An Advanced Case of Calculus

James S. Cook
Liberty University
Department of Mathematics

Spring 2025

introduction and scope

Calculus has many chapters. In this course I endeavor to share methods of calculus which you have probably not encountered elsewhere. In contrast, there are other versions of *Advanced Calculus* which are more in the direction of an analysis class where the focus is less on calculation and more on existence and uniqueness and generality. We are less concerned with questions of abstractness and rigor here and more concerned with the simple question of *how do we calculuate ?*. That said, we do make some attempt at abstractness in our treatment of calculus over finite dimensional normed vector spaces. The last half of the course is more inutitive and calculational and we leave the proper development of the ideas in deeper rigor to graduate texts such as John M. Lee's *Smooth Manifolds.* The main theme for the first half of the course is to describe calculus on a vector space. The main theme for the second half of the course is to describe what differential forms are and to showcase their incredible application to calculus.

These notes are intended for someone who has completed the introductory calculus sequence and has some experience with matrices or linear algebra. This set of notes covers the first month or so of Math 332 at Liberty University. I intend this to serve as a supplement to our required text: *Manifolds, Tensors, and Forms: An Introduction for Mathematicians and Physicists* by Paul Renteln. These notes are mostly disconnected from Renteln; Chapters 1,2,3 and 6 of these notes stand alone from Renteln. On the other hand, there are sections in Renteln which I have not typed corresponding sections in these notes. In addition, on occasion I add a handout to round out lecture. Please consult the course planner for by best attempt at giving the day-by-day path of the course.

This course is primarily concerned with abstractions of calculus and geometry which are accessible to the undergraduate. This is not a course in real analysis and it does not have a prerequisite of real analysis so my typical students are not prepared for topology or measure theory. We defer the topology of manifolds and exposition of abstract integration in the measure theoretic sense to another course. Our focus for course as a whole is on what you might call the *linear algebra* of abstract calculus.

In particular, we study: basic linear algebra, spanning, linear independennce, basis, coordinates, norms, distance functions, inner products, metric topology, limits and their laws in a NLS, continuity of mappings on NLS's, linearization, Frechet derivatives, partial derivatives and continuous differentiability, linearization, properties of differentials, generalized chain and product rules, intuition and statement of inverse and implicit function theorems, implicit differentiation via the method of differentials, manifolds in $\mathbb{R}^n$ from an implicit or parametric viewpoint, tangent and normal spaces of a submanifold of Euclidean space, Lagrange multiplier technique, compact sets and the extreme value theorem, theory of quadratic forms, proof of real Spectral Theorem via method of Lagrange multipliers, higher derivatives and the multivariate Taylor theorem, multivariate power series, critical point analysis for multivariate functions, introduction to variational calculus. Multilinear algebra, symmetric and antisymmetric tensors, tensor products as multilinear maps, wedge products, Hodge duality, pull-backs and push-forwards, tangent and cotangent spaces to manifolds, tensor fields, exterior derivatives, integration of forms, the generalized Stokes' Theorem, Poincare's Lemma, potential theory, electromagnetism in differential form, 4D electromagnetism with Coulomb-type field.

The current set of notes was contructed from joining several past works. Basically, Chapters 1-6 are taken with very little modification from my 2017 notes for Advanced Calculus (the last half of the course we used McInerny's First Steps in Differential Geometry). Then Chapter 7-9 are taken from my 2015 notes for Advanced Calculus. Finally, Chapter 10 I wrote last semester for an second course in analysis. I don't intend to cover Chapter 10 in lecture, but I include it here for the elusive creature, the interested reader[1].

In contrast to some previous versions of this course, I do not study contraction mappings, differentiating under the integral and other questions related to uniform convergence. However, I have written a final chapter on these matters which contains some fairly complete proofs of the major existence theorems. In the regular flow of this course I present such theorems more intuitively and leave the rigor of the last chapter for another course. Likewise, we leave the complete exposition of manifold theory to a later course, I only touch on topological issues in these notes. We may have time to discuss more in class thanks to the very accessible entry level topology given in Renteln. Finally, I admit, and make no apology for the fact that the motivations and bent of much of this course is inspired by my interest in mathematical physics. I do hope this course provides a reasonable on-ramp to studying General Relativity. Certainly elementary differential geometry and manifold theory are natural extensions of our work here. If all goes as planned the final arc of lectures survey some important results from future courses and simultaneously showcase the important role differential forms play in the story of calculus on curved space and physics.

There are many excellent texts on calculus of many variables. Three which have had significant influence on my thinking and the creation of these notes are:

1. *Advanced Calculus of Several Variables* revised Dover Ed. by C.H. Edwards,

2. *Mathematical Analysis II*, Vladimir A. Zorich,

3. *Foundations of Modern Analysis*, by J. Dieudonné, Academic Press Inc. (1960)

These notes are a work in progress, do let me know about the errors. Thanks!

James S. Cook, December 18, 2024.

[1] credit to my advisor for this horrible saying

Contents

Chapter 1

on norms and limits

A normed linear space is a vector space which also has a concept of vector length. We use this length function to set-up limits for maps on normed linear spaces. The idea of the limit is the same as it was in first semester calculus; we say the map approaches a value when we can make values of the map arbitrary close to the value by taking inputs sufficiently close to the limit point. A map is continuous at a limit point in its domain if and only if its limiting value matches its actual value at the limit point. We derive the usual limit laws and work out results which are based on the component expansion with respect to a basis. We try to provide a fairly complete account of why common maps are continuous. For example, we argue why the determinant map is a continuous map from square matrices to real numbers.

We also introduce elementary concepts of topology. Open and closed sets are defined in terms of the metric topology induced from a given norm. We also discuss inner products and the more general concept of a distance function or metric. We explain why the set of invertible matrices is topologically open.

This Chapter concludes with a brief introduction into sequential methods. We define completeness of a normed linear space and hence introduce the concept of a Banach Space. Finally, the matrix exponential is shown to exist by an analytical appeal to the completeness of matrices.

Certain topics are not covered in depth in this work, I survey them here to attempt to provide context for the larger world of math I hope my students soon discover. In particular, while I introduce inner products, metric spaces and the rudiments of functional analysis, there is certainly far more to learn and indicate some future reading as we go. For future chapters we need to understand both linear algebra and limits carefully so my focus here is on normed linear spaces and limits. These suffice for us to begin our study of Frechet differentiation in the next chapter.

History is important and I must admit failure on this point. I do not know the history of these topics as deeply as I'd like. Similar comments apply to the next Chapter. I believe most of the linear algebra and analysis was discovered between about 1870 and 1910 by the likes of Frobenius, Frechet, Banach and other great analysts of that time, but, I have doubtless left out important work and names.

1.1 linear algebra

A real vector space is a set with operations of addition and scalar multiplication which satisfy a natural set of axioms. We call elements of the vector space **vectors**. We are primarily focused on **real** vector spaces which means the scalars are real numbers. Typical examples include:

(1.) $\mathbb{R}^n = \{(x_1, \dots, x_n) \ | x_1, \dots, x_n \in \mathbb{R}\}$ where for $x, y \in \mathbb{R}^n$ and $c \in \mathbb{R}$ we define $(x+y)_i = x_i + y_i$ and $(cx)_i = cx_i$ for each $i = 1, \dots, n$. In words, these are real n-tuples formed as column vectors. The notation $(x_1, \dots, x_n)$ is shorthand for $[x_1, \dots, x_n]^T$ in order to ease the typesetting.

