f}{\partial x_i} v_i \\ &= \left(\sum_{i=1}^{m} v_i \frac{\partial}{\partial x_i} \right)(f) = Xf. \end{aligned}$$

□

Given these facts, we can now justify a formula which connects L_X, $X \lrcorner$, and d — all using simple algebraic manipulations.

Proposition 9.12 (Cartan's magic formula)**.** *For any smooth vector field X and any smooth differential form ω,*

$$L_X\omega = X \lrcorner (d\omega) + d(X \lrcorner (\omega)).$$

That is,

$$L_X = X \lrcorner \circ d + d \circ X \lrcorner .$$

Mnemonic 9.13. The d operator raises the degree of forms, the $X \lrcorner$ operator lowers degree, and the L_X operator preserves degree. So, with regard to degrees, we have $0 = (1-1) + (-1+1)$. Or,

$$\longrightarrow \quad = \quad \begin{array}{ccc} \nearrow & & \searrow \\ & & \end{array} \quad + \quad \begin{array}{ccc} & & \\ \searrow & & \nearrow \end{array} .$$

Proof. The proof is in three steps: (1) for 0-forms; (2) for 1-forms; (3) for k-forms, using induction.

First suppose ω is a 0-form. Rename ω to f for clarity. Then

$$L_X(f) = Xf$$

whereas

$$X \lrcorner (df) + d(X \lrcorner (f)) = Xf + 0 = Xf.$$

The first term is Xf because of proposition 9.11. The second term is 0 because $X \lrcorner$ is degree-lowering and f is already of degree 0; our convention [xxx write and xref backward] is that $X \lrcorner$ of 0-forms is 0.

Now suppose ω is a 1-form. Then [xxx write and xref backward] ω is a linear combination of 1-forms of the form $u\ dv$, where u and v are 0-forms. Recall that by convention, we omit the wedge in wedge products when one factor is a scalar — e.g. $u\ dv$ means $u \wedge dv$. Using the above propositions, the left-hand side of the magic formula is

$$\begin{aligned} L_X(u\ dv) &= L_X(u) \wedge dv + u \wedge L_X(dv) \\ &= Xu\ dv + u\ L_X(dv) \\ &= Xu\ dv + u\ dL_X(v) \\ &= Xu\ dv + u\ d(Xv). \end{aligned}$$

Recall that while the Lie derivative is a derivation, the $X \lrcorner$ and d operators are antiderivations. For the right-hand side, we have

$$\begin{aligned} d(u\ dv) &= du \wedge dv \\ X \lrcorner (d(u\ dv)) &= X \lrcorner (du \wedge dv) \\ &= X \lrcorner (du) \wedge dv - du \wedge X \lrcorner (dv) \\ &= Xu\ dv - du\ Xv \\ X \lrcorner (u\ dv) &= X \lrcorner (u) \wedge dv + u\ X \lrcorner (dv) \\ &= u\ Xv \\ d(X \lrcorner (u\ dv)) &= d(u\ Xv) \\ &= du\ Xv + u\ d(Xv) \\ X \lrcorner (d(u\ dv)) + d(X \lrcorner (u\ dv)) &= Xu\ dv - du\ Xv + du\ Xv + u\ d(Xv) \\ &= Xu\ dv + u\ d(Xv) \\ &= L_X(u\ dv). \end{aligned}$$

[xxx finish up the induction part.] □

9.3 Lie derivatives and flows

xxx attribute:

Proposition 9.14. *Let V and W be vector fields with associated flows ϕ_s and ψ_t. The following are equivalent:*

- $L_V(W) = 0$.
- $L_W(V) = 0$.
- $[V, W] = 0$
- *W is invariant under the flow of V.*
- *V is invariant under the flow of W.*
- $\psi_t \circ \phi_s = \phi_s \circ \psi_t$.

xxx examples here, or xref to prolrevqual.

9.4 Computations using Lie derivatives

An idea which can be used once is a trick. If it can be used more than once it becomes a method.
— George Polya and Gabor Szego

Let's summarize the results of the previous sections. If an exam question asks you if a form ω is "**invariant** under the flow" of a vector field X, or asks you whether the "flow of X **preserves**" ω, what can you do? Compute the Lie derivative of ω or Y with respect to X: the Lie derivative $L_X(\omega)$ is zero if and only if ω is invariant under the flow of X (proposition 18.16 of [**Lee2**]). Here are our options for carrying that out that computation:

(i) For a vector field Y, $L_X(Y) = [X, Y]$. [xxx type up qual problem and xref to it.]

(ii) For a 0-form f, $L_X(f) = X(f)$.

(iii) For an exact form $d\omega$, $L_X(d\omega) = dL_X(\omega)$.

(iv) For a wedge of forms $\omega \wedge \eta$, $L_X(\omega \wedge \eta) = L_X(\omega) \wedge \eta + \omega \wedge L_X(\eta)$.

This allows us to split up a lot of things. Perhaps not everything imaginable, but hopefully anything we will see on an exam in our first year.

Common qualifier questions are: is a given form ω **invariant** under the flow corresponding to a given vector field X, or does a given flow **preserve** a given form. We can answer these questions in the affirmative by showing that the Lie derivative $L_X(\omega)$ is zero.

Example 9.15. $\triangleright$ On $\mathbb{S}^2$, does the flow of $\partial/\partial\theta$ preserve the area form dA?

Using spherical coordinates, the area form is $dA = \sin\phi\, d\theta \wedge d\phi$. Then we need to compute

$$L_{\partial/\partial\theta}(\sin\phi\, d\theta \wedge d\phi).$$

Remember from [xxx write and xref] that there is an implicit wedge between 0-forms and higher-level forms, so this is really

$$L_{\partial/\partial\theta}(\sin\phi \wedge d\theta \wedge d\phi).$$

Using rule (iv), this splits up as

$$L_{\partial/\partial\theta}(\sin\phi) \wedge d\theta \wedge d\phi \quad + \quad \sin\phi \wedge L_{\partial/\partial\theta}(d\theta) \wedge d\phi \quad + \quad \sin\phi \wedge d\theta \wedge L_{\partial/\partial\theta}(d\phi).$$

Using rule (ii) we can set the first term to zero. Can we use rule (iii) for the remaining two terms? Just because a form traditionally is written with a d in front doesn't mean it is exact. Locally, at least, θ and ϕ are 0-forms. Then we have

$$0 \quad + \quad \sin\phi \wedge dL_{\partial/\partial\theta}(\theta) \wedge d\phi \quad + \quad \sin\phi \wedge d\theta \wedge dL_{\partial/\partial\theta}(\phi) = 0.$$

Since this is identically zero, $\partial/\partial\theta$ preserves the area form dA.

Suppose the question asked about the flow of $\partial/\partial\phi$. The second and third terms would still go away, but the first would not. $\triangleleft$

Example 9.16. $\triangleright$ (This example revisits example 8.2, using the Lie derivative.) Show that the two-form $\omega = x\, dy \wedge dz + y\, dz \wedge dx + z\, dx \wedge dy$ is invariant under the flow of ϕ_t.

Solution. We will compute the Lie derivative [xxx for prolrev, xref forward or reorg] of ω in the direction of $\mathbf{v}$ and [xxx]. Since

$$\mathbf{v} = -y\,\partial/\partial x + x\,\partial/\partial y,$$

we have

$$L_{\mathbf{v}}(x) = -y, \quad L_{\mathbf{v}}(y) = x, \quad L_{\mathbf{v}}(z) = 0.$$

Recall that $L_{\mathbf{v}}$ is a derivation, i.e. it follows the product rule through wedges. Also, it commutes with d for exact forms. For the first term of ω, namely $x\,dy \wedge dz$, we then have

$$\begin{aligned} L_{\mathbf{v}}(x\,dy \wedge dz) &= L_{\mathbf{v}}(x) dy \wedge dz \;+\; x\,d(L_{\mathbf{v}}(y)) \wedge dz \;+\; x\,dy \wedge d(L_{\mathbf{v}}(z)) \\ &= -y\,dy \wedge dz \;+\; x\,dx \wedge dz \;+\; x\,dy \wedge 0 \\ &= -y\,dy \wedge dz \;-\; x\,dz \wedge dx. \end{aligned}$$

For the second term:

$$\begin{aligned} L_{\mathbf{v}}(y\,dz \wedge dx) &= L_{\mathbf{v}}(y) dz \wedge dx \;+\; y\,d(L_{\mathbf{v}}(z)) \wedge dx \;+\; y\,dz \wedge d(L_{\mathbf{v}}(x)) \\ &= x\,dz \wedge dx \;+\; y\,0 \wedge dx \;-\; y\,dz \wedge dy \\ &= x\,dz \wedge dx \;+\; y\,dy \wedge dz. \end{aligned}$$

For the third term:

$$\begin{aligned} L_{\mathbf{v}}(z\,dx \wedge dy) &= L_{\mathbf{v}}(z) dx \wedge dy \;+\; z\,d(L_{\mathbf{v}}(x)) \wedge dy \;+\; z\,dx \wedge d(L_{\mathbf{v}}(y)) \\ &= 0\,dx \wedge dy \;-\; z\,dy \wedge dy \;+\; z\,dx \wedge dx \\ &= 0. \end{aligned}$$

Combining these, we have

$$\begin{aligned} L_{\mathbf{v}}(\omega) &= -y\,dy \wedge dz \;-\; x\,dz \wedge dx \\ &+ \; x\,dz \wedge dx \;+\; y\,dy \wedge dz \\ &= 0. \end{aligned}$$

which shows that ω is invariant under the flow of ϕ_t. $\triangleleft$

[xxx do it again using the magic formula. Much nicer.]

[xxx type up the August 06 qual problem and xref to it.]

10 Problems

Each problem that I solved became a rule which served afterwards to solve other problems.

— René Descartes (1596-1650).

Here are various problems, along with solutions. The problems are taken from homework assignments, exam reviews, and exams.

10.1 Geometry homework 1

10.1.1 Geometry homework 1 problem 2d

Find Df where $f : \mathcal{L}(\mathbb{R}^n) \to \mathbb{R}$ is given by $A \mapsto \det(A)$.

Hint: First consider the case when A is invertible, i.e. $\det(A) \neq 0$. Then by the multiplicativity of the determinant,

$$\left.\frac{d}{dt}\right|_{t=0} \det(A + tB) = \left.\frac{d}{dt}\right|_{t=0} \det(I + tA^{-1}B)\det(A).$$

Next, recall that the exponential map from $\mathcal{L}(\mathbb{R}^n)$ to $\mathrm{GL}(\mathbb{R}^n)$ has the form

$$e^A = I + A + A^2/2! + \ldots$$

and so

$$e^{tA^{-1}B} = I + tA^{-1}B + (\text{higher-order terms}).$$

Thus

$$\left.\frac{d}{dt}\right|_{t=0} \det(I + tA^{-1}B)\det(A) = \left.\frac{d}{dt}\right|_{t=0} \det(e^{tA^{-1}B})\det(A).$$

Answer (due largely to Andy Lebovitz). Recall that $\det(e^X) = e^{\mathrm{tr}(X)}$. So,

$$D\det|_A(B) = \left.\frac{d}{dt}\right|_{t=0} e^{\mathrm{tr}(tA^{-1}B}\det(A) = \det(A)\left.\frac{d}{dt}\right|_{t=0} e^{\mathrm{tr}(tA^{-1}B)}.$$

Since the trace is linear, $\mathrm{tr}(tA^{-1}B = t \cdot \mathrm{tr}(A^{-1}B)$. Also note that $A^{-1}B$ is constant with respect to t. Then

$$\begin{aligned} D\det|_A(B) &= \det(A)\left.\frac{d}{dt}\right|_{t=0} e^{t\cdot\mathrm{tr}(A^{-1}B)} \\ &= \det(A)\mathrm{tr}(A^{-1}B)\left.\frac{d}{dt}\right|_{t=0} e^t \\ &= \det(A)\mathrm{tr}(A^{-1}B). \end{aligned}$$

Continuing to follow the hint, consider non-invertible A. Recall that when A is invertible [xxx have defined adjugate somewhere above],

$$A^{-1} = \frac{\mathrm{adju}(A)}{\det(A)}.$$

From above,

$$D\det|_A(B) = \mathrm{tr}\left(\frac{\mathrm{adju}(A)B}{\det(A)}\right)\det(A).$$

Since the trace is linear,

$$D \det|_A(B) = \operatorname{tr}(\operatorname{adju}(A)B).$$

This formula applies for non-singular A. Note that tr is linear, the adjugate is polynomial in the coefficients of A, and $\mathrm{GL}(\mathbb{R}^n)$ is dense in $\mathcal{L}(\mathbb{R}^n)$. Thus by continuity, this formula holds for all A.

10.2 Geometry homework 6

10.2.1 Geometry homework 6 problem 1

Suppose $z = g(y)$ is a positive function, and let M denote the surface of revolution obtained by revolving the graph of g around the y axis in $\mathbb{R}^3$. Show that the tangent bundle of M is trivial by finding an explicit trivialization.