(2.) $\mathbb{R}^{m \times n}$ the set of $m \times n$ real matrices. If $A, B \in \mathbb{R}^{m \times n}$ and $c \in \mathbb{R}$ then $(A+B)_{ij} = A_{ij} + B_{ij}$ and $(cA)_{ij} = cA_{ij}$ for all $1 \leq i \leq m$ and $1 \leq j \leq n$. Notice, an $m \times n$ matrix can be thought of as n-columns from $\mathbb{R}^m$ glued together, or as m-rows from $\mathbb{R}^{1 \times n}$ glued together (sometimes I say the rows or columns are concatenated)

$$A = \left[\begin{array}{cccc} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{m1} & A_{m2} & \cdots & A_{mn} \end{array} \right] = [col_1(A)|col_2(A)|\cdots|col_n(A)] = \left[\begin{array}{c} row_1(A) \\ \hline row_2(A) \\ \hline \vdots \\ \hline row_m(A) \end{array} \right] \tag{1.1}$$

In particular, it is at times useful to note: $(col_j(A)_i = A_{ij}$ and $(row_i(A))_j = A_{ij}$. Furthermore, in addition to the vector space structure, we also have a **multiplication** of matrices; for $A \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{n \times p}$ the matrix $AB \in \mathbb{R}^{m \times p}$ is defined by $(AB)_{ij} = row_i(A) \bullet col_j(B)$ which can be written in index notation as:

$$(AB)_{ij} = \sum_{k=1}^{n} A_{ik} B_{kj}. \tag{1.2}$$

(3.) $\mathbb{C}^n = \{(z_1, \dots, z_n) \ | \ z_1, \dots, z_n \in \mathbb{C}\}$. Once more define addition and scalar multiplication component-wise; for $z, w \in \mathbb{C}^n$ and $c \in \mathbb{C}$ define $(z+w)_i = z_i + w_i$ and $(cz)_i = cz_i$. Since $\mathbb{R} \subseteq \mathbb{C}$ the complex scalar multiplication in $\mathbb{C}^n$ also provides a real scalar multiplication. We can either view $\mathbb{C}^n$ as a real or complex vector space.

(4.) $\mathbb{C}^{m \times n}$ is the set of $m \times n$ complex matrices. If $Z, W \in \mathbb{C}^{m \times n}$ and $c \in \mathbb{C}$ then $(Z+W)_{ij} = Z_{ij} + W_{ij}$ and $(cZ)_{ij} = cZ_{ij}$ for $1 \leq i \leq m$ and $1 \leq j \leq n$. Just as in the previous example, we can view $\mathbb{C}^{m \times n}$ either as a complex vector space or as a real vector space.

(5.) If V and W are real vector spaces then $Hom_{\mathbb{R}}(V, W)$ is the set of all linear transformations from V to W. This forms a vector space with respect to the usual pointwise addition of functions. If $V = W$ then we denote $Hom_{\mathbb{R}}(V, V) = End_{\mathbb{R}}(V)$ for **endomorphisms** of V. The set of endomorphisms forms an algebra with respect to composition of functions since the composite of linear maps is once more linear. The set of invertible endomorphisms of V forms $GL(V)$. In the particular case that $V = \mathbb{R}^n$ we denote $GL(\mathbb{R}^n) = GL(n, \mathbb{R})$. Notice $GL(V)$ is not a subspace since $Id_V \in GL(V)$ where $Id_V(x) = x$ for all $x \in V$ and $Id_V - Id_V = 0 \notin GL(V)$.

Definition 1.1.1.

If V is real vector space and $S \subseteq V$ then define the **span of** S by

$$span(S) = \{c_1 s_1 + \cdots + c_k s_k \ | \ s_1, \dots, s_k \in S, c_1, \dots, c_k \in \mathbb{R}, k \in \mathbb{N}\}.$$

In words, *span*(S) is the set of all finite $\mathbb{R}$-linear combinations of vectors from S. Since the scalar multiple and linear combination of linear combinations is once more a linear combination we find that *span*$(S) \leq V$. That is, *span*(S) forms a **subspace** of V. The set S is called a **spanning set** or **generating set** for *span*(S).

Definition 1.1.2.

Let V be a real vector space and $S \subseteq V$. If $c_1, \dots, c_k \in \mathbb{R}$ and $s_1, \dots, s_k \in S$ with $c_1 s_1 + \cdots + c_k s_k = 0$ imply $c_1 = 0, \dots, c_k = 0$ for each $k \in \mathbb{N}$ then S is **linearly independent** (LI). Otherwise, we say S is linearly dependent.

When generating sets are linearly independent they are minimal, if you remove any vector from a minimal spanning set then the resulting span is smaller. In contrast, if S is linearly dependent then there exists $S' \subset S$ for which $span(S') = span(S)$. Our convention is that $span(\emptyset) = \{0\}$.

Definition 1.1.3.

Let V be a real vector space. If β is a linearly independent spanning set for V then we say β is a **basis** for V. Furthermore, using # to denote cardnality, $\#(\beta)$ is the **dimension** of V. If $\#(\beta) = n \in \mathbb{N}$ then we say V is an n-dimensional vector space and write $dim(V) = n$.

Bases are very helpful for calculations. In particular, if $\beta = \{v_1, \dots, v_n\}$ then

$$x_1 v_1 + \cdots + x_n v_n = y_1 v_1 + \cdots + y_n v_n \quad \Rightarrow \quad x_i = y_i \text{ for } i = 1, \dots, n. \tag{1.3}$$

We call this calculation **equating coefficients** with respect to the basis β.

Definition 1.1.4.

Let V be a real finite dimensional vector space with basis $\beta = \{v_1, \dots, v_n\}$ then for each $x \in V$ there exist $c_i \in \mathbb{R}$ for which $x = c_1 v_1 + \cdots + c_n v_n$. We write $[x]_\beta = (c_1, \dots, c_n)$ and say $[x]_\beta$ is the **coordinate vector** of x with respect to the β basis. We also denote $\Phi_\beta(x) = [x]_\beta$ and say $\Phi_\beta : V \to \mathbb{R}^n$ is the **coordinate map** with respect to the basis β.

If $\beta = \{v_1, \dots, v_n\}$ is a basis for the real vector space V and $\psi \in GL(V)$ then $\psi(\beta) = \{\psi(v_1), \dots, \psi(v_n)\}$ forms a basis for V. Clearly the choice of basis is far from unique. That said, it is useful for us to settle on a **standard basis** for our usual real examples:

(1.) Let $(e_i)_j = \delta_{ij}$ hence $e_1 = (1, 0, \dots, 0)$, $e_2 = (0, 1, 0, \dots, 0)$ and $e_n = (0, \dots, 0, 1)$. If $\beta = \{e_1, \dots, e_n\}$ then $x = (x_1, \dots, x_n) = \sum_{i=1}^n x_i e_i$ and $[x]_\beta = x$. We say β is the **standard basis of column vectors** and note $\#(\beta) = n = dim(\mathbb{R}^n)$.