Answer. It suffices to apply the **global frame criterion** as defined in section 6.2.4. Let $\mathbf{q} \in M$. Then in rectangular coordinates for $\mathbb{R}^3$ we can write

$$\mathbf{q} = \begin{pmatrix} g(y)\cos\theta \\ y \\ g(y)\sin\theta \end{pmatrix}.$$

We can parameterize this surface using y and θ coordinates, where θ is the angle from the x axis to $\mathbf{q}$. As in section 6.3.4, we compute

$$\frac{\partial}{\partial y} = \begin{pmatrix} g'(y)\cos\theta \\ 1 \\ g'(y)\sin\theta \end{pmatrix}, \quad \frac{\partial}{\partial \theta} = \begin{pmatrix} -g(y)\sin\theta \\ 0 \\ g(y)\cos\theta \end{pmatrix}.$$

It suffices to show that the sections $\partial/\partial y$ and $\partial/\partial\theta$ form a global frame, i.e. that they are linearly independent and globally defined on M. Clearly, they are globally defined on M. (Unlike, say, [xxx xref] $\mathbb{S}^2$, where our various coordinates leave some part of the sphere uncovered.) It is also clear that $\partial/\partial y$ and $\partial/\partial\theta$ are smooth on M, as long as $g(y)$ is smooth on $\mathbb{R}$.

For linear independence, we can show that (1) both sections are non-zero, and (2) their dot product is identically zero. For the first claim, $\frac{\partial}{\partial y}$ is nowhere zero since $g(y)$ is assumed non-zero, and since sine and cosine have no common zeroes. Likewise, $\partial/\partial\theta$ is nowhere zero due to the 1 in the y component. For the second claim, we compute the dot product

$$\frac{\partial}{\partial y} \cdot \frac{\partial}{\partial \theta} = \begin{pmatrix} g'(y)\cos\theta \\ 1 \\ g'(y)\sin\theta \end{pmatrix} \cdot \begin{pmatrix} -g(y)\sin\theta \\ 0 \\ g(y)\cos\theta \end{pmatrix} = -g(y)g'(y)\sin\theta\cos\theta + g(y)g'(y)\sin\theta\cos\theta = 0.$$

10.3 Geometry exam 1

10.3.1 Geometry exam 1 problem 6

Fix $0 < r < R$. Consider the manifold M obtained by revolving the circle $(y-R)^2 + z^2 = r^2$ around the z-axis in $\mathbb{R}^3$. As we discussed in class, there is an explicit diffeomorphism

$$\mathbb{S}^1 \times \mathbb{S}^1 \to M : (e^{i\theta}, e^{i\phi}) \longmapsto (R + r\cos\phi)) \begin{pmatrix} \cos\theta \\ \sin\theta \\ 0 \end{pmatrix} + r\sin\phi \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Express the Euclidean inner product (or metric), viewed as a 2-tensor field on M, in terms of the coordinates θ, ϕ.

Answer. See example 4.123. Here we are doing the same thing, but with varying coefficients in the sense of sections 6.2.1 and 6.50. A point $\mathbf{q}$ of M is

$$\mathbf{q} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = (R + r\cos\phi) \begin{pmatrix} \cos\theta \\ \sin\theta \\ 0 \end{pmatrix} + r\sin\phi \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

so we can write

$$\frac{\partial}{\partial\theta} = (R + r\cos\phi) \begin{pmatrix} -\sin\theta \\ \cos\theta \\ 0 \end{pmatrix} \quad \text{and} \quad \frac{\partial}{\partial\phi} = (-r\sin\phi) \begin{pmatrix} \sin\theta \\ \cos\theta \\ 0 \end{pmatrix} + r\cos\phi \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

As in example 4.123, we want to write the metric tensor g as

$$g = g\left(\frac{\partial}{\partial\theta}, \frac{\partial}{\partial\theta}\right) d\theta \otimes d\theta + g\left(\frac{\partial}{\partial\phi}, \frac{\partial}{\partial\phi}\right) d\phi \otimes d\phi$$

where we can omit the diagonal terms $g(\partial/\partial\theta, \partial/\partial\phi)$ etc. since g is diagonal and $\{\partial/\partial\theta, \partial/\partial\phi\}$ is an orthogonal basis. First we find $g(\partial/\partial\theta, \partial/\partial\theta)$. Recalling that dx extracts the x coordinate, etc., this is

$$\begin{aligned} g\left(\frac{\partial}{\partial\theta}, \frac{\partial}{\partial\theta}\right) &= (dx \otimes dx + dy \otimes dy + dz \otimes dz)\left(\frac{\partial}{\partial\theta}, \frac{\partial}{\partial\theta}\right) \\ &= dx \otimes dx\left(\frac{\partial}{\partial\theta}, \frac{\partial}{\partial\theta}\right) + dy \otimes dy\left(\frac{\partial}{\partial\theta}, \frac{\partial}{\partial\theta}\right) + dz \otimes dz\left(\frac{\partial}{\partial\theta}, \frac{\partial}{\partial\theta}\right) \\ &= (R + r\cos\phi)^2 \sin^2\theta + (R + r\cos\phi)^2 \cos^2\theta + (0)^2 \\ &= (R + r\cos\phi)^2. \end{aligned}$$

Next we find $g(\partial/\partial\phi, \partial/\partial\phi)$:

$$\begin{aligned} g\left(\frac{\partial}{\partial\phi}, \frac{\partial}{\partial\phi}\right) &= dx \otimes dx\left(\frac{\partial}{\partial\phi}, \frac{\partial}{\partial\phi}\right) + dy \otimes dy\left(\frac{\partial}{\partial\phi}, \frac{\partial}{\partial\phi}\right) + dz \otimes dz\left(\frac{\partial}{\partial\phi}, \frac{\partial}{\partial\phi}\right) \\ &= r^2 \sin^2\theta \sin^2\phi + r^2 \cos^2\theta \sin^2\phi + r^2 \cos^2\phi \\ &= r^2(\sin^2\theta + \cos^2\theta \sin^2\phi) \sin^2\phi + r^2 \cos^2\phi \\ &= r^2 \sin^2\phi + r^2 \cos^2\phi \\ &= r^2. \end{aligned}$$

So, we have

$$\begin{aligned} g &= g\left(\frac{\partial}{\partial\theta}, \frac{\partial}{\partial\theta}\right) d\theta \otimes d\theta + g\left(\frac{\partial}{\partial\phi}, \frac{\partial}{\partial\phi}\right) d\phi \otimes d\phi \\ &= (R + r\cos\phi)^2 d\theta \otimes d\theta + r^2 d\phi \otimes d\phi. \end{aligned}$$

10.4 Geometry homework 11

10.4.1 Geometry homework 11 problem 2

Let x, y, z denote linear coordinates for $\mathbb{R}^3$. Let

$$\nabla = \begin{pmatrix} \partial/\partial x \\ \partial/\partial y \\ \partial/\partial z \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \quad \mathbf{A} = \begin{pmatrix} A \\ B \\ C \end{pmatrix}, \quad \text{and} \quad dV = dx \wedge dy \wedge dz.$$

Show that for the exterior derivative

$$\begin{array}{lrcl} (a) & d(\mathbf{A} \cdot d\mathbf{x}) & = & (\nabla \times \mathbf{A}) \lrcorner\, dV \\ (b) & d(\mathbf{A} \lrcorner\, dV) & = & \operatorname{div}(\mathbf{A}) dV. \end{array}$$

Answer. The key is that dV is the **volume operator**, and as such is represented by a determinant. (Without this observation this problem can be become a time-consuming mess.)

For part (a), first we compute $\mathbf{A} \cdot d\mathbf{x}$. Using the shorthand notation $A_x = \partial A/\partial x$, etc., we have

$$\mathbf{A} \cdot d\mathbf{x} = A dx + B dy + C dz.$$

Next we find d of this. Recalling that [xxx cite]

$$dA = A_x dx + A_y dy + A_z dz,$$

etc., we have

$$\begin{array}{rcl} d(\mathbf{A} \cdot d\mathbf{x}) & = & (A_x dx + A_y dy + A_z dz) \wedge dx \\ & + & (B_x dx + B_y dy + B_z dz) \wedge dy \\ & + & (C_x dx + C_y dy + C_z dz) \wedge dz \\ & & \\ & = & (A_y dy + A_z dz) \wedge dx \\ & + & (B_x dx + B_z dz) \wedge dy \\ & + & (C_x dx + C_y dy) \wedge dz \\ & & \\ & = & A_y dy \wedge dx + A_z dz \wedge dx \\ & + & B_x dx \wedge dy + B_z dz \wedge dy \\ & + & C_x dx \wedge dz + C_y dy \wedge dz \\ & & \\ & = & (B_x - A_y) dx \wedge dy + (C_y - B_z) dy \wedge dz + (A_z - C_x) dz \wedge dx. \end{array}$$

We also need to write out $\nabla \times \mathbf{A}$. This is

$$\nabla \times \mathbf{A} = \begin{vmatrix} \hat{\mathbf{x}} & \hat{\mathbf{y}} & \hat{\mathbf{z}} \\ \partial/\partial x & \partial/\partial y & \partial/\partial z \\ A & B & C \end{vmatrix} = \begin{pmatrix} C_y - B_z \\ A_z - C_x \\ B_x - A_y \end{pmatrix}$$

Let $\mathbf{v}, \mathbf{w} \in \mathbb{R}^3$. Then, since dV of m vectors in $\mathbb{R}^m$ is the volume of the parallepiped spanned by them, and using the definition of the contraction operator from [xxx cite], we have

$$\begin{aligned}
(\nabla \times \mathbf{A}) \lrcorner\, dV(\mathbf{v}, \mathbf{w}) = \det \begin{pmatrix} \nabla \times \mathbf{A} \\ \hline \mathbf{v} \\ \hline \mathbf{w} \end{pmatrix} &= \begin{vmatrix} C_y - B_z & A_z - C_x & B_x - A_y \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix} \\
&= (C_y - B_z)(v_2 w_3 - v_3 w_2) \\
&+ (A_z - C_x)(v_3 w_1 - v_1 w_3) \\
&+ (B_x - A_y)(v_1 w_2 - v_2 w_1).
\end{aligned}$$

At this point we should simply *see* that

$$v_2 w_3 - v_3 w_2 = dy \wedge dz(\mathbf{v}, \mathbf{w}),$$

etc. so

$$\begin{aligned}
(\nabla \times \mathbf{A}) \lrcorner\, dV &+ (B_x - A_y) dx \wedge dy \\
&+ (C_y - B_z) dy \wedge dz \\
&= (A_z - C_x) dz \wedge dx.
\end{aligned}$$

as desired.

* * *

For part (b), first we will find out what $\mathbf{A} \lrcorner\, dV$ is, then compute d of it, then compare that to $\operatorname{div}(A) dV$.

As in part (a), let $\mathbf{v}, \mathbf{w} \in \mathbb{R}^3$. Then

$$\begin{aligned}
\mathbf{A} \lrcorner\, dV(\mathbf{v}, \mathbf{w}) = \det \begin{pmatrix} \mathbf{A} \\ \hline \mathbf{v} \\ \hline \mathbf{w} \end{pmatrix} &= \begin{vmatrix} A & B & C \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix} \\
&= A(v_2 w_3 - v_3 w_2) + B(v_3 w_1 - v_1 w_3) + C(v_1 w_2 - v_2 w_1).
\end{aligned}$$

Again observe that

$$v_2 w_3 - v_3 w_2 = dy \wedge dz(\mathbf{v}, \mathbf{w}),$$

etc., so

$$\mathbf{A} \lrcorner\, dV(\mathbf{v}, \mathbf{w}) = A\, dy \wedge dz + B\, dz \wedge dx + C\, dx \wedge dy.$$

Then, it is easy to compute d of this [xxx cite justifcations for steps]:

$$\begin{aligned}
d(\mathbf{A} \lrcorner\, dV(\mathbf{v}, \mathbf{w})) &= d(A\, dy \wedge dz + B\, dz \wedge dx + C\, dx \wedge dy) \\
&= d(A\, dy \wedge dz) + d(B\, dz \wedge dx) + d(C\, dx \wedge dy)
\end{aligned}$$

Now [xxx cite], and using the shorthand $A_x = \partial A / \partial x$, etc.,

$$dA = A_x dx + A_y dy + A_z dz,$$

etc. When we FOIL this out, we'll get 9 terms, but 6 of them are zero since $dx \wedge dx = 0$ etc. So,

$$\begin{aligned}
d(\mathbf{A} \lrcorner\, dV(\mathbf{v}, \mathbf{w})) &= A_x dx \wedge dy \wedge dz + B_y dx \wedge dy \wedge dz + C_z dx \wedge dy \wedge dz \\
&= (A_x + B_y + C_z) dx \wedge dy \wedge dz \\
&= \operatorname{div}(\mathbf{A}) dV.
\end{aligned}$$

10.5 Geometry exam 2

10.5.1 Geometry exam 2 problem 1

Let $\mathbf{v}$ denote the vector field on $\mathbb{R}^2$ given by

$$\mathbf{v}\Big|_{\binom{x}{y}} = \begin{pmatrix} x^2 y \\ -y \end{pmatrix}\Big|_{\binom{x}{y}}.$$

Compute the flow of this vector field.

Answer. We have

$$\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = \begin{pmatrix} x^2 y \\ -y \end{pmatrix}$$

which immediately gives $\dot{y} = y$ with solution $y = y_0 e^t$. Plugging this into the first equation gives

$$\begin{aligned} \frac{dx}{dt} &= x^2 y_0 e^t \\ x^{-2} dx &= y_0 e^t dt \\ \int x^{-2} dx &= y_0 e^t + C \\ -1/x &= y_0 e^t + C \end{aligned}$$

...finish typing up later.

10.5.2 Geometry exam 2 problem 2

Recall that we can identify the tangent space of $\mathbb{S}^2$ at a point $\mathbf{q}$ with $\{(\mathbf{v}, \mathbf{q}) : \mathbf{v} \in \mathbb{R}^3, \mathbf{q} \perp \mathbf{v}\}$. Let η denote the 1-form (or 1-tensor) which is given by

$$\eta((\mathbf{q}, \mathbf{v})) = v_1$$

where v_1 is the first component of $\mathbf{v}$.

(a) Express η in terms of spherical coordinates ϕ, θ for $\mathbb{S}^2$, where $0 < \phi < \pi$ and $0 < \theta < 2\pi$.

(b) Find two reasons that $d\eta = 0$.

Answer. See example 4.120. Here we are doing the same thing, but with varying coefficients in the sense of section 6.2.1. Using that approach, we have

$$\eta = \eta\left(\frac{\partial}{\partial \theta}\right) d\theta + \eta\left(\frac{\partial}{\partial \phi}\right) d\phi.$$

The point $\mathbf{q}$ is

$$\mathbf{q} = \begin{pmatrix} \cos\theta \sin\phi \\ \sin\theta \sin\phi \\ \cos\phi \end{pmatrix}.$$

Thus

$$\partial/\partial\theta = \mathbf{q} = \begin{pmatrix} -\sin\theta \sin\phi \\ \cos\theta \sin\phi \\ 0 \end{pmatrix}, \qquad \partial/\partial\phi = \mathbf{q} = \begin{pmatrix} \cos\theta \cos\phi \\ \sin\theta \cos\phi \\ -\sin\phi \end{pmatrix}.$$

Now, η extracts the first coordinate, so

$$\eta = -\sin\theta\sin\phi d\theta + \cos\theta\cos\phi d\phi.$$

For part (b), we might like to say that $d^2 = 0$. However, it is an unfortunate fact [xxx xref to the appropriate section once I write it] that not everything that begins with d is exact. [in the xref, say whether $d\theta$ and $d\phi$ are in fact exact. I guess that's the same as asking whether θ and ϕ are 0-forms. Also in the xref, talk about dA, etc.] Let's compute $d\eta$ using the above expression. We use the fact that d is an antiderivation, so it is linear and has a product rule. We get

$$\begin{aligned} d\eta &= -d(\sin\theta\sin\phi d\theta) + d(\cos\theta\cos\phi d\phi) \\ &= -d(\sin\theta)\sin\phi d\theta - \sin\theta d(\sin\phi)d\theta - \sin\theta\sin\phi dd\theta \\ &\quad +d(\cos\theta)\cos\phi d\phi + \cos\theta d(\cos\phi)d\phi + \cos\theta\cos\phi dd\phi \end{aligned}$$

Now, dd of anything is zero, and $d\theta \wedge d\theta = 0$ and likewise for $d\phi \wedge d\phi$, so we have

$$d\eta = -\sin\theta\cos\phi d\phi \wedge d\theta - \sin\theta d\theta\cos\phi \wedge d\phi$$

But $d\phi \wedge d\theta = -d\theta \wedge d\phi$ so this is zero.

10.5.3 Geometry exam 2 problem 3

Consider the manifold $M = \{(x, y, z) \in \mathbb{R}^3 : z = x^2 + y^2\}$ with the upward-pointing orientation.

(a) Compute the area form of M in x, y coordinates.

(b) Let

$$h : M \to \mathbb{R}^2 : \begin{pmatrix} x \\ y \\ z \end{pmatrix} \longmapsto \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} xy \\ z \end{pmatrix}.$$

If $du \wedge dv$ denotes the standard area form on $\mathbb{R}^2$, then $h^*(du \wedge dv)$ is a two-form on M. Express this two-form in x, y coordinates.

Answer to part (a). Proceeding as in [xxx xref], we note that M is a warp of the plane, so the graph coordinates x, y are appropriate. Then a point $\mathbf{q} \in M$ is, in the x, y parameterization,

$$\mathbf{q} = \begin{pmatrix} x \\ y \\ x^2 + y^2 \end{pmatrix}.$$

Since we are using x, y coordinates, a basis for the tangent space of M is $\{\partial/\partial x, \partial/\partial y\}$, and a basis for the cotangent space of M is $\{dx, dy\}$. As in [xxx], we compute

$$\frac{\partial}{\partial x} = \begin{pmatrix} 1 \\ 0 \\ 2x \end{pmatrix} \quad \text{and} \quad \frac{\partial}{\partial y} = \begin{pmatrix} 0 \\ 1 \\ 2y \end{pmatrix}.$$

As in example 4.121, for the area form dA we have

$$dA = dA\left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}\right) dx^\wedge dy.$$

so it remains to find out what $dA(\partial/\partial x, \partial/\partial y)$ is.

Unlike in example [xxx], here we need a little trick: from vector calculus we know that the area of the parallelogram spanned by two vectors is the signed magnitude of their cross product. The cross product, in turn, is just the normal vector to the surface:

$$\mathbf{n} = \frac{\partial}{\partial x} \times \frac{\partial}{\partial y} = \begin{vmatrix} \hat{\mathbf{x}} & \hat{\mathbf{y}} & \hat{\mathbf{z}} \\ 1 & 0 & 2x \\ 0 & 1 & 2y \end{vmatrix} = \begin{pmatrix} -2x \\ -2y \\ 1 \end{pmatrix}$$

The magnitude of this is

$$\|\mathbf{n}\| = \sqrt{4x^2 + 4y^2 + 1}.$$

To get the sign correct, use the right-hand rule: moving from $\partial/\partial x$ to $\partial/\partial y$ we would expect $\mathbf{n}$ to point upward, so we want the positive sign, so we can leave the sign as is. In conclusion, we have

$$dA = (4x^2 + 4y^2 + 1)^{1/2} dx \wedge dy.$$

Answer to part (b). We have

$$h : M \to \mathbb{R}^2 : \begin{pmatrix} x \\ y \\ z \end{pmatrix} \longmapsto \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} xy \\ z \end{pmatrix}.$$

Then using [xxx write and xref; also include a pullback diagram], we have

$$\begin{aligned} h^*(du \wedge dv) &= h^*(du) \wedge h^*(dv) = d(u \circ h) \wedge d(v \circ h) \\ &= d(xy) \wedge dz = (y\,dx + x\,dy) \wedge dz = (y\,dx + x\,dy) \wedge d(x^2 + y^2) \\ &= (y\,dx + x\,dy) \wedge (2x\,dx + 2y\,dy) = 2y^2 dx \wedge dy - 2x^2 dx \wedge dy \\ &= (2y^2 - 2x^2) dx \wedge dy. \end{aligned}$$

Note that the key to this problem is pulling back correctly; the rest is straightforward computation.

10.6 Geometry final review

10.6.1 Geometry final review problem 1

(a) Calculate the derivative of $\det : \mathcal{L}(\mathbb{R}^n) \to \mathbb{R}$.

(b) Determine the set of critical points for det.

Answer to part (a). This was done in homework 1, section 10.1.1. I remark that the derivation of the formula required several non-obvious tricks, so for an exam one would perhaps be advised to simply memorize the formulas

$$D\det|_A(B) = \det(A)\mathrm{tr}(A^{-1}B) \qquad (\det(A) \neq 0)$$

and

$$D\det|_A(B) = \mathrm{tr}(\mathrm{adju}(A)B) \qquad (\text{general } A).$$

Answer to part (b). Critical points A of det occur when $D\det|_A$ is not surjective. Since the target space $\mathbb{R}$ has dimension 1 over $\mathbb{R}$, the rank of $D\det|_A$ is either 0 or 1. We need to find points A such that $D\det|_A$ is zero. This is precisely when $\mathrm{adju}(A)$ is 0. That is, if A is $n \times n$, then $\mathrm{adju}(A)$ is zero when all $(n-1)\times(n-1)$ submatrices of A are singular.

This result is perhaps surprising: we might have guessed that the critical points of det would be all the singular matrices, i.e. those of rank less than n. However, the above tells us that the matrices of rank $n-1$ are also regular points for the det function. It is the matrices of rank $n-2$ or less which are critical.

10.6.2 Geometry final review problem 2

(a) State the regular value theorem.

(b) Prove that $\mathrm{SL}(n, \mathbb{R}) = \{A \in \mathcal{L}(\mathbb{R}^n) : \det(A) = 1\}$ is an embedded submanifold of $\mathcal{L}(\mathbb{R}^n)$. What is its dimension?

(c) Calculate the tangent space to $\mathrm{SL}(\mathbb{R}^n)$ at the identity matrix.

Answer for part (a). See theorem 6.17 in section 6.1.5: if $f : M \to N$ is a map of manifolds, and if $\mathbf{c} \in N$ is a regular value of f, then $f^{-1}(\mathbf{c})$ is either the empty set, or an embedded submanifold of M. In the latter case, $f^{-1}(\mathbf{c})$ has dimension $m - n$.

Answer for part (b). Apply the regular value theorem, with $M = \mathcal{L}(\mathbb{R}^n)$, $f = \det$, $N = \mathbb{R}$, and $\mathbf{c} = 1$. It remains only to show that 1 is a regular value of det. In turn, 1 is a regular value of det if all A in $\det^{-1}(A)$ are regular points. This would mean $D\det|_A$ is surjective when $\det(A) = 1$. Since $\mathbb{R}$ is a one-dimensional vector space over $\mathbb{R}$, it suffices to show that $D\det|_A$ is not identically zero when $\det(A) = 1$. From the previous problem, we can use

$$D\det|_A(B) = \det(A)\mathrm{tr}(A^{-1}B).$$

Since $\det(A) = 1$, $D\det|_A(B)$ won't vanish for that reason. The only remaining task is to argue that for fixed A, the B-linear function $\mathrm{tr}A^{-1}B$ is not identically zero. If it were identically zero, it would be true for $B = A$. But $\mathrm{tr}A^{-1}A = \mathrm{tr}I = n \neq 0$ where n is the dimension of A. This is a contradiction, so $D\det|_A$ is not identically zero.

By the regular value theorem, since the source space $\mathcal{L}(\mathbb{R}^n)$ has dimension n^2 and the destination space $\mathbb{R}$ has dimension 1, $\mathrm{SL}(\mathbb{R}^n)$ has dimension $n^2 - 1$.

Answer for part (c) (Pickrell). As discussed in remark 6.22, the tangent space to $\mathrm{SL}(\mathbb{R}^n)$ at I is $\ker(D\det|_I)$. Now,

$$D\det|_I(B) = tr(B)$$

so the tangent space at I consists of the matrices with zero trace.

10.6.3 Geometry final review problem 3

(a) Determine the critical points and critical values for the function

$$F : \mathbb{R}^3 \to \mathbb{R}^2 : \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} z^2 - y^2 - x^2 \\ z - y \end{pmatrix}.$$

(b) Show that *all* of the level sets of F are embedded submanifolds (of various dimensions).

Answer for part (a). From section 6.1.5 we know that F is critical at all points $\mathbf{q} \in \mathbb{R}^3$ such that DF has rank less than 2. We can use the adjugate criterion (proposition 6.23) to find such points. The Jacobian of F is

$$DF = \begin{pmatrix} -2x & -2y & 2z \\ 0 & -1 & 1 \end{pmatrix}.$$

The determinants of 2×2 submatrices are

$$2x, -2x, \quad \text{and} \quad 2z - 2y.$$

The critical points of F are where these are simultaneously zero, i.e. when

$$x = 0 \quad \text{and} \quad z = y.$$

Inserting these into the expression for F gives

$$F\begin{pmatrix} 0 \\ y \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Answer for part (b). By the regular value theorem, the level sets for regular values are embedded submanifolds. It remains to show that the level set of the critical value $(0, 0)$ is also an embedded submanifold. But that level set is just the line $z = y$. This is certainly an embedded submanifold of $\mathbb{R}^3$, since the inclusion map is globally 1-1.

Remark 10.1 (Pickrell)**.** This does not contradict the regular value theorem. The regular value theorem says that if $\mathbf{c}$ is a regular value of the map $f : M \to N$, then $f^{-1}(\mathbf{c})$ is an embedded submanifold of M. Here we have an example of a critical value, the inverse image of which also happens to be an embedded submanifold. This shows that the converse to the regular value theorem does not hold.