(2.) Let $(E_{ij})_{kl} = \delta_{ik}\delta_{jl}$ for $1 \leq i, k \leq m$ nad $1 \leq j, l \leq n$ define $E_{ij} \in \mathbb{R}^{m \times n}$. The matrix E_{ij} has a 1 in the ij-th entry and zeros elsewhere. For any $A \in \mathbb{R}^{m \times n}$ we have

$$A = \sum_{i=1}^{m} \sum_{j=1}^{n} A_{ij} E_{ij} \tag{1.4}$$

We order the **standard** $m \times n$ **matrix basis** $\beta = \{E_{ij} \mid 1 \leq i \leq m, 1 \leq j \leq n\}$ by the usual lexographic ordering. For example, in the case of 2×3 matrices,

$$\beta = \{E_{11}, E_{12}, E_{13}, E_{21}, E_{22}, E_{23}\} \tag{1.5}$$

Following the notation from Equation 1.1,

$$\Phi_\beta(A) = (A_{11}, A_{12}, \dots, A_{1n}, A_{21}, A_{21}, \dots, A_{2n}, \dots, A_{mn}) \tag{1.6}$$

The coordinate vector for $A \in \mathbb{R}^{m \times n}$ w.r.t. the standard basis is given by listing out the components of A row-by-row. Also, $\#(\beta) = mn = dim(\mathbb{R}^{m \times n})$.

Viewing $\mathbb{C}^n$ and $\mathbb{C}^{m\times n}$ as real vector spaces there are at least two natural choices for the basis,

(3.) For $\mathbb{C}^n$ notice $\beta = \{e_1, ie_1, \dots, e_n, ie_n\}$ and $\gamma = \{e_1, \dots, e_n, ie_1, \dots, ie_n\}$ serve as natural bases. If $z = x + iy$ where $x, y \in \mathbb{R}^n$ then we **define** $Re(z) = x$ and $Im(z) = y$. Hence,[1]

$$\Phi_\gamma(z) = (x, y), \quad \& \quad \Phi_\beta(z) = (x_1, y_1, x_2, y_2, \dots, x_n, y_n). \tag{1.7}$$

Note $dim_{\mathbb{R}}(\mathbb{C}^n) = 2n$.

(4.) For $\mathbb{C}^{m\times n}$ notice $\beta = \{E_{11}, iE_{11}, \dots, E_{mn}, iE_{mm}\}$ and $\gamma = \{E_{11}, \dots, E_{mn}, iE_{11}, \dots, iE_{mn}\}$ serve as natural bases. If $Z = X + iY$ where $X, Y \in \mathbb{R}^{m\times n}$ then we **define** $Re(Z) = X$ and $Im(Z) = Y$. With this notation,

$$[Z]_\gamma = (X_{11}, \dots, X_{mn}, Y_{11}, \dots, Y_{mn}) \quad \& \quad [Z]_\beta = (X_{11}, Y_{11}, \dots, X_{mn}, Y_{mn}). \tag{1.8}$$

For example,

$$A = \begin{bmatrix} 1+i & 2 \\ 3+4i & 5i \end{bmatrix} \quad \Rightarrow \quad [A]_\beta = (1,1,2,0,3,4,0,5) \ \ \& \ \ [A]_\gamma = (1,2,3,0,1,0,4,5).$$

Finally, note $dim_{\mathbb{R}}(\mathbb{C}^{m\times n}) = 2mn$

Naturally, $dim_{\mathbb{C}}(\mathbb{C}^n) = n$ and $dim_{\mathbb{C}}(\mathbb{C}^{m\times n}) = mn$, but, our primary interest is in the calculus of real vector spaces so we just need such formulas as a conceptual backdrop.

1.2 norms, metrics and inner products

The concept of norm, metric and inner product all strike a the same issue; how to describe distance abstractly. Of these the inner product is most special and the metric or **distance function** is most general. In particular, both norms and inner products require a background vector space. In constrast, distance functions can be given to all sorts of sets where there is no well-defined addition which closes on the set. The general study of distance functions belongs to real or functional analysis, however, I think it is important to mention them here for context.

1.2.1 normed linear spaces

This definition abstracts the concept of vector length:

Definition 1.2.1. *Normed Linear Space (NLS):*

Suppose V is a real vector space. If $||\cdot|| : V \times V \to \mathbb{R}$ is a function such that for all $x, y \in V$ and $c \in \mathbb{R}$:

(1.) $||cx|| = |c|\,||x||$

(2.) $||x + y|| \leq ||x|| + ||y||$ (triangle inequality)

(3.) $||x|| \geq 0$

(4.) $||x|| = 0$ if and only if $x = 0$

then we say $(V, ||\cdot||)$ is a normed vector space. When there is no danger of ambiguity we also say that V is a **normed vector space** or a **normed linear space (NLS)**.

Notice that we did not assume V was finite-dimensional in the definition above. Our current focus is on finite-dimensional cases.

[1] technically, this is an abuse of notation, I'm ignoring the distinction between a vector of vectors and a vector

(1.) the standard **euclidean** norm on $\mathbb{R}^n$ is defined by $||v|| = \sqrt{v \bullet v}$.

(2.) the **taxicab** norm on $\mathbb{R}^n$ is defined by $||v||_1 = \sum_{j=1}^{n} |v_j|$.

(3.) the **sup** norm on $\mathbb{R}^n$ is defined by $||v||_\infty = \max\{|v_j| \mid j = 1, \dots, n\}$

(4.) the p-norm on $\mathbb{R}^n$ is defined by $||v||_p = \left(\sum_{j=1}^{n} |v_j|^p\right)^{1/p}$ for $p \in \mathbb{N}$. Notice $||v|| = ||v||_2$ and the taxicabl is the $p = 1$ case and finally the sup-norm appears as $p \to \infty$.

If we identify points with vectors based at the origin then it is natural to think about a **circle** of radius 1 as the set of vectors (points) which have norm (distance) one. Focusing on $n = 2$,

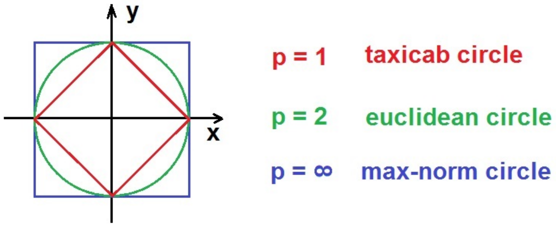
(1.) in the euclidean case a circle is the circle for $\mathbb{R}^2$ with $||v||_2 = 1$.

(2.) if we use the taxicab norm on $\mathbb{R}^2$ then the circle is a diamond.

(3.) for $\mathbb{R}^2$ with the p-norm the circle is something like a blown-up circle.

(4.) as $p \to \infty$ the circle expands to a square.

In other words, to **square the circle** we need only study p-norms in the plane.



We could play similar games for our other favorite examples, but primarily we just use the analog of the $p = 1, 2$ or ∞ norms in our application of norms. Let us agree by **convention** that $||x|| = \sqrt{x \bullet x}$ for $x \in \mathbb{R}^n$, since the coordinate map yields real column vectors the r.h.s. makes use of this convention in each of the following examples:

(2.) the standard norm for $\mathbb{R}^{m\times n}$ is given by $||A|| = ||\Phi_\beta(A)||$ where Φ_β is the standard coordinate map for $\mathbb{R}^{m\times n}$ as defined in Equation 1.6.

(3.) the standard norm for $\mathbb{C}^n$ is given by $||z|| = ||\Phi_\beta(z)||$ where Φ_β is the standard coordinate map described in Equation 1.7.

(4.) the standard norm for $\mathbb{C}^{m\times n}$ is given by $||Z|| = ||\Phi_\beta(Z)||$ where Φ_β is the standard coordinate map for $\mathbb{C}^{m\times n}$ as defined in Equation 1.8.