10.6.4 Geometry final review problem 4

Determine whether the function

$$\chi : \mathbb{S}^2 \to \mathbb{R}^2 : \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} x^2 - y^2 \\ 2xy + z \end{pmatrix}$$

is a coordinate for $\mathbb{S}^2$ in the neighborhood of the north pole.

Answer. It would be tempting to compute the rank of χ by taking the 2×3 Jacobian, but note that we have $\mathbb{S}^2 \hookrightarrow \mathbb{R}^3 \to \mathbb{R}^2$, and χ as shown above is a map from $\mathbb{R}^3$ to $\mathbb{R}^2$. We need to somehow restrict χ to $\mathbb{S}^2$.

My approach is that a coordinate chart on $\mathbb{S}^2$ does give a two-dimensional surface. Consider a known-good coordinate chart, namely the graph coordinates x, y. If χ were also a coordinate chart, then the transition function from graph coordinates to χ would be a 1-1 function. [xxx include a picture here, with two coordinate charts mapping out of the upper hemisphere.] Let ψ be the transition function:

$$\psi \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x^2 - y^2 \\ 2xy + (1 - x^2 - y^2)^{1/2} \end{pmatrix}$$

Now, ψ is a non-linear function, but we can use the inverse function theorem. That is, by linearizing, we can use the Jacobian test:

$$D\psi = \begin{pmatrix} 2x & -2y \\ 2y - \frac{x}{(1-x^2-y^2)^{1/2}} & 2x - \frac{y}{(1-x^2-y^2)^{1/2}} \end{pmatrix}$$

At the point $x = y = 0$, this is the zero matrix. So, ψ is not 1-1 at the north pole, which is all that remained to be shown.

10.6.5 Geometry final review problem 5

(a) Compute the tangent bundle of the manifold

$$N = \{(x, y, z) \in \mathbb{R}^3 : z^2 - x^2 - y^2 = 1\}.$$

(b) What is the dimension of TN, as a manifold?

Remark. [xxx insert picture here.] To get an idea for what this looks like, fix $y = 0$ and graph that cross-section on the plane. Then by rotational symmetry you'll see that near the origin $x = y = 0$, it looks like a paraboloid, but away from the origin $x = y = 0$ it straightens out. So, it's like a cone with a rounded end.

First answer to part a. This solution uses the technique shown in section 6.3.2. This uses vector calculus and linear algebra, and is a bit tedious. The next, more efficient, solution uses the technique of section 6.3.4.

The tangent plane is perpendicular to the normal vector. To get the normal vector, in turn, we can observe that N is the level set of 1 and $F(x, y, z) = z^2 - x^2 - y^2$. As in section 1.5.4, the equation for the normal is given by ∇F. This is

$$\nabla F = \begin{pmatrix} -2x \\ -2y \\ 2z \end{pmatrix}.$$

Let $\mathbf{q} = (x, y, z)$ be an arbitrary point on N. Then the normal vector footed at $\mathbf{q}$ is $\mathbf{n_q}$, as given above. Any element $\mathbf{v_q}$ of the tangent space $TM_\mathbf{q}$ is perpendicular to $\mathbf{n_q}$, i.e.

$$\mathbf{n_q} \cdot \mathbf{v_q} = 0.$$

This means

$$\begin{pmatrix} n_1 & n_2 & n_3 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} -2x & -2y & 2z \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = (-2x)v_1 + (-2y)v_2 + (2z)v_3 = 0.$$

This is a linear system, where $\mathbf{v_q}$ is the variable we want to solve for. The coefficients are (x, y, z). These look like variables, but they're constant for each point $\mathbf{q} = (x, y, z)$. Since the matrix has rank 1, and we are computing its kernel, from the rank-nullity theorem we expect kernel dimension 2. So, since there are two degrees of freedom, we first choose $v_1 = 1$ and $v_2 = 0$, then $v_1 = 0$ and $v_2 = 1$. It remains to compute what the v_3 coordinate must be in each case. The factor of -2 is irrelevant, so the equation to be solved is

$$xv_1 + yv_2 - zv_3 = 0.$$

For the first and second choices, respectively, we have

$$x - zv_3 = 0 \quad \text{and} \quad y - zv_3 = 0.$$

These give $v_3 = x/z$ and y/z, respectively, and so a basis for the tangent space at $\mathbf{q}$ is

$$\{(1, 0, x/z), (0, 1, y/z)\} \qquad \text{or} \qquad \{(z, 0, x), (0, z, y)\}.$$

Let's do a quick sanity check. Pick $x = 0$ and $y = 1$ on N. Then $z^2 = x^2 + y^2 + 1$, i.e. $z^2 = 2$. Take the upper point $\mathbf{q} = (x, y, z) = 0, 1, \sqrt{2}$. Then the tangent space at $(0, 1, \sqrt{2})$ should be spanned by $(\sqrt{2}, 0, 0)$ and $(\sqrt{0}, 2, 1)$. This makes sense: [xxx need a picture here.] As a second check, take $x = y = 0$ and $z = 1$. Then the basis vectors for the tangent space are just $(1, 0, 0)$ and $(0, 1, 0)$, which is what we would intuitively expect.

Second answer to part a. This solution uses the technique shown in section 6.3.4. From the graph, we know that M looks like a cone with a rounded end. In particular, it looks like a deformation of the plane. So, the key observation is to parameterize M using the graph coordinates x and y. (How do we know to use these coordinates? We have to intuitively know, although we haven't had algebraic topology until the next semester, that this manifold looks like a plane.) Since we are parameterizing using x and y, the tangent space will be spanned by $\partial/\partial x$ and $\partial/\partial y$. A point $\mathbf{q} \in M$ has coordinates

$$\mathbf{q} = \begin{pmatrix} x \\ y \\ \sqrt{x^2 + y^2 + 1} \end{pmatrix}.$$

As in section 6.3.4, we apply $\partial/\partial x$ and $\partial/\partial y$ to the rectangular coordinates of $\mathbf{q}$ to obtain

$$\frac{\partial}{\partial x} = \begin{pmatrix} 1 \\ 0 \\ \frac{x}{\sqrt{x^2+y^2+1}} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ x/z \end{pmatrix}, \quad \frac{\partial}{\partial y} = \begin{pmatrix} 0 \\ 1 \\ \frac{y}{\sqrt{x^2+y^2+1}} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ x/z \end{pmatrix}.$$

We can (and should) do the sanity check as in the first solution. Note that the z in the denominator is non-singular, since all points on this surface have non-zero z (in fact, $z \geq 1$).

* * *

Answer to part b. We have found that, at all points $\mathbf{q}$ of N, the tangent space $TN_{\mathbf{q}}$ has dimension 2, i.e. it is a tangent plane. Thus, TN has dimension 2. Furthermore, the dimension of the tangent space is the same as the dimension of the manifold [xxx write and xref].

10.6.6 Geometry final review problem 6

(a) What is a trivial vector bundle? Give an example of a vector bundle which is not trivial.

(b) Suppose that M is an embedded submanifold of a manifold X. How is the normal bundle of M inside of X defined? What is its dimension as a manifold (in terms of the dimensions of X and M)?

(c) Consider $\mathbb{S}^n \subseteq \mathbb{R}^{n+1}$. Show that the normal bundle is a trivial bundle.

Answer to part (a). xxx move up to main body of paper: I believe the following to be correct: Let E be a k-dimensional bundle over a manifold M. A (global) trivialization is a diffeomorphism $E \cong M \times \mathbb{R}^k$ which is linear on fibers. I believe that, in practice, it suffices to construct smoothly varying, linearly independent sections which are defined globally on M.

xxx include the additional commuting-diagram criterion.

xxx move up to main body: surface of revolution and $\mathbb{S}^1$ are examples of manifolds with trivial tangent bundles. $\mathbb{S}^2$, on the other hand, has a non-trivial tangent bundle. Proof of this requires appealing to the sphere-combing theorem, but we can provide some motivation by noting that spherical coordinates are not defined at the poles, graph coordinates are only defined on hemispheres, stereographic projections are defined everywhere except *one* pole, etc.

Answer to part (b).

Answer to part (c).

10.6.7 Geometry final review problem 7

Compute the flow for the vector field on $\mathbb{S}^2$ given by

$$\mathbf{v}\Big|_{\left(\begin{smallmatrix} x \\ y \\ z \end{smallmatrix}\right)} = \begin{pmatrix} -z \\ 0 \\ x \end{pmatrix}.$$

Answer. We are given

$$\begin{pmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \end{pmatrix} = \begin{pmatrix} -z \\ 0 \\ x \end{pmatrix}.$$

Ideally we would want univariate differential equations. Clearly we already have $\dot{y} = 0$ from which $y = y_0$ for all t. Differentiating twice gives

$$\begin{pmatrix} \ddot{x} \\ \ddot{y} \\ \ddot{z} \end{pmatrix} = \begin{pmatrix} -\dot{z} \\ 0 \\ \dot{x} \end{pmatrix} = \begin{pmatrix} -x \\ 0 \\ -z \end{pmatrix}.$$

The univariate ODEs $\ddot{x} + x = 0$ and $\ddot{z} + z = 0$ have solutions

$$\begin{aligned} x &= a\cos t + b\sin t \\ z &= c\cos t + d\sin t. \end{aligned}$$

Putting in initial conditions gives $x(0) = a$ and $z(0) = c$, so

$$\begin{aligned} x &= x_0\cos t + b\sin t \\ z &= z_0\cos t + d\sin t. \end{aligned}$$

Doing this required only looking at x and z independently. Now, to find b and d, we can use the coupling of x and z from the original equation to obtain

$$\begin{array}{rclclcl} \dot{x} &=& -x_0\sin t + b\cos t &=& -z &=& -z_0\cos t - d\sin t \\ \dot{z} &=& -z_0\sin t + d\cos t &=& x &=& x_0\cos t + b\sin t \end{array}$$

Equating like terms (namely, sin and cos) gives $d = x_0$ and $b = -z_0$, so

$$\begin{aligned} x &= x_0\cos t - z_0\sin t \\ y &= y_0 \\ z &= z_0\cos t + x_0\sin t. \end{aligned}$$

or

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \cos t & 0 & -\sin t \\ 0 & 1 & 0 \\ \sin t & 0 & \cos t \end{pmatrix} \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix}.$$

Note that this motion is rotating about the y axis.

10.6.8 Geometry final review problem 9

(a) Suppose that $M = \mathbb{R}^3$. Explain the sense in which $d : \Omega^k(M) \to \Omega^{k+1}(M)$ (the exterior derivative) is equivalent to the gradient, curl, and divergence.

(b) Suppose that $\mathbf{A}$ is a vector field on $\mathbb{R}^m$. Show that

$$d(\mathbf{A} \lrcorner\, dx_1 \wedge \ldots \wedge dx_m) = \operatorname{div}(\mathbf{A}) dx_1 \wedge \ldots \wedge dx_m.$$

(Hence the exterior derivative on $m-1$ forms in $\mathbb{R}^m$ is equivalent to the divergence.)

Answer to part (a).

Answer to part (b). This is the same as problems 10.4.1 and 10.4.1 part (b), except generalized from 3 to m dimensions. Proceeding as in the $m = 3$ case, we let

$$\mathbf{A} = \begin{pmatrix} A_1 \\ \vdots \\ A_m \end{pmatrix}.$$

Then, again observing that dV of m vectors $\mathbf{v}_1, \ldots, \mathbf{v}_m$ is equal to the area of the parallelepiped spanned by them, we have

$$\mathbf{A} ctt dx_1 \wedge \ldots \wedge dx_m = \begin{vmatrix} A_1 & \cdots & A_m \\ v_{21} & \cdots & v_{2m} \\ \vdots & & \vdots \\ v_{m1} & \cdots & v_{mm} \end{vmatrix} = \sum_{i=1}^{m} A_i dx_1 \wedge \ldots \wedge \widehat{dx_i} \wedge \ldots \wedge \ldots dx_m$$

where the $\widehat{\cdot}$ notation indicates omission. Recall [xxx cite] that

$$dA_i = \sum_{j=1}^{m} \frac{\partial A_i}{\partial x_j} dx_j.$$

Then using the linearity of the derivative,

$$\begin{aligned} d(\mathbf{A} \lrcorner dx_1 \wedge \ldots \wedge dx_m) &= \sum_{i=1}^{m} d\left(A_i dx_1 \wedge \ldots \wedge \widehat{dx_i} \wedge \ldots \wedge \ldots dx_m\right) \\ &= \sum_{i=1}^{m} \left[\sum_{j=1}^{m} \frac{\partial A_i}{\partial x_j} dx_j\right] dx_1 \wedge \ldots \wedge \widehat{dx_i} \wedge \ldots \wedge \ldots dx_m. \end{aligned}$$

Now, $dx_j \wedge dx_j = 0$ so, as in the $m = 3$ case above, these terms are zero except when $i = j$, where the dx_i from the inner sum fills the empty spot in the outer sum, making up the full dV form each time. We now have

$$d(\mathbf{A} \lrcorner dx_1 \wedge \ldots \wedge dx_m) = \left(\sum_{i=1}^{m} \frac{\partial A_i}{\partial x_i}\right) dx_1 \wedge \ldots \wedge dx_m.$$

which is precisely $\operatorname{div}(\mathbf{A}) dV$.

10.6.9 Geometry final review problem 10

Let ω denote the two-form defined on $\mathbb{R}^3$ by

$$\omega = y^2 dx \wedge dz.$$

Let $i : \mathbb{S}^2 \to \mathbb{R}^3$ denote the inclusion.

(a) Calculate $i^*\omega$ in spherical coordinates.

(b) Orient $\mathbb{S}^2$ using the outward-pointing normal vector. Set up an explicit integral expression for

$$\int_{(\mathbb{S}^2)^+} i^*\omega$$

in spherical coordinates.

Answer to part (a). Convert directly to spherical coordinates for $\mathbb{S}^2$ (i.e. use spherical coordinates for $\mathbb{R}^3$

with $\rho = 1$):

$$\begin{aligned}
i^*(\omega) &= (i^*y)^2 d(i^*x) \wedge d(i^*z) \\
&= (i^*y)^2 d(i^*x) \wedge d(i^*z) \\
&= (\sin\theta\sin\phi)^2 d(\cos\theta\sin\phi) \wedge d(\cos\phi) \\
&= \sin^2\theta\sin^2\phi\,(-\sin\theta\sin\phi d\theta + \cos\theta\cos\phi d\phi) \wedge (-\sin\phi d\phi) \\
&= \sin^2\theta\sin^2\phi\,(\sin\theta\sin\phi d\theta - \cos\theta\cos\phi d\phi) \wedge (\sin\phi d\phi) \\
&= \sin^3\theta\sin^4\phi d\theta \wedge d\phi.
\end{aligned}$$

Alternate answer to part (a). We want to write $\omega = y^2\, dx \wedge \mathrm{d}z$ in spherical coordinates on $\mathbb{S}^2$. As in problem 10.5.2, we should be able to write

$$\omega = \omega\left(\frac{\partial}{\partial\theta}, \frac{\partial}{\partial\phi}\right) d\theta \wedge d\phi.$$

We know that a point $\mathbf{q}$ of $\mathbb{S}^2$ has coordinates

$$\mathbf{q} = \begin{pmatrix} \cos\theta\sin\phi \\ \sin\theta\sin\phi \\ \cos\phi \end{pmatrix}$$

so

$$\frac{\partial}{\partial\theta} = \begin{pmatrix} -\sin\theta\sin\phi \\ \cos\theta\sin\phi \\ 0 \end{pmatrix}, \quad \frac{\partial}{\partial\phi} = \begin{pmatrix} \cos\theta\cos\phi \\ \sin\theta\cos\phi \\ -\sin\phi \end{pmatrix}.$$

Then, recalling that dx and dz are the coordinate-extractor functions for x and z, respectively, we have

$$\begin{aligned}
\omega\left(\frac{\partial}{\partial\theta}, \frac{\partial}{\partial\phi}\right) &= y^2 dx \wedge dz\left(\frac{\partial}{\partial\theta}, \frac{\partial}{\partial\phi}\right) \\
&= (\sin\theta\sin\phi)^2\left[dx\left(\frac{\partial}{\partial\theta}\right) dz\left(\frac{\partial}{\partial\phi}\right) - dz\left(\frac{\partial}{\partial\theta}\right) dx\left(\frac{\partial}{\partial\phi}\right)\right] \\
&= \sin^2\theta\sin^2\phi\left[\sin\theta\sin^2\phi - 0\right] \\
&= \sin^3\theta\sin^4\phi
\end{aligned}$$

which is in agreement with the previous result.

Answer to part (b). It turns out that what Pickrell means by "explicit integral expression" is a Riemann integral of the form

$$\int_{\phi=0}^{\pi/2}\int_{\theta=0}^{2\pi} \sin^3\theta\sin^4\phi d\theta\, d\phi.$$

It remains to get the sign right. We have $\partial/\partial\theta$ going latitudinally and counterclockwise (i.e. west to east) and $\partial/\partial\phi$ going longitudinally toward the north pole. Thus to get an outward-pointing normal, using the right-hand rule, we would need to order our basis for $T\mathbb{S}^2$ as $\{\partial/\partial\theta, \partial/\partial\phi\}$. So, no sign change is necessary.

10.6.10 Geometry final review problem 11

Let $\partial/\partial\rho$ denote the outward-pointing vector field in $\mathbb{R}^n \setminus \{0\}$.

(a) Express this vector field in Euclidean coordinates.

(b) Compute $w = (\partial/\partial\rho) \lrcorner\, dx_1 \wedge \ldots \wedge dx_n$ in Euclidean coordinates.

(c) Explain why the $(n-1)$-form ω restricts to the standard volume form on $\mathbb{S}^{n-1}$.

Answer to part (a). The vector pointing outward from the origin at a point $\mathbf{q}$ of $\mathbb{R}^n$ is simply $\mathbf{q}$ itself. Thus we can get an outward-pointing normal vector by normalizing $\mathbf{q}$:

$$\partial/\partial\rho = \hat{\mathbf{q}} = \frac{\mathbf{q}}{\|\mathbf{q}\|} = \frac{(x_1, \ldots, x_n)}{\sqrt{x_1^2 + \ldots + x_n^2}}.$$

Here we think of x_i as the coordinates of $\partial/\partial\rho$ with respect to the basis $\{\partial/\partial x_1, \ldots, \partial/\partial x_n\}$ for $T\mathbb{R}^n$, which are the same n real numbers as the coordinates of $\mathbf{q}$ with respect to the basis $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ for $\mathbb{R}^n$.

Answer to part (b). If I am understanding this correctly, I can use problem 9b. Recall that the divergence is defined in $\mathbb{R}^n$, not just $\mathbb{R}^3$: just dot the gradient with the vector in question. The ith component of $\partial/\partial\rho$ is

$$\frac{x_i}{(x_1^2 + \ldots + x_n^2)^{1/2}}.$$

Then $\partial/\partial x_i$ of that is

$$\begin{aligned}
\frac{\partial}{\partial x_i}\left(\frac{x_i}{(x_1^2 + \ldots + x_n^2)^{1/2}}\right) &= \frac{1}{(x_1^2 + \ldots + x_n^2)^{1/2}} + \frac{x_i \cdot 2x_i \cdot (\frac{-1}{2})}{(x_1^2 + \ldots + x_n^2)^{3/2}} \\
&= \frac{1}{(x_1^2 + \ldots + x_n^2)^{1/2}} - \frac{x_i^2}{(x_1^2 + \ldots + x_n^2)^{3/2}} \\
&= \frac{x_1^2 + \ldots + x_{i-1}^2 + x_{i+1}^2 + \ldots + x_n^2}{(x_1^2 + \ldots + x_n^2)^{3/2}}
\end{aligned}$$

Answer to part (c). This is clear geometrically: think of the area form on $\mathbb{S}^2$, which along with $d\rho$ gives the volume form on $\mathbb{R}^3$. To be more formal, though, ... [xxx type up Pickrell's remarks from the review session.]

10.6.11 Geometry final review problem 12

Let M be a manifold.

(a) Define $H^*_{dR}(M)$.

(b) Attempt to define a product (the *cup product*) on $H^*_{dR}(M)$ by

$$[\omega] \smile [\eta] = [\omega \wedge \eta].$$

Show that this is well-defined, and hence cohomology is an associative algebra. (An example: Next semester we will see that in the case of the torus $M = \mathbb{S}^1 \times \mathbb{S}^1$, $H^*(M)$ is the exterior algebra generated by $[d\theta_1]$ and $[d\theta_2]$.

Answer for part (a). For any given r, the rth de Rham cohomology of M, written $H^r_{dR}(M)$, is the quotient space of closed r-forms mod exact r-forms. (How can one remember this? There are two ways: (1) c comes alphabetically before e; (2) Since $d^2 = 0$, an exact form is closed. That is, if $\omega = d\eta$, then $d\omega = d^2\eta = 0$. So, taking exact forms mod closed forms wouldn't make sense.)

Now, the r forms, for each r, form a real vector space. The rth de Rham cohomology is a quotient of vector spaces, but we sometimes forget (in the sense of section 4.1.8) about scalar multiplication and think of this as a quotient of abelian groups. The set of all r forms, for varying r, form a tensor algebra. When we form the quotient, we lose the multiplication, but this can be restored by considering the cup product (see below).

Thus $H^*_{dR}(M)$ is the union of $H^r_{dR}(M)$ for all r.

Answer for part (b). To show that the cup product is well-defined, we need to show that it is independent of choice of coset representative. By the symmetry of the problem it suffices to show that the cup product is independent of the choice of ω; showing that it is independent of the choice of η is essentially the same problem.

Suppose

$$\omega_1 \sim \omega_2, \qquad \text{i.e.} \qquad [\omega_1] = [\omega_2]$$

and fix η. We need to show that

$$[\omega_1 \wedge \eta] = [\omega_2 \wedge \eta].$$

Since H^* is the quotient of closed forms by exact forms, this means (see also section 6.6.1) that $\omega_1 - \omega_2$ is exact. That is, there is a form α, of order one less than the orders of the ω_i's, such that

$$\omega_1 - \omega_2 = d\alpha.$$

We need to show that

$$[\omega_1 \wedge \eta] = [\omega_2 \wedge \eta].$$

For this to be true, we would need

$$\omega_1 \wedge \eta - \omega_2 \wedge \eta = d\beta$$

for some β. But recall that the set of forms is an algebra, in particular a ring, with the wedge operator as its multiplication. So, we can undistribute:

$$(\omega_1 - \omega_2) \wedge \eta = d\alpha \wedge \eta = d\beta.$$

So our remaining task is to find such a form β. One might suggest $\beta = \alpha \wedge \eta$. Then

$$d\beta = d\alpha \wedge \eta + (-1)^k \alpha \wedge d\eta$$

where $k = \text{ord}(\alpha)$. All we need is $\alpha \wedge d\eta$ is 0. But $d\eta = 0$ since η is closed: recall that homology is *closed* forms mod exact ones.

10.6.12 Geometry final review problem 13

(a) State and prove Stokes' theorem for the standard cube in $\mathbb{R}^n$. (On the exam I might ask you to do this for $n = 2$, which amounts to Green's theorem.)

(b) Explain why the classical divergence and Stokes' theorems are special cases of our Stokes' theorem.

Answer to part (a). First, I will do $n = 2$. (The solution for the $n = 2$ case is due largely to Dr. Pickrell and Yuliya Gorlina.)

We need to prove

$$\int_c d\omega = \int_{\partial c} \omega$$

where c is the unit square I^2. Since $n = 2$, $n - 1 = 1$, ω is a 1-form and hence [xxx write and xref to basis statement] is of the form

$$\omega = f(x, y)dx + g(x, y)dy.$$

To find the left-hand side of the equation to be proved, we need $d\omega$. This is

$$\begin{aligned} d(\omega) &= d(f\,dx + g\,dy) = d(f\,dx) + d(g\,dy) \\ &= \left(\frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy\right) \wedge dx + \left(\frac{\partial g}{\partial x}dx + \frac{\partial g}{\partial y}dy\right) \wedge dy \\ &= \frac{\partial f}{\partial y}dy \wedge dx + \frac{\partial g}{\partial x}dx \wedge dy \\ &= \left(\frac{\partial g}{\partial x} - \frac{\partial f}{\partial y}\right) dx \wedge dy. \end{aligned}$$

Now we can integrate this. At this point we switch [xxx write and xref] from the chains-and-forms notation to the Riemann integral. Once the double integral is in that form, we can switch the order of integration, which is justifiable by **Fubini's theorem**. We have

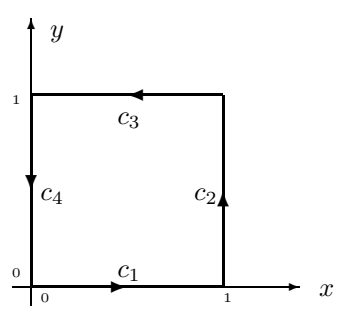
$$\begin{aligned} \int_c d\omega &= \int_{y=0}^{y=1} \int_{x=0}^{x=1} \left(\frac{\partial g}{\partial x} - \frac{\partial f}{\partial y}\right) dx\,dy \\ &= \int_{y=0}^{y=1} \left[\int_{x=0}^{x=1} \frac{\partial g}{\partial x}dx\right] dy - \int_{x=0}^{x=1} \left[\int_{y=0}^{y=1} \frac{\partial f}{\partial y}dy\right] dx. \end{aligned}$$

By the fundamental theorem of calculus (using it in reverse), this is

$$\int_c d\omega = \int_{y=0}^{y=1} [g(1, y) - g(0, y)]\,dy - \int_{x=0}^{x=1} [f(x, 1) - f(x, 0)]\,dx.$$

Since I'm not sure how to continue with this, it seems like a good time to leave it as is, and work on the right-hand side for a while.