In each case above there is some slick formula which hides the simple truth I described above; the length of matrices and complex vectors is simply the Euclidean length of the corresponding coordinate vectors.

$$||v||^2 = v^T\bar{v}, \qquad ||A||^2 = \text{trace}(A^TA), \qquad ||Z||^2 = \text{trace}(Z^\dagger Z)$$

where the complex vector $v = (v_1, \dots, v_n)$ has conjugate vector $\bar{v} = (\bar{v}_1, \dots, \bar{v}_n)$ and the complex matrix Z has conjugates $\bar{Z}$ defined by $(\bar{Z})_{ij} = \bar{Z}_{ij}$ and $Z^\dagger = \bar{Z}^T$ is the **Hermitian conjugate**. Again, to be clear, there is not just one choice of norm for $\mathbb{C}^n, \mathbb{R}^{m\times n}$ or $\mathbb{C}^{m\times n}$. The set paired with the norm is what gives us the structure of a normed space. We conclude this Section with norms which are a bit less obvious.

Example 1.2.2. *Let $C([a,b],\mathbb{R})$ denote the set of continuous real-valued functions with domain $[a,b]$. If $f \in C([a,b],\mathbb{R})$ then we define $\|f\| = max\{|f(x)| \mid x \in [a,b]\}$. It is not too difficult to check this defines a norm on the infinite dimensional vector space $C([a,b],\mathbb{R})$.*

Example 1.2.3. *Suppose V, W are normed linear spaces and $T : V \to W$ is a linear transformation. Then we may define the norm of $\|T\|$ as follows:*

$$\|T\| = sup\{\|T(x)\| \mid x \in V, \ \|x\| = 1\}$$

When V is infinite dimensional there is no reason that $\|T\|$ must be finite. In fact, the linear transformations with finite norm are special. I leave the completion of this thought to your functional analysis course. On the other hand, for finite dimensional V we can argue $\|T\|$ is finite.

Incidentally, given $T : V \to W$ with $\|T\| < \infty$ you can show $\|T(x)\| \leq \|T\|\|x\|$ for all $x \in V$. To see this claim, consider $x \neq 0$ has $\|x\| \neq 0$ hence:

$$\begin{aligned} \|T(x)\| &= \left\| T\left(\frac{\|x\|}{\|x\|}x\right)\right\| \qquad (1.9) \\ &= \left\| \|x\| T\left(\frac{x}{\|x\|}\right)\right\| \\ &= \|x\| \left\| T\left(\frac{x}{\|x\|}\right)\right\| \\ &\leq \|x\|\|T\| \end{aligned}$$

as $\|x/\|x\|\| = 1$ so $\|T\|$ certainly provides the claimed bound.

I include the next example to give you a sense of what sort of calculation takes the place of coordinates in infinite dimensions. I'm mostly including these examples so we can appreciate the technical meaning of **continuously differentiable** in our later work.

Example 1.2.4. *Assume $a < b$. Define $T(f) = \int_a^b f(x)\ dx$ for each $f \in C([a, b], \mathbb{R})$. Observe T is a linear transformation. Also,*

$$|T(f)| = \left| \int_a^b f(x)dx \right| \leq \int_a^b |f(x)|dx.$$

Use the max-norm of Example 1.2.2. If $\|f\| = max\{|f(x)| \mid x \in [a, b]\} = 1$ then $|f(x)| \leq 1$ for $a \leq x \leq b$. Thus $|T(f)| \leq \int_a^b dx = b - a$. However, the constant function $f(x) = 1$ has $\|f\| = 1$ and $T(1) = \int_a^b dx = b - a$ thus $\|T\| = sup\{\|T(f)\| \mid f \in C([a, b], \mathbb{R}),\ \|f\| = 1\} = b - a$.

At this point I introduce some notation I found in Zorich. I think it's a useful addition to my standard notations. Pay attention to the semi-colon.

Definition 1.2.5. *multilinear maps*

Let $V_1, V_2, \dots, V_k, W$ be real vector spaces then $T : V_1 \times V_2 \times \cdots \times V_k \to W$ is a **multilinear map** if T is linear in each of its k-variables while holding the other variables fixed. We write $T \in \mathcal{L}(V_1, V_2, \dots, V_k; W)$ in this case.

In the case $k = 1$ and $V_1 = V_2 = V$ we say $T \in \mathcal{L}(V, V; W)$ is a W-valued bilinear map on V.

Example 1.2.6. *If $T \in \mathcal{L}(V_1, \dots, V_k; W)$ where $V_1, \dots, V_k, W$ are normed linear spaces then define*[2]

$$\|T\| = sup\{\|T(u_1, \dots, u_k)\| \mid \|u_i\| = 1, i = 1, 2, \dots, k\}. \qquad (1.10)$$

[2] see page 52 of Zorich's *Mathematical Analysis II* for further discussion

Then we can argue, much as we did in Equation 1.9 that

$$||T(x_1, x_2, \dots, x_n)|| \leq ||T|| ||x_1|| ||x_2|| \cdots ||x_n||. \tag{1.11}$$

Notice $det : \mathbb{R}^n \times \cdots \times \mathbb{R}^n \to \mathbb{R} \in \mathcal{L}(\mathbb{R}^n, \dots, \mathbb{R}^n; \mathbb{R})$. *Hence, as*

$$||det|| = sup\{|det[u_1| \cdots |u_n]| \mid ||u_i|| = 1, i = 1, \dots, n\} \tag{1.12}$$

and

$$|det(x_1, x_2, \dots, x_n)| \leq ||det|| ||x_1|| ||x_2|| \cdots ||x_n||. \tag{1.13}$$

But, $det(I) = 1$ *thus* $||det|| = 1$.

I've probably done a bit more than we need here, I hope it is not too disturbing.

1.2.2 inner product space

There are generalized dot-products on many abstract vector spaces, we call them **inner-products**.

Definition 1.2.7. *Inner product space*

Suppose V is a real vector space. If $\langle \ , \ \rangle : V \times V \to \mathbb{R}$ is a function such that for all $x, y, z \in V$ and $c \in \mathbb{R}$:

(1.) $\langle x, y \rangle = \langle y, x \rangle$ (symmetric)

(2.) $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$ (additive in the first slot)

(3.) $\langle cx, y \rangle = c\langle x, y \rangle$ (together with (2.) gives linearity of the first slot)

(4.) $\langle x, x \rangle \ \geq \ 0$ and $\langle x, x \rangle = 0$ if and only if $x = 0$.

then we say $(V, \langle \ , \ \rangle)$ is an **inner-product space** with inner product $\langle \ , \ \rangle$.

Given an inner-product space $(V, \langle \ , \ \rangle)$ we can easily induce a norm for V by the formula $||x|| = \sqrt{\langle x, x \rangle}$ for all $x \in V$. Properties (1.), (3.) and (4.) in the definition of the norm are fairly obvious for the induced norm. Let's think throught the triangle inequality for the induced norm:

$$\begin{array}{rcll}
||x+y||^2 & = & \langle x+y, x+y \rangle & \text{def. of induced norm} \\
& = & \langle x, x+y \rangle + \langle y, x+y \rangle & \text{additive prop. of inner prod.} \\
& = & \langle x+y, x \rangle + \langle x+y, y \rangle & \text{symmetric prop. of inner prod.} \\
& = & \langle x, x \rangle + \langle y, x \rangle + \langle x, y \rangle + \langle y, y \rangle & \text{additive prop. of inner prod.} \\
& = & ||x||^2 + 2\langle x, y \rangle + ||y||^2 &
\end{array}$$

At this point we're stuck. A nontrivial identity[3] called the **Cauchy-Schwarz** identity helps us proceed; $\langle x, y \rangle \leq ||x|| ||y||$. It follows that $||x+y||^2 \leq ||x||^2 + 2||x|| ||y|| + ||y||^2 = (||x|| + ||y||)^2$. However, the induced norm is clearly positive so we find $||x+y|| \leq ||x|| + ||y||$.