For the right-hand side we have $\int_{\partial c} \omega$ where ∂c is the counterclockwise path around the unit square. I have labeled the four pieces of the path path c_1, c_2, c_3, and c_4:



Again switching to the Riemann integral, we have

$$\begin{aligned}
\int_{\partial c} \omega &= \int_{c_1} (f(x,y)dx + g(x,y)dy) + \int_{c_2} (f(x,y)dx + g(x,y)dy) \\
&+ \int_{c_3} (f(x,y)dx + g(x,y)dy) + \int_{c_4} (f(x,y)dx + g(x,y)dy) \\
&= \int_{x=0}^{x=1} (f(x,0)dx + g(x,0)dy) + \int_{y=0}^{y=1} (f(1,y)dx + g(1,y)dy) \\
&+ \int_{x=1}^{x=0} (f(x,1)dx + g(x,1)dy) + \int_{y=1}^{y=0} (f(0,y)dx + g(0,y)dy)
\end{aligned}$$

Half of these vanish as discussed in remark 2.29. Then

$$\begin{aligned}
\int_{\partial c} \omega &= \int_{x=0}^{x=1} f(x,0)dx + \int_{y=0}^{y=1} g(1,y)dy + \int_{x=1}^{x=0} f(x,1)dx + \int_{y=1}^{y=0} g(0,y)dy \\
&= \int_{x=0}^{x=1} f(x,0)dx + \int_{y=0}^{y=1} g(1,y)dy - \int_{x=0}^{x=1} f(x,1)dx - \int_{y=0}^{y=1} g(0,y)dy) \\
&= \int_{x=0}^{x=1} (f(x,0) - f(x,1))dx + \int_{y=0}^{y=1} (g(1,y) - g(0,y))dy \\
&= \int_{y=0}^{y=1} [g(1,y) - g(0,y)]\, dy - \int_{x=0}^{x=1} [f(x,1) - f(x,0)]\, dx
\end{aligned}$$

which is what we wanted to show.

10.6.13 Geometry final review problem 14

(a) Let M be an oriented manifold. Prove that there exists a volume form on M which is positive at all points of M.

(b) Suppose that M is an oriented compact n-dimensional manifold without boundary. Explain why the map

$$H^n_{dR}(M, \mathbb{R}) \to \mathbb{R} : [\omega] \mapsto \int_M \omega$$

is well-defined. Use (a) to explain why, in this case, $H^n_{dR}(M, \mathbb{R})$ is not zero.

(c) It turns out that the map in part (b) is an isomorphism. This has the following consequence: if ω is an n-form on M, then there is an $n-1$ form η on M such that $d\eta = \omega$ iff $\int_M \omega = 0$. Explain why this is true for $M = \mathbb{S}^1$.

10.6.14 Geometry final review problem 15

Calcluate the area of $\mathbb{S}^2$ (the 2-sphere with radius r), for the Euclidean area form. Calculate the volume of the ball of radius r in $\mathbb{R}^3$, for the standard volume form. How are these two quantities abstractly related? (Hint: consider a derivative with respect to r of something, and use units. This is a first-semester calculus question related to the second fundamental theorem of calculus.)

Answer. Let B^2 be the filled unit ball in $\mathbb{R}^3$ (i.e. $\mathbb{S}^2 = \partial B^2$). Recall that

$$A = \int_{r\mathbb{S}^2} dA = 4\pi r^2$$

and

$$V = \int_{rB^2} dV = \frac{4}{3}\pi r^3.$$

This gives us $V/A = r/3$, from which $V = Ar/3$. I will use this fact to check my work.

To use the second fundamental theorem of calculus (theorem 2.36), though, we need to write something of the form

$$F(r) = \int_0^r f(t)dt$$

which by the SFTC will tell us

$$F'(r) = f(r).$$

So, think of the ball as the union of nested shells, and take the limit as the number of shells goes to infinity. The volume of a thin shell from radius a to radius b may be approximated by

$$4\pi a^2(b-a) = 4\pi a^2 h$$

where $h = b - a$. Summing up n such shells, from radius r_i to r_{i+1}, with $\Delta r = r_{i+1} - r_i$, gives

$$V_n = \sum_{i=0}^{n-1} 4\pi r_i^2 \Delta r.$$

Taking the limit of this finite sum gives

$$V = \int_0^r 4\pi t^2 dt = \frac{4\pi r^3}{3}.$$

Since $A(t) = 4\pi t^2$, we can also think of this integral as

$$V = \int_0^r A(t)dt.$$

The SFTC says

$$V'(r) = A(r),$$

which we can verify by writing

$$\frac{dV}{dr} = \frac{d}{dr}\left(\frac{4}{3}\pi r^3\right) = 4\pi r^2 = A(r).$$

Alternative answer (Pickrell). Start with

$$V(r) = \frac{4\pi r^3}{3}.$$

Use the definition of derivative to obtain

$$\begin{aligned} V'(r) &= \lim_{h\to 0} \frac{V(r+h) - V(r)}{h} = \lim_{h\to 0} \frac{4\pi}{3} \frac{(r+h)^3 - r^3}{h} \\ &= \left(\frac{4\pi}{3}\right) \lim_{h\to 0} \left(\frac{r^3 + 3r^2h + 3rh^2 + h^3 - r^3}{h}\right) = \frac{4\pi}{3} 3r^2 = 4\pi r^2. \end{aligned}$$

Remark. I'm a bit puzzled by both of these solutions. Pickrell asked for how the volume and area are "abstractly related", but both solutions are quite concrete.

10.6.15 Geometry final review problem 16

(a) State a change-of-variables theorem for integration.

(b) Explain why it is true for a constant integrand, the standard cube as domain, and a linear change of variables.

Answer. For part (a), see [**Lee2**], p. 593. [xxx move to main body of paper and xref.] Let A, B be compact domains of integration in $\mathbb{R}^m$. Let $f : B \to \mathbb{R}$ be continuous, and let $G : A \to B$ be a smooth bijection with smooth inverse. Then

$$\int_{G(A)} f dV = \int_A (f \circ G)|\det(DG)|dV.$$

xxx include a nice pullback diagram here: $G : A \to B$, $f : B \to \mathbb{R}$, $f \circ G : A \to \mathbb{R}$.

For part (b), I will first do a particular example. The emphasis here is on having a quick computation we can do in our heads, or on scratch paper, to make sure we have the det on the correct side of the equation.

Let x, y be coordinates for $\mathbb{R}^2$. Let I be the unit square in $\mathbb{R}^2$ and let u, v be given by

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} 2x \\ 3y \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = G \begin{pmatrix} x \\ y \end{pmatrix}.$$

Note that G is already linear, so $DG = G$, and $\det(G) = 6$. Integrate the constant function 1. From freshman calculus we have

$$\int_I dA = \int_{y=0}^{y=1} \int_{x=0}^{x=1} dx\, dy = 1.$$

Using the change of variables $u = 2x$ and $v = 3y$, we have

$$\int_{v=0}^{v=3} \int_{u=0}^{u=2} du\, dv = 6.$$

So, we may safely remember that the $\det(G)$ goes on the left-hand side, and we have

$$\int_I \det(G) dA = \int_{G(I)} dA.$$

[xxx replace $v = 3x$ with $v = -3x$ and check that the signs are correct.]

Now for part (b) per se. We are integrating a constant function $c(\mathbf{x})$ over the standard cube I of $\mathbb{R}^m$, with linear change of variables $G : \mathbb{R}^m \to \mathbb{R}^m$. Again, G is already linear so $DG = G$. The function c is from $G(I)$ to $\mathbb{R}$, so to get a map from I to $\mathbb{R}$ we need to pull c back as $c \circ G$. This helps us get the c and the $c \circ G$ in the right places. Then the left-hand side is

$$\int_I (c \circ G) \det(G) dV$$

while the right-hand side is

$$\int_{G(I)} c dV.$$

The composition $c \circ G$ is again a constant function. (I.e. c is not the identity function.) Thus the constants pull out of both sides, and the constant $\det(G)$ pulls out of the left-hand side. We have

$$\det(G) \int_I dV, \qquad \int_{G(I)} dV.$$

All we need to do to prove these two integrals are equal is recognize that

$$\int_{G(I)} dV = \det(G).$$

10.7 Geometry final

10.7.1 Geometry final problem 1

(a) Show that

$$M = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 : x^2 + xy + y^2 + xz + z^2 = 1 \right\}$$

is an embedded submanifold of $\mathbb{R}^3$.

(b) The point $\mathbf{q} = (1, -1, -1)$ is in M. Find a parameterization for M in a neighborhood of $\mathbf{q}$.

Answer to part (a). When we hear *embedded submanifold*, we should think *regular value theorem.* In particular, let $f : \mathbb{R}^3 \to \mathbb{R}$ be given by $f(x, y, z) = x^2 + xy + y^2 + xz + z^2$. Then M is the level set of f and 1, i.e. $M = f^{-1}(1)$. By the regular value theorem (theorem 6.17), if 1 is a regular value of f, then $f^{-1}(1)$ is either empty, or is an embedded submanifold of $\mathbb{R}^3$. Since part (b) gives us a point on M, $f^{-1}(1)$ is not empty. So, it remains to show that 1 is a regular value of f. This is the case if all preimages of 1, i.e. all points $\mathbf{p}$ of M, are regular points of f. This in turn is the case if $Df|_{\mathbf{p}}$ is surjective for each p. Since $f : \mathbb{R}^3 \to \mathbb{R}$, $Df|_{\mathbf{p}}$ has full rank when it is non-zero. We compute

$$Df = \begin{pmatrix} 2x + y + z, & x + 2y, & x + 2z \end{pmatrix}.$$

For all of these to simultaneously vanish at a point $\mathbf{p} = (x, y, z)$, we would have

$$x = -2y, x = -2z, 2x + y + z = 0.$$

The first two force $y = z$, so the last becomes $2x + 2y = 0$. This combined with $x = -2y$ gives $x = 0$, and so we have $\mathbf{p} = (0, 0, 0)$. But this point is not on M. Since any critical point lies off M, all points p on M are regular.

Answer to part (b). When we hear *parameterization in a neighborhood*, we should think *implicit function theorem.* We know by the regular value theorem that M is two-dimensional, since it is defined by one equation in three unknowns. As in example 3.9, we can use the theorem to find which variable to solve for in terms of the other two. We compute Df as above, and substitute $\mathbf{q} = (1, -1, -1)$:

$$Df|_{\mathbf{q}} = \begin{pmatrix} 0, & -1, & -1 \end{pmatrix}.$$

So, we can solve for y or z. Let's use z. We have a quadratic and a linear term in z in the equation

$$x^2 + xy + y^2 + xz + z^2 = 1,$$

suggesting that we will need to complete the square. Solving for the z terms gives

$$\begin{aligned} z^2 + xz &= 1 - x^2 - xy - y^2 \\ z^2 + zx + \frac{x^2}{4} &= 1 - x^2 - xy - y^2 + \frac{x^2}{4} \\ \left(z + \frac{x}{2}\right)^2 &= 1 - xy - y^2 - \frac{3x^2}{4} \\ z &= -\frac{x}{2} - \sqrt{1 - xy - y^2 - \frac{3x^2}{4}} \end{aligned}$$

where we select the negative square root so that z will be -1 at $x = 1$, $y = -1$.

The problem asked for a parameterization (see definition 6.3). So, to finish up, we write

$$\begin{pmatrix} x \\ y \end{pmatrix} \longmapsto \begin{pmatrix} x \\ y \\ -\frac{x}{2} - \sqrt{1 - xy - y^2 - \frac{3x^2}{4}} \end{pmatrix}.$$

10.7.2 Geometry final problem 2

Consider the function

$$f : \mathbb{S}^2 \to \mathbb{R} : \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto xy.$$

(a) Find the critical points and critical values for this function.

(b) Consider the point $\mathbf{q} = (1, 0, 0)$. Find the direction at this point in which f increases most rapidly.

Here I give two solutions, one using projections as discussed in section 1.3.2, and the other using Lagrange multipliers as discussed in section 2.7.