Most linear algebra texts have a whole chapter on inner-products and their applications, you can look at my notes for a start if you're curious. That said, this is a bit of a digression for this course. Primarily we use the dot-product paired with $\mathbb{R}^n$ in certain applications. I should mention, $\mathbb{R}^n$ with the usual dot-product forms **Euclidean n-space**. We'll say more just before we use the theory of orthogonal complements to understand how to find extreme values on curves or surfaces.

[3] I prove this for the dot-product in my linear notes, however, the proof is written in such a way it equally well applies to a general inner-product

1.2.3 metric as a distance function

Given a set S a distance function describes the **distance** between points in S. This definition is a natural abstraction of our everyday idea of distance.

Definition 1.2.8.

A function $d : S \times S \to \mathbb{R}$ is a **metric** or **distance function** on S if d satisfies the following: for all $x, y, z \in S$,

(1.) $d(x, y) \geq 0$ (non-negativity)

(2.) $d(x, y) = d(y, x)$ (symmetric)

(3.) $d(x, z) \leq d(x, y) + d(y, z)$ (triangle inequality)

(4.) $d(x, y) = 0$ if and only if $x = y$

then we say (S, d) is a **metric space**.

There are many strange examples one may study, I leave those to your future courses. For our purposes, note any subset of a NLS forms a metric space via the distance function $d(x, y) = \|y-x\|$. Geometrically, the idea is that the distance from the point x to the point y is the length of the displacement vector $y - x$ which goes from x to y. Of course, we could just as well write $d(x, y) = \|x - y\|$ since $\|x - y\| = \|(-1)(y - x)\| = |-1|\|y - x\|$.

Remark 1.2.9. *There is another use of the term* **metric**. *In particular, $g : V \times V \to \mathbb{R}$ is a metric if it is symmetric, bilinear and nondegenerate. Then (V, g) forms a* **geometry**. *We say $T : V \to V$ is an* **isometry** *if $g(T(x), T(y)) = g(x, y)$ for all $x, y \in V$. For example, if $g(x, y) = -x^0y^0 + x^1y^1 + x^2y^2 + x^3y^3$ for $x, y \in \mathbb{R}^4$ then g is the* **Minkowski metric** *and isometries of this metric are called* **Lorentz transformations**. *To avoid confusion, I try to use the term* **scalar product** *rather than metric. An inner product is a scalar product which is positive definite. Riemannian geometry is based on an abstraction of inner products to curved space whereas semi-Riemannian geometry generalizes the Minkowski metric to curved space. The geometry of Eistein's General Relativity is semi-Riemannian geometry.*

1.3 topology and limits in normed linear spaces

The limit we describe here is the natural extension of the $\epsilon - \delta$-limit from elementary calculus. Recall, we say $f(x) \to L$ as $x \to a$ if for each $\epsilon > 0$ there exists $\delta > 0$ such that $0 < |x - a| < \delta$ implies $|f(x) - L| < \epsilon$. Essentially, this limit is constructed to zoom in on the values taken by f as they get close to a, yet, not $x = a$ itself. Avoidance of the limit point itself allows us to extend the algebra of limits past the confines of unqualified algebra. The same holds for NLS we simply need to replace absolute values with norms. This goes back to the metric space structure involved. For $\mathbb{R}$ we have distance function given by abolute value of the difference $d(x, y) = |x - y|$. For a NLS $(V, \| \, \|)$ we have distance function given by the norm of the difference $d(x, y) = \|x - y\|$. Keeping this analogy in mind it is not hard to see all the definitions in what follows as simple extensions of the analysis we learned in first semester calculus[4]

[4] at Liberty University we still cover elementary $\epsilon - \delta$ proofs in the beginning calculus course

Definition 1.3.1. *open and closed sets in an NLS*

Let $(V, \| \ \|)$ be a NLS. An **open ball centered at x_o with radius R** is:

$$B_R(x_o) = \{x \in V \mid \|x - x_o\| < R\}.$$

Likewise, a **closed ball centered at x_o with radius R**[5]

$$\overline{B_R(x_o)} = \{x \in V \mid \|x - x_o\| \leq R\}.$$

If $x_o \in U$ and there exists $R > 0$ for which $B_R(x_o) \subseteq U$ then we say x_o is an **interior point** of U. When each point in $U \subseteq V$ is an interior point we say U is an **open set**. If $S \subset V$ has $V - S$ open then we say S is a **closed set**.

In the case $V = \mathbb{R}^n$ with $n = 1, 2, 3$ we have other terms.

(1.) $n = 1$: an open ball is an open interval; $B_r(a) = (a - r, a + r)$,

(2.) $n = 2$: an open ball is an open disk,

(3.) $n = 3$: an open ball is an open ball,

Intuitively, an open set either has no edges, or, has only fuzzy edges whereas a closed set either has no edges, or, has solid edges. The larger problem of studying which sets are open and how that relates to the continuity of functions is known as **topology**. Briefly, a **topology** is a set paired with the set of all sets declared to be open. The topology we study here is **metric topology** as it is derived from a distance function. Moving on,

Definition 1.3.2. *limit points, isolated points and boundary points in an NLS.*

Let $(V, \| \ \|)$ be a NLS. We define a **deleted open ball centered at x_o with radius R** by:

$$B_R(x_o) - \{x_o\} = \{x \in V \mid 0 < \|x - x_o\| < R\}.$$

We say x_o is a **limit point** of a function f if and only if there exists a deleted open ball which is contained in the $dom(f)$. If $y_o \in dom(f)$ and there exists an open ball centered at y_o which contains no other points in $dom(f)$ then y_o is called an **isolated point** of $dom(f)$. A **boundary point** of $S \subseteq V$ is a point in $x_o \in V$ for which every open ball centered at x_o contains points outside S.

Notice a limit point of f need not be in the domain of f. Also, a boundary point of S need not be in S. Furthermore, if we consider $g : \mathbb{N} \to V$ then each point in $dom(g) = \mathbb{N}$ is isolated.

Definition 1.3.3. *limits and continuity in an NLS.*

If $f : dom(f) \subseteq V \to W$ is a function from normed space $(V, || \cdot ||_V)$ to normed vector space $(W, || \cdot ||_W)$ and x_o is either a limit point or an isolated point of $dom(f)$ and $L \in W$ then we say $\lim_{x \to x_o} f(x) = L$ if and only if for each $\epsilon > 0$ there exists $\delta > 0$ such that if $x \in V$ with $0 < ||x - x_o||_V < \delta$ then $||f(x) - f(x_o)||_W < \epsilon$. If $\lim_{x \to x_o} f(x) = f(x_o)$ then we say that f is a continuous function at x_o.

The definition above indicates functions are by default continuous at isolated points, my apologies if you find this bothersome. Let me give a few examples then we'll turn our attention to proving limit laws for an NLS.