Answer to part (a) using projections. For part (a), let $\mathbf{q} = (x, y, z)$ vary over $\mathbb{S}^2$. On $\mathbb{R}^3$, we have

$$Df_{\mathbf{q}} = (y, x, 0)_{\mathbf{q}}.$$

We need to restrict this to $\mathbb{S}^2$. Proceeding as in section 1.3.2, we recognize that $\hat{\mathbf{n}} = \mathbf{q}$ is normal to $\mathbb{S}^2$. So, we can **decompose** $Df_{\mathbf{q}}$ into components parallel and perpendicular to $\hat{\mathbf{n}}$. Writing $\mathbf{u}$ for $Df_{\mathbf{q}}$, and recognizing that $\hat{\mathbf{n}}$ already has unit length to we don't need to scale by $\hat{\mathbf{n}} \cdot \hat{\mathbf{n}}$, the parallel part (which we need first) is

$$\mathbf{u}_{\parallel} = (\mathbf{u} \cdot \hat{\mathbf{n}})\hat{\mathbf{n}} = \left[\begin{pmatrix} y \\ x \\ 0 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} \right] \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2x^2y \\ 2xy^2 \\ 2xyz \end{pmatrix}.$$

Then

$$Df|_{T\mathbb{S}^2} = \mathbf{u}_{\perp} = \mathbf{u} - \mathbf{u}_{\parallel} = \begin{pmatrix} y - 2x^2y \\ x - 2xy^2 \\ -2xyz \end{pmatrix}.$$

The critical points of f are those for which $Df|_{T\mathbb{S}^2}$ has rank less than 1, i.e. zero. For this to happen, all three entries must be zero. So, the following must be simultaneously true:

$$\begin{cases} y(1 - 2x^2) &= 0, \\ x(1 - 2y^2) &= 0, \\ -2xyz &= 0. \end{cases}$$

That is,

$$\begin{array}{lllll} [y = 0 & \text{or} & x = \pm 1/\sqrt{2}\,], & \text{and} & \\ [x = 0 & \text{or} & y = \pm 1/\sqrt{2}\,], & \text{and} & \\ [x = 0 & \text{or} & y = 0 & \text{or} & z = 0\,]. \end{array}$$

Enumerating the possibilities, we obtain the 6 critical points

$$\begin{pmatrix} 0 \\ 0 \\ \pm 1 \end{pmatrix}, \quad \begin{pmatrix} \pm 1/\sqrt{2} \\ \pm 1/\sqrt{2} \\ 0 \end{pmatrix}.$$

The critical values of f are the images of these 6 points, which are

$$0, \quad \pm 1/2.$$

Answer to part (a) using Lagrange multipliers. The manifold $\mathbb{S}^2$ is the zero set of $g(x, y, z) = x^2+y^2+z^2-1$; the function to be maximized is $f(x, y, z) = xy$. We put

$$\begin{aligned} \nabla f &= \lambda \nabla g \\ \begin{pmatrix} y \\ x \\ 0 \end{pmatrix} &= \begin{pmatrix} 2\lambda x \\ 2\lambda y \\ 2\lambda z \end{pmatrix}. \end{aligned}$$

If $\lambda = 0$, then $x = y = 0$ and $z = \pm 1$. If $\lambda \neq 0$, then $z = 0$ and we can solve for λ in the first two equations to find x and y. Note that if $x = 0$ then $y = 0$ and we have $x = y = z = 0$ which is not a point on $\mathbb{S}^2$. So, we may divide by x and y. We obtain

$$\begin{aligned} \lambda &= y/2x \\ \lambda &= x/2y \\ 2x^2 &= 2y^2. \end{aligned}$$

Since $z = 0$ and $x^2 + y^2 = 1$, we obtain for $\lambda \neq 0$ the same four points as in the previous solution. These four, along with the north and south poles for $\lambda = 0$, give us the critical points

$$\begin{pmatrix} 0 \\ 0 \\ \pm 1 \end{pmatrix}, \quad \begin{pmatrix} \pm 1/\sqrt{2} \\ \pm 1/\sqrt{2} \\ 0 \end{pmatrix}.$$

as before. Of course, the critical values are the same as well.

Answer to part (b). The direction of greatest change of f on $\mathbb{S}^2$ is simply the gradient of f restricted to $\mathbb{S}^2$, which is what we just computed. So all we need to do is evaluate $Df_{T\mathbb{S}^2}$ at $\mathbf{q} = (1, 0, 0)$. We obtain

$$Df|_{T\mathbb{S}^2} = \begin{pmatrix} y - 2x^2y \\ x - 2xy^2 \\ -2xyz \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}.$$

10.7.3 Geometry final problem 5

Let

$$\eta = \frac{-y\,dx + x\,dy}{x^2 + y^2}$$

on $\mathbb{R}^2 \setminus \{0\}$.

(a) Show that η is closed.

(b) Calculate $\int_{\mathbb{S}^1} \eta$, where $\mathbb{S}^1$ has the counterclockwise orientation.

(c) Why does your answer to (b) imply that there does not exist a function f on $\mathbb{R}^2 \setminus \{0\}$ such that $df = \eta$?

Answer to part (a).

Answer to part (b).

Answer to part (c).

10.7.4 Geometry final problem 6

Let X denote the upper hemisphere of $\mathbb{S}^2$ with the orientation induced by the upward-pointing normal vector. Compute the integral

$$\int_X z\,dx \wedge dy.$$

Answer, due in part to Tommy Occhipinti. My (the author's) natural inclination is to treat X as the surface, with perimeter ∂X being the unit circle in the x, y plane. Then Stokes' theorem would give us

$$\int_X d\omega = \int_{\partial X} \omega.$$

In this scenario, the given form $z\,dx \wedge dy$ is $d\omega$, and we need to find out what ω is. But will this search be futile? That is, is $z\,dz \wedge dy$ exact? Recall that exact forms are closed, since $d^2 = 0$: if $d\omega = z\,dx \wedge dy$ then $d^2\omega = 0$. But this is

$$d(z\,dx \wedge dy) = dz \wedge dz \wedge dy = dx \wedge dy \wedge dz = dV,$$

which is not identically zero. So, $z\,dx \wedge dy$ is not closed, and therefore is not exact.

Since I want to invoke Stokes' theorem, I can't have X being the surface and the perimeter being the boundary. Instead, I must think of X as being the boundary. This requires a little leap of insight: we have to also include the closed disk on the x, y plane, the floor of the dome if you will, in the boundary. (One hopes this will not affect the value of the integral, or at least that the contribution will easily be taken back out. However, the integral over the floor of the dome of $z\,dx \wedge dy$ is zero, since $z = 0$ on the floor.) Here, the volume to be integrated is the upper half of the filled sphere: call it H. If

$$\omega = z\,dx \wedge dy,$$

then (as above)

$$d\omega = dV.$$

Then this problem is easy, since

$$\int_X \omega = \int_H d\omega = \int_H dV.$$

Now, the integral of the volume form over the whole filled unit sphere is $4\pi/3$, so half that is $2\pi/3$.

10.7.5 Geometry final problem 9

Suppose that $z = g(y)$ is a smooth positive function, defined for $y \in [a, b]$. Consider the surface M of revolution with boundary, obtained by revolving the graph of g around the y axis in $\mathbb{R}^3$. Orient M using the outward-pointing normal vector $\hat{\mathbf{n}}$, and consider the coordinates θ, y for M. [xxx orientation of θ. or xref to prev sec.]

(a) Briefly explain why the oriented area form of M is given by the expression

$$dA = g(y)\sqrt{1 + g'(y)^2}\, d\theta \wedge dy.$$

(b) Suppose that $G(y)$ is an antiderivative of $g(y)\sqrt{1 + g'(y)^2}$ for $y \in [a, b]$. Show that

$$\eta = G(y)\, d\theta$$

is a well-defined one-form on M, and that $d\eta = dA$.

[xxx rmk and xref that not everything starting with a d is exact.]

(c) Use part (b) and Stokes' theorem to compute the area of M in terms of G. (Remark: this just recovers a standard formula for the area from first-year calculus.)

10.7.6 Geometry final problem 10

Determine whether the following statements are true or false. Briefly explain your answers.

(a) The form $\omega_1 = y^2\, dx \wedge dz$ is exact in $\mathbb{R}^3$, by the Poincaré Lemma, because $\mathbb{R}^3$ is contractible.

(b) If ω is exact in $\mathbb{R}^2 \setminus \{0\}$, then for any closed oriented curve c in $\mathbb{R}^2 \setminus \{0\}$, $\int_c \omega = 0$.

(c) On $\mathbb{S}^2$, $L_{\partial/\partial\theta} dA = 0$.

(d) If X is a vector field on M and $f : M \to N$ is a map, then we can use f to push X forward to a vector field on N.

10.8 Topology homework 1

10.9 Topology exam 1 review

10.9.1 Topology exam 1 review problem 5

For each real number α, define a group action

$$\mathbb{Z} \times \mathbb{S}^1 \longrightarrow \mathbb{S}^1 : (n, e^{2\pi i\theta}) \mapsto e^{2\pi i(\theta+n\alpha)}.$$

(a) Determine the values of α for which this is a free action, i.e. the stabilizers are all trivial.

(b) Fix a value of α as in (a) such that $\mathbb{Z}$ is acting freely. Show that $\mathbb{Z}$ is not acting properly discontinuously.

Answer to part (a). Either α is rational, or it is not. First consider the former case. Put $\alpha = a/b$ in lowest terms. Then for $n = b$, we have

$$(b, e^{2\pi i\theta}) \mapsto e^{2\pi i(\theta+b(a/b))} = e^{2\pi i(\theta+a)} = e^{2\pi i\theta}$$

since a is an integer. Thus the stabilizer of α is non-trivial. (Note in particular that for $\alpha = 0$, the stabilizer is all of $\mathbb{Z}$.)

Now consider the case when α is irrational. For

$$e^{2\pi i\theta} = e^{2\pi i(\theta+n\alpha)}$$

requires

$$e^{2\pi i(\theta+n\alpha)}e^{-2\pi i\theta} = 0 = e^{2\pi i0} = e^{2\pi in\alpha}$$

which puts $n\alpha \equiv 0 \pmod 1$, i.e. $n\alpha$ is an integer. This of course happens when $n = 0$. Now suppose (seeking a contradiction) that there is some integer m such that $n\alpha = m$ and $n \neq 0$. Then we may divide to obtain $\alpha = n/m$. This shows that an irrational number is equal to a rational number, which is the desired contradiction.

Thus we have shown that the action has trivial stabilizer if and only if α is irrational.

Answer to part (b).

10.10 Topology final

11 About the references

As mentioned in the preface, this is not a standalone work. My focus is on providing intuition and examples that have been short-changed elsewhere; I offer little rigor in this paper since there is no lack of that in the literature — as described below. The following annotations are based on my experiences and the experiences of my fellow graduate students. This is an admittedly small sample.

- Calculus:
  - [**Anton**] is a standard approach to calculus.
  - [**HHGM**] is a reform approach to calculus; the present work is in part inspired by this text.
  - [**Hildebrand**] is advanced calculus.
- Linear algebra:
  - [**FIS**] A senior undergraduate text. The determinant is done entirely without exterior algebra.
- Abstract algebra:
  - [**Herstein**]: A senior undergraduate text.
  - [**DF**], [**Grove**], [**Hungerford**], [**Lang**]: Graduate texts. In particular, you can find information here about sequences, homology, and cohomology.
- Analysis:
  - [**Rudin**]. Contains some information about differential forms.
- Physics approach to geometry/topology:
  - [**Abr**], [**Frankel**]: I find the latter more readable; also, the former is more rigorous.
  - [**Pen**]: Really a pop-science book, but with plenty of mathematical content.
- Differential geometry:
  - [**Spivak1**] is a slim, senior undergraduate text which manages to include much of the content of the geometry course. It includes chains, forms, and duality, ending in a proof of generalized Stokes.
  - [**Boothby**] and [**Lee2**] are textbooks for a geometry/topology course. I have received positive reviews of the former and mixed reviews of the latter. The latter is fussy with proofs, which some readers find obscures the big picture. [**Lee2**] is intended to be preceded by [**Lee1**] (below).
  - I have received very few if any positive student reviews of [**Conlon**] and [**Spivak2**]. Fans of the latter tend to be faculty members.
  - [**Lee3**] is a follow-on to a geometry/topology course.
- Algebraic topology:
  - [**Hatcher**] and [**Massey**]: standalone sources for algebraic topology.
  - [**Lee1**] is intended to precede [**Lee2**].
  - All three are textbooks for a graduate geometry/topology course.
- Classical differential geometry:

  - [**Guggenheimer**]: Focuses on concepts such as curvature and torsion which, for reasons I don't understand, we no longer consider in a first-year geometry course.

- Math history:
  - [**BMA**]: history of math from ancient times. For this reason, information on the development of calculus is limited.
  - [**Dunham**]: initial development and rigorization of single-variable calculus.
  - [**Crowe**]: the advent of the vectorial system.
  - I am still looking for a good reference on the historical development of differential forms.

- Works similar to this one:
  - [**Lamb**]: succinct notes for qualifier preparation.
  - [**Bachman**]: a work very similar to the present one, although less ambitious. Intended for an undergraduate audience (immediately post-calculus) although there is plenty of information for graduate students. I have discovered this work only recently; its geometrical approach to forms is very similiar to mine.

- Encyclopediae:
  - [**CRC**], [**PDM**], [**PM**], and of course Wikipedia.

12 Typesetting and computation

The soul cannot think without a picture. — Aristotle (384-322 B.C.).

This section presents software-related issues related to learning differential geometry and the preparation of this paper. Such topics may seem ephemeral: software changes quickly. However, central concepts and algorithms, happily, do not.

12.1 Files and tools

The source for this paper is (as of this writing) located at

`http://math.arizona.edu/~kerl/doc/prolrev/`.

The top-level file is `prolrev.tex`; most of the contents are in `body.tex`. Figures and artwork are in the `figures/` subdirectory. The index was prepared using my Perl script `kmkidx`: see

`http://math.arizona.edu/~kerl/index.php?v=software`

for more information. PostScript was created using `dvips`; The PDF was created using `ps2pdf`. The `psfrag` package was used to include LaTeX math symbols in the figures. (Thanks to David Bachman for the tip!)

12.2 Figures and artwork

The figures in this document were created using Inkscape and Matlab. The latter is discussed in more detail in section 12.3.

The cover art suggests that we can approach geometry/topology using the familiar circle as a starting point, as discussed in section [write and xref].

The back art (page 209) depicts the following:

- The large and imposing edifice of geometry-topology is constructed of simple building blocks.
- The black flag is the flag of anarchy: this paper was written by a graduate student, not a faculty member. This paper emphasizes issues of interest and/or difficulty *for the learner*.
- The flag is flapping and warped by the breeze — the same breeze which turns Green's theorem into classical Stokes.
- The flag is a 2-dimensional manifold with boundary, with rank-one fundamental group.
- Although the edifice may appear well-armed and impenetrable, the central feature is an entryway, through which we are invited.
- The stones of the archway are reminiscent of a chain of singular 3-cubes.
- Inscribed on the keystone is the generalized Stokes theorem, in pairing notation.

12.3 Matlab

xxx paste in some examples here. Include gradient, surface, contour, and quiver.

How-to's.

`http://math.arizona.edu/~kerl/doc/prolrev/figures`

list.

See [**Kerl**].

13 Under construction and to be merged

13.1 temp

xxx rm2r:

Definition 13.1. Let f be a vector-to-scalar function and let $\mathbf{u} = (u, v, w) \in \mathbb{R}^3$. The **directional derivative** of f in the direction of $\mathbf{u}$ is the dot product

$$D_{\mathbf{u}}(f) = \mathbf{u} \cdot \nabla f = u\frac{\partial f}{\partial x} + v\frac{\partial f}{\partial y} + w\frac{\partial f}{\partial z}.$$

The directional derivative $\mathbf{u} \cdot \nabla(f)$ is the **rate of change** of f in the arbitrary direction $\mathbf{u}$. Note that the partial derivative with respect to x is the directional derivative of f in the direction of $\hat{\mathbf{x}}$ and likewise for y and z. That is, the partial derivative is the fundamental notion; the directional derivative is the more general notion which has the partial derivative as a special case.

Now let $\mathbf{u}$ vary over the unit sphere, i.e. consider all $\mathbf{u}$ such that $\|\mathbf{u}\| = 1$. Then $\mathbf{u} \cdot \nabla f$ is biggest when $\mathbf{u}$ points in the same direction as ∇f. This is true because

$$\mathbf{u} \cdot \nabla f = \|\mathbf{u}\| \, \|\nabla f\| \cos\theta = \|\nabla f\| \cos\theta$$

where θ is the angle between $\mathbf{u}$ and ∇f. This is biggest when $\theta = 0$, i.e. when $\mathbf{u}$ and ∇f point in the same direction. This means that the gradient of f points in the **direction of greatest change** of f.

13.2 Jacobian matrix TBD

xxx ways to motivate:

- Normal to surface (gradient as column vector)
- Tangent hyperplane to embedded submanifold: kernel of gradient.
- Regular-value thm (express in vr. calc. terms here)
- Impl. fcn. thm. (cf. Frankel)
- Inv. fcn. thm. (cf. Frankel p. 29)

13.3 Level sets TBD

Definition 13.2. Let $f : \mathbb{R}^3 \to \mathbb{R}$ be a vector-to-scalar function. The **level set** of f and c is the set of points

$$\{\mathbf{q} : f(\mathbf{q}) = c\}.$$

Example 13.3. $\triangleright$ If $f(x, y, z) = x^2 + y^2 + z^2$, then the level set of f and 1 is the unit sphere $\mathbb{S}^2$. $\triangleleft$

Remark 13.4. Recall that ∇f points in the direction of greatest change of f. When evaluated at $\mathbf{q}$, it is also the normal vector to the level sets at $\mathbf{q}$, written $\hat{\mathbf{n}}_{\mathbf{q}}$. Perpendicular to the normal is along the direction of the surface.

Remark 13.5. The **point-normal form** gives an equation for the **tangent plane** to the level-set surface. That is, if

$$\mathbf{q} = (x_0, y_0, z_0)$$

then the point-normal form of the tangent plane at $\mathbf{q}$ is

$$\frac{\partial f}{\partial x}(x - x_0) + \frac{\partial f}{\partial y}(y - y_0) + \frac{\partial f}{\partial z}(z - z_0) = 0.$$

Write this [xxx to do] in terms of (1) dot product, when ∇f is thought of as a vector, and (2) kernel of linear transformation $Df = \nabla f$, when ∇f is thought of as a 1×3 matrix. Conclude that the tangent space at q is the kernel of $Df|q$.

xxx to do: header style!!

xxx rcol/rowvectwo/three macros.

xxx really need to understand $(\nabla \times \mathbf{F}) \cdot \hat{\mathbf{n}} dA$.

If $\hat{\mathbf{n}} = (a, b, c)^t$ then what is dA? Try to work in point-normal form somehow ....

xxx notation: col/row stack.

References

[**Abr**] Abraham, R. and Marsden, J.E. *Foundations of Mechanics* (2nd ed.). Benjamin/Cummings, 1978.

[**Anton**] Anton, H.A. *Calculus with Analytic Geometry* (3rd ed.). Wiley, 1988.

[**Bachman**] Bachman, D. *A Geometric Approach to Differential Forms* Birkhäuser, 2006.

[**Boothby**] Boothby, W.M. *An Introduction to Differentiable Manifolds and Riemannian Geometry* (revised 2nd ed.). Academic Press, 2002.

[**BMA**] Boyer, C.B. and Merzbach, U.C. *A History of Mathematics* (revised 2nd ed.). Wiley, 1991.

[**Conlon**] Conlon, L. *Differentiable Manifolds* (2nd ed.). Birkhäuser, 2001.

[**CRC**] Zwillinger, D., editor. *CRC Standard Mathematical Tables and Formulae* (30th ed.). CRC Press, 1996.

[**Crowe**] Crowe, M.J. *A History of Vector Analysis: The Evolution of the Idea of a Vectorial System*. Dover, 1994.

[**DF**] Dummit, D.S., and Foote, R.M. *Abstract Algebra* (2nd ed.). John Wiley and Sons, 1999.

[**Dunham**] Dunham, W. *The Calculus Gallery: Masterpieces from Newton to Lebesgue*. Princeton, 2004.

[**FIS**] Friedberg, H., Insel, A., and Spence, L. *Linear Algebra* (3rd ed). Prentice Hall, 1997.

[**Frankel**] Frankel, T. *The Geometry of Physics: An Introduction* (2nd ed). Cambridge University Press, 2004.

[**Grove**] Grove, L.C. *Algebra*. Dover, 2004.

[**Guggenheimer**] Guggenheimer, H.W. *Differential Geometry*. Dover, 1977.

[**Hatcher**] Hatcher, A. *Algebraic Topology* (1st ed.). Cambridge, 2001.

[**Herstein**] Herstein, I.N. *Abstract Algebra*. Macmillan, 1990.

[**Hildebrand**] Hildebrand, F.B. *Advanced Calculus for Applications* (2nd ed). Prentice-Hall, 1976.

[**HHGM**] Hughes-Hallett, Gleason, McCallum, et al. *Calculus: Single and Multivariable* (4th ed.). Wiley, 2005.

[**Hungerford**] Hungerford, T.W. *Algebra*. Springer, 1997.

[**Kerl**] Kerl, J. *Numerical Differential Geometry in Matlab*. `http://math.arizona.edu/~kerl/gdg/gdgprojnotes.pdf`.

[**Lamb**] Lamb, M. *Differential Geometry Notes*. `http://math.arizona.edu/~mlamb/geometry.pdf`.

[**Lang**] Lang, S. *Algebra* (3rd ed.). Springer, 2002.

[**Lee1**] Lee, J. *Introduction to Topological Manifolds*. Springer, 2000.

[**Lee2**] Lee, J. *Introduction to Smooth Manifolds*. Springer, 2003.

[**Lee3**] Lee, J. *Riemannian Manifolds: An Introduction to Curvature*. Springer, 1997.

[**Massey**] Massey, W.S. *Algebraic Topology: An Introduction* (4th ed.). Springer, 1977.

[**PDM**] Nelson, D., editor. *Penguin Dictonary of Mathematics* (2nd ed.). Penguin Books, 1989.

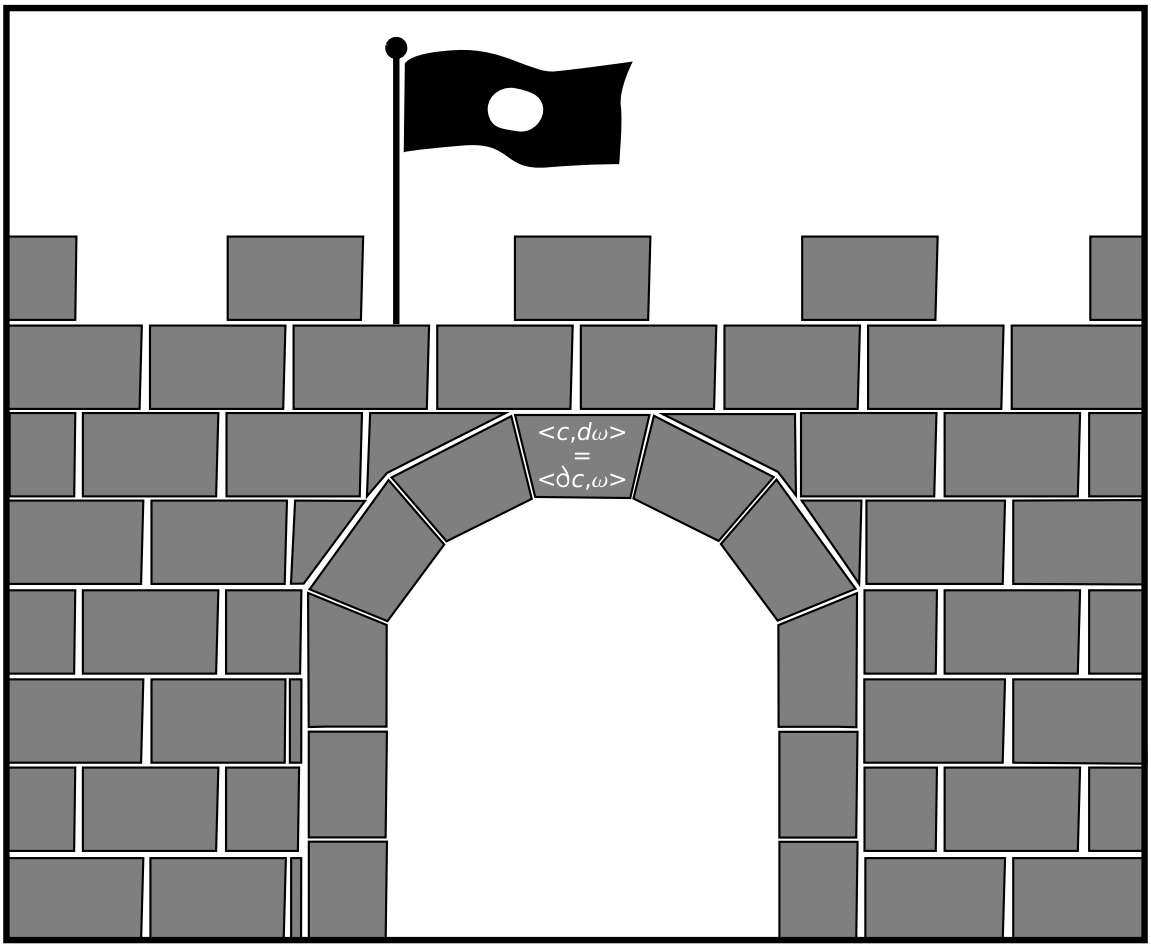
[**Pen**] Penrose, R. *The Road to Reality: A Complete Guide to the Laws of the Universe*. Knopf, 2005.

[**PM**] PlanetMath. `http://planetmath.org`.

[**Rudin**] Rudin, W. *Principles of Mathematical Analysis* (3rd ed.). McGraw-Hill, 1976.

[**Spivak1**] Spivak, M. *Calculus on Manifolds*. Perseus, 1965.

[**Spivak2**] Spivak, M. *A Comprehensive Introduction to Differential Geometry*. Publish or Perish, 2005.


<c,dω>
=
<∂c,ω>

Index

T

U

V

W

X