Example 1.3.4. *Suppose V is an NLS and let $c \in \mathbb{R}$ with $c \neq 0$. Also fix $b_o \in V$. Let $F(x) = cx + b_o$ for each $x \in V$. We wish to calculate $\lim_{x\to a} F(x)$. Naturally, we expect the limit is simply $ca + b_o$ hence we work towards proving our intutiion is correct. If $\epsilon > 0$ then choose $\delta = \epsilon/|c|$ and note $0 < \|x - a\| < \delta = \epsilon/|c|$ provides $0 < |c|\|x - a\| < \epsilon$. With this estimate in mind we calculate:*

$$\|F(x) - F(a)\| = \|cx + b_o - (ax + b_o)\| = \|c(x - a)\| = |c|\|x - a\| < \epsilon.$$

Thus $F(x) \to F(a) = ca + b_o$ as $x \to a$. As $a \in V$ was arbitrary we've shown F is continuous on V. Specializing a bit, if we set $c = 1$ and $b_o = 0$ then $F = Id_V$ thus the **identity function** *on V is everywhere continuous.*

Example 1.3.5. *Let V and W be normed linear spaces. Fix $w_o \in W$ and define $F(x) = w_o$ for each $x \in V$. I leave it to the reader to prove $\lim_{x\to a}(F(x)) = w_o$ for any $a \in V$. In other words, a constant function is everywhere continuous in the context of a NLS.*

Example 1.3.6. *Let $F : \mathbb{R}^n - \{a\} \to \mathbb{R}^n$ be defined by $F(x) = \frac{1}{||x-a||}(x - a)$. In this case, certainly a is a limit point of F but geometrically it is clear that $\lim_{x\to a} F(x)$ does not exist. Notice for $n = 1$, the discontinuity of F at a can be understood by seeing that left and right limits exist, but are not equal. On the other hand, $G(x) = \frac{||x-a||}{||x-a||}(x - a)$ clearly has $\lim_{x\to a} G(x) = 0$ and we could classify the discontinuity of G at $x = a$ as removeable. Clearly $\tilde{G}(x) = x - a$ is a continuous extension of G to all of $\mathbb{R}^n$*

On occasion it is helpful to keep the following observation in mind:

Proposition 1.3.7. *norm is continuous with respect to itself.*

Suppose V has norm $||\cdot||$ then $f : V \to \mathbb{R}$ defined by $f(x) = ||x||$ is continuous.

Proof: Suppose $a \in V$ and let $\epsilon > 0$. Choose $\delta = \epsilon$ and consider $x \in V$ such that $0 < ||x - a|| < \delta$. Observe $||x|| = ||x - a + a|| \leq ||x - a|| + ||a|| = \delta + ||a||$ and hence

$$|f(x) - f(a)| = |||x|| - ||a||| < |\delta + ||a|| - ||a||| = |\delta| = \epsilon.$$

Thus $f(x) \to f(a)$ as $x \to a$ and as $a \in V$ was arbitrary the proposition follows $\square$.

It is generally quite challenging to prove limits directly from the definition. Fortunately, there are many useful properties which typically allow us to avoid direct attack.[6] One fun point to make here, if you missed the proof of the so-called *limit laws* in calculus then you can retroactively apply the arguments we soon offer here.

Proposition 1.3.8. *Linearity of the limit on a NLS.*

Let V, W be normed vector spaces. Let a be a limit point of mappings $F, G : U \subseteq V \to W$ and suppose $c \in \mathbb{R}$. If $\lim_{x\to a} F(x) = b_1 \in W$ and $\lim_{x\to a} G(x) = b_2 \in W$ then

(1.) $\lim_{x\to a}(F(x) + G(x)) = \lim_{x\to a} F(x) + \lim_{x\to a} G(x)$.

(2.) $\lim_{x\to a}(cF(x)) = c \lim_{x\to a} F(x)$.

Moreover, if F, G are continuous then $F + G$ and cF are continuous.

[6] of course, some annoying instructor probably will ask you to calculate a couple from the definition so you can learn the definition more deeply

Proof: Let $\epsilon > 0$ and suppose $\lim_{x\to a} f(x) = b_1 \in W$ and $\lim_{x\to a} g(x) = b_2 \in W$. Choose $\delta_1, \delta_2 > 0$ such that $0 < ||x-a|| < \delta_1$ implies $||f(x)-b_1|| < \epsilon/2$ and $0 < ||x-a|| < \delta_2$ implies $||g(x)-b_2|| < \epsilon/2$. Choose $\delta = min(\delta_1, \delta_2)$ and suppose $0 < ||x-a|| < \delta \leq \delta_1, \delta_2$ hence

$$||(f+g)(x) - (b_1+b_2)|| = ||f(x) - b_1 + g(x) - b_2|| \leq ||f(x) - b_1|| + ||g(x) - b_2|| < \epsilon/2 + \epsilon/2 = \epsilon.$$

Item (2.) follows. To prove (2.) note that if $c = 0$ the result is clearly true so suppose $c \neq 0$. Suppose $\epsilon > 0$ and choose $\delta > 0$ such that $||f(x) - b_1|| < \epsilon/|c|$. Note that if $0 < ||x-a|| < \delta$ then

$$||(cf)(x) - cb_1|| = ||c(f(x) - b_1)|| = |c|||f(x) - b_1|| < |c|\epsilon/|c| = \epsilon.$$

The claims about continuity follow immediately from the limit properties. □

Induction easily extends the result above to linear combinations of three or more functions;

$$\lim_{x\to a} \sum_{i=1}^{n} c_i F_i(x) = \sum_{i=1}^{n} c_i \lim_{x\to a} F_i(x). \tag{1.14}$$

We now turn to analyzing limits of a map in terms of the limits of its component functions. First a Lemma which is a slight twist on what we already proved.

Lemma 1.3.9. *Constant vectors pull out of limit.*

Let V be a NLS and suppose $f : dom(f) \subseteq V \to \mathbb{R}$ is a function with $\lim_{x\to a} f(x) = L$. If W is a NLS with $w_o \in W$ then $\lim_{x\to a}(f(x)w_o) = Lw_o$.

Proof: if $w_o = 0$ then the Lemma is clearly true. Hence suppose $w_o \neq 0$ thus $\|w_o\| \neq 0$. Also, we assume $f(x) \to L$ as $x \to a$. Let $\epsilon > 0$ and note we are free to choose $\delta > 0$ such that $0 < \|x-a\| < \delta$ implies $|f(x) - L| < \epsilon/\|w_o\|$. Thus, for $x \in V$ with $0 < \|x-a\| < \delta$

$$\|f(x)w_o - Lw_o\| = \|(f(x) - L)w_o\| = |f(x) - L|\|w_o\| < \frac{\epsilon}{\|w_o\|}\|w_o\| = \epsilon. \tag{1.15}$$

Consequently $\lim_{x\to a}(f(x)w_o) = \lim_{x\to a}(f(x))w_o$. □

We soon need this Lemma to pull basis vectors out of a limit in the proof of Theorem 1.3.11.

Definition 1.3.10. *Component functions of map with values in an NLS.*

Suppose V, W are normed linear spaces and $dim(W) = m$. If $F : dom(F) \subseteq V \to W$ and $\gamma = \{w_1, w_2, \dots, w_m\}$ is a basis for W and there exist $F_i : dom(F) \subseteq V \to \mathbb{R}$ for $i = 1, 2, \dots, m$ such that $F = F_1w_1 + \cdots + F_mw_m$ and we call $F_1, \dots, F_m$ the **component functions** of F with respect to the γ basis.

Given the limits of each component function we may assemble the limit of the function. Notice, this is a comment about breaking up the limit in the range of the map. In contrast, there is no easy way to break a multivariate limit into one-dimensional limits in the domain, hopefully you saw examples in multivariable calculus which illustrate this subtle point. Only in one dimension to we have the luxury of reducing a full limit to a pair of path limits. See this question and answer, beware wolfram alpha not so good here, Maple wins and this master list of advice on how to calculate multivariate limits that arise in calculus III. There are many examples linked there if you need to see evidence of my claim here.

Theorem 1.3.11.

Suppose V, W are NLSs where W has basis $\gamma = \{w_1, \dots, w_n\}$ and $F : dom(F) \subseteq V \to W$ has component functions $F_i : dom(F) \subseteq V \to \mathbb{R}$ for $i = 1, \dots, m$. If $\lim_{x\to a} F_i(x) = L_i \in \mathbb{R}$ for $i = 1, \dots, m$ then $\lim_{x\to a} F(x) = \sum_{i=1}^m L_i w_i$.

Proof: assume F and its components are as described in the Proposition,

$$\begin{aligned} \lim_{x\to a} F(x) &= \lim_{x\to a}\left(\sum_{i=1}^m F_i(x)w_i\right) && : \text{defn. of component functions} \quad (1.16)\\ &= \sum_{i=1}^m \lim_{x\to a}\left(F_i(x)w_i \right) && : \text{additivity of the limit}\\ &= \sum_{i=1}^m \left(\lim_{x\to a} F_i(x)\right) w_i && : \text{applying Lemma 1.3.9}\\ &= \sum_{i=1}^m L_i w_i. \end{aligned}$$

Therefore, the limit of a map may be assembled from the limits of its component functions. □

It turns out the converse of this Theorem is also true, but, I need to prepare some preliminary ideas to give the proof in the desired generality. Basically, the trouble is that at one point in my proof I need the magnitude of a component to a vector $x = x_1v_1 + \cdots + x_nv_n$ to be smaller than the norm of the whole vector; $|x_i| \leq \|x\|$. Certainly this is true for orthonormal bases, but, notice $\beta = \{(1, \varepsilon), (1, 0)\}$ is a basis for $\mathbb{R}^2$ which is not orthonormal in the euclidean sense for any $\varepsilon \neq 0$ and:

$$x = (1, \varepsilon) - (1, 0) = (0, \varepsilon) \tag{1.17}$$

hence $\|x\| = |\varepsilon|$ and $[x]_\beta = (1, -1)$ so both components of x in the β basis have magnitude 1. But, we can make $|\varepsilon|$ as small as we like. So, clearly, I cannot just assume for any basis of a NLS we have this property $|x_i| \leq \|x_1v_1 + \cdots + x_nv_n\|$. It is a special property for certain nice bases. In fact, it is true for most examples we consider. You use it a great deal in study of complex analysis as it says $|Re(z)|, |Im(z)| \leq |z|$. But, we're trying to study abstract NLSs, so we must face the difficulty.

Lemma 1.3.12. *Coordinate change for component functions.*

Suppose $F : dom(F) \subseteq V \to W$ is a map on NLS where $dim(W) = m$ and W has bases $\overline{\gamma} = \{\overline{w}_1, \dots, \overline{w}_m\}$ and $\gamma = \{w_1, \dots, w_m\}$. Let $P_{ij} \in \mathbb{R}$ be such that $\overline{w}_i = \sum_{j=1}^m P_{ij}w_j$. If $F_1, \dots, F_m$ are the component functions of F with respect to γ and $\overline{F_1}, \dots, \overline{F}_m$ are the component functions of F with respect to $\overline{\gamma}$ then $F_j = \sum_{i=1}^m \overline{F}_i P_{ij}$ for $j = 1, \dots, m$.

Proof: Since $\gamma, \overline{\gamma}$ are given bases of W we know there exist $P_{ij} \in \mathbb{R}$ such that $\overline{w}_i = \sum_{j=1}^m P_{ij}w_j$. Therefore, we can relate the component expansions in both bases as follows:

$$F = \sum_{i=1}^m \overline{F}_i\overline{w}_i = \sum_{j=1}^m F_jw_j \quad\Rightarrow\quad \sum_{i=1}^m \overline{F}_i \sum_{j=1}^m P_{ij}w_j = \sum_{j=1}^m F_jw_j \tag{1.18}$$

thus

$$\sum_{j=1}^m \left(\sum_{i=1}^m \overline{F}_iP_{ij}\right) w_j = \sum_{j=1}^m F_jw_j \quad\Rightarrow\quad \sum_{i=1}^m \overline{F}_iP_{ij} = F_j \tag{1.19}$$

where we equated coefficients of w_j to obtain the result above. □

It always amuses me to see how the basis and components transform inversely. Continuing to use the notation of the previous Theorem and Lemma,

Proposition 1.3.13.

If $\lim_{x\to a} \overline{F}(x) = \overline{L}_i$ for $i = 1, \dots, m$ then $\lim_{x\to a} F(x) = \sum_{i,j=1}^{m} P_{ij}\overline{L}_i$.

Proof: use Lemma 1.3.12 to see $F_i(x) = \sum_{j=1}^{m} P_{ij}\overline{F}_j(x)$. Then, by linearity of the limit,

$$\lim_{x\to a}(F_i(x)) = \sum_{j=1}^{m} P_{ij}\lim_{x\to a}\left(\overline{F}_j(x)\right) = \sum_{j=1}^{m} P_{ij}\overline{L}_j. \tag{1.20}$$

The Proposition follows by application of Theorem 1.3.11. □

The coordinate change results above are most interesting when paired with an additional freedom to analyze limits in finite dimensional vector spaces.

(1.) The metric topology for a finite dimensional normed linear space is independent of our choice of norm[7]. For example, in $\mathbb{R}^2$, if we find a point is interior with respect the euclidean norm then it's easy to see the point is also interior w.r.t. the taxicab or sup norm. I might assign a homework which helps you prove this claim.

(2.) Given normed linear spaces V, W and a function $F : dom(F) \subseteq V \to W$, we find F is continuous if and only if the inverse image under F of each open set in W is open in V.[8]

(3.) Since different choices of norm provide the same open sets it follows that the calculation of a limit in a finite dimensional NLS is in fact independent of the choice of norm.

Given any basis for finite dimensional real vector space we can construct an inner product by essentially mimicking the dot-product.

Lemma 1.3.14. *existence of inner product which makes given basis orthonormal.*

If $(V, \|\ \|)$ is a normed linear space with basis $\beta = \{v_1, \dots, v_n\}$ then $\langle v_i, v_j\rangle = \delta_{ij}$ extended bilinearly serves to define an inner product for V where β is an orthonormal basis. Furthermore, if $\|x\|_2 = \sqrt{\langle x, x\rangle}$ and $x = x_1v_1 + \cdots + x_nv_n$ then

$$\|x\|_2 = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$$

hence $|x_i| \leq \|x\|_2$ for any $x \in V$ and for each $i = 1, 2, \dots, n$.

Proof: left to reader, essentially the claim is immediate once we show $\langle x, y\rangle = x_1y_1 + \cdots + x_ny_n$ where x_i, y_i are the coordinates of x, y with respect to β basis. □

[7] see this question and the answers for some interesting discussion of this point

[8] Notice, we insist that $\emptyset$ is open, my apologies if my earlier wording was insufficiently clear on this point.

Theorem 1.3.15.

Let V, W be normed vector spaces and suppose W has basis $\beta = \{w_j\}_{j=1}^m$. Let $a \in V$ then

$$\lim_{x \to a} F(x) = B = \sum_{j=1}^m B_j w_j \qquad \Leftrightarrow \qquad \lim_{x \to a} F_j(x) = B_j \ \text{ for all } j = 1, 2, \dots m.$$

Proof: Suppose $\lim_{x \to a} F(x) = B \in W$. Construct the inner product $\langle , \rangle : W \times W \to \mathbb{R}$ which forces orthonormality of $\beta = \{w_1, \dots, w_m\}$. That is, let $\langle w_i, w_j \rangle = \delta_{ij}$ and extend bilinearly. Let $\|y\| = \sqrt{\langle y, y \rangle}$ hence $y = y_1 w_1 + \cdots + y_m w_m$ has $\|y\| = \sqrt{y_1^2 + \cdots + y_n^2}$ thus $\|w_i\| = 1$ and $|y_i| \leq \|y\|$ for each $y \in W$ and $i = 1, \dots, m$. Therefore,

$$|F_i(x) - B_i| = |F_i(x) - B_i| \|w_i\| = \|(F_i(x) - B_i) w_i\| \leq \|F(x) - B\|. \tag{1.21}$$

Hence, for $\epsilon > 0$ choose $\delta > 0$ such that $0 < \|x - a\| < \delta$ implies $\|F(x) - B\| < \epsilon$. Hence, by Inequality 1.21 we find $0 < \|x - a\| < \delta$ implies $|F_i(x) - B_i| < \epsilon$ for each $i = 1, 2, \dots, m$. Thus $\lim_{x \to a} F_j(x) = B_j$ for each $j = 1, \dots, m$ and this remains true when a different norm is given to W (here I use the result that the limit calculated in a finite dimensional NLS is independent of our choice of norm since all norms produce the same topology).

The converse direction follows from Theorem 1.3.11, but I include argument below since it's good to see. Conversely, suppose $\lim_{x \to a} F_j(x) = B_j$ as $x \to a$ for all $j \in \mathbb{N}_n$. Let $\epsilon > 0$ and choose $\delta_j > 0$ such that $0 < ||x - a|| < \delta_j$ implies $||F_j(x) - B_j|| < \frac{\epsilon}{||w_j|| m}$. We are free to choose such δ_j by the given limits as clearly $\frac{\epsilon}{||w_j|| m} > 0$ for each j. Choose $\delta = min(\delta_j \mid j \in \mathbb{N}_m\}$ and suppose $x \in V$ such that $0 < ||x - a|| < \delta$. Using properties $||x + y|| \leq ||x|| + ||y||$ and $||cx|| = |c| ||x||$ multiple times yield:

$$||F(x) - B|| = ||\sum_{j=1}^m (F_j(x) - B_j) w_j|| \leq \sum_{j=1}^m |F_j(x) - B_j| ||w_j|| < \sum_{j=1}^m \frac{\epsilon}{||w_j|| m} ||w_j|| = \sum_{j=1}^m \frac{\epsilon}{m} = \epsilon.$$

Therefore, $\lim_{x \to a} F(x) = B$ and this completes the proof $\square$.

Our next goal is to explain why polynomials in coordinates of an NLS are continuous. Many examples fall into this general category so it's worth the effort. The first result we need is the observation that we are free to pull limits out of continuous functions on an NLS:

Proposition 1.3.16. *Limit of composite functions.*

Suppose V_1, V_2, V_3 are normed vector spaces with norms $|| \cdot ||_1, || \cdot ||_2, || \cdot ||_3$ respective. Let $f : dom(f) \subseteq V_2 \to V_3$ and $g : dom(g) \subseteq V_1 \to V_2$ be mappings. Suppose that $\lim_{x \to x_o} g(x) = y_o$ and suppose that f is continuous at y_o then

$$\lim_{x \to x_o} (f \circ g)(x) = f\left(\lim_{x \to x_o} g(x) \right).$$

Proof: Let $\epsilon > 0$ and choose $\beta > 0$ such that $0 < ||y - b||_2 < \beta$ implies $||f(y) - f(y_o)||_3 < \epsilon$. We can choose such a β since Since f is continuous at y_o thus it follows that $\lim_{y \to y_o} f(y) = f(y_o)$. Next choose $\delta > 0$ such that $0 < ||x - x_o||_1 < \delta$ implies $||g(x) - y_o||_2 < \beta$. We can choose such a δ because we are given that $\lim_{x \to x_o} g(x) = y_o$. Suppose $0 < ||x - x_o||_1 < \delta$ and let $y = g(x)$

note $||g(x) - y_o||_2 < \beta$ yields $||y - y_o||_2 < \beta$ and consequently $||f(y) - f(y_o)||_3 < \epsilon$. Therefore, $0 < ||x - x_o||_1 < \delta$ implies $||f(g(x)) - f(y_o)||_3 < \epsilon$. It follows that $\lim_{x \to x_o}(f(g(x)) = f(\lim_{x \to x_o} g(x))$. □

The following functions are suprisingly useful as we seek to describe continuity of functions.

Definition 1.3.17.

The **sum** and **product** are functions from $\mathbb{R}^2$ to $\mathbb{R}$ defined by

$$s(x,y) = x + y \qquad p(x,y) = xy$$

Proposition 1.3.18.

The sum and product functions are continuous.

Proof: I leave to the reader. □

The proof that the product is continuous is not entirely trivial, but, once you have it, so many things follow:

Proposition 1.3.19.

Let V be an NLS. If $f : dom(f) \subseteq V \to \mathbb{R}$ and $g : dom(g) \subseteq V \to \mathbb{R}$ and $\lim_{x \to a} f(x), \lim_{x \to a} g(x) \in \mathbb{R}$ then $\lim_{x \to a}(f(x) \cdot g(x)) = \lim_{x \to a} f(x) \cdot \lim_{x \to a} g(x)$.

Proof: Combining Propositions 1.3.19 and 1.3.16 we find

$$\begin{aligned} \lim_{x \to a}(f(x) \cdot g(x)) &= \lim_{x \to a}(p(f(x), g(x))) \\ &= p\left(\lim_{x \to a}(f(x), g(x))\right) \\ &= p\left(\lim_{x \to a} f(x), \lim_{x \to a} g(x)\right) \\ &= \lim_{x \to a} f(x) \cdot \lim_{x \to a} g(x). \quad \square \end{aligned} \tag{1.22}$$

Of course, we can continue to products of three or more factors by iterating the product:

$$fgh = p(fg, h) = p(p(f,g), h) \tag{1.23}$$

and by an argument much like that given in Equation 1.22 we can argue that the product of three continous real-valued functions on a subset of a NLS V is once more continuous. It should be clear we can extend by induction this result to any product of finitely many real-valued continuous functions.

Lemma 1.3.20.

Let V be a NLS with basis $\{v_1, \dots, v_n\}$. Define coordinate function $x_i : V \to \mathbb{R}$ as follows: given $a = a_1 v_1 + \cdots + a_n v_n$ set $x_i(a) = a_i$. Then $\Phi_\beta = (x_1, x_2, \dots, x_n)$ and each coordinate function is continuous on V.

Proof: if $a = a_1 v_1 + \cdots + a_n v_n$ then $\Phi_\beta(a) = (a_1, \dots, a_n) = (x_1(a), \dots, x_n(a))$ therefore $\Phi_\beta = (x_1, \dots, x_n)$. I leave the proof that $x_i : V \to \mathbb{R}$ is continuous for each $i = 1, \dots, m$ as a likely homework for the reader. □

Definition 1.3.21.

Let $x_1, \dots, x_n$ be coordinate functions with respect to basis β for a NLS V then a function $f : V \to \mathbb{R}$ such that for constants $c_0, c_i, c_{ij}, \dots, c_{i_1,\dots,i_k} \in \mathbb{R}$,

$$f(x) = c_0 + \sum_{i=1}^{n} c_i x_i + \sum_{i,j=1}^{n} c_{ij} x_i x_j + \cdots + \sum_{i_1,\dots,i_k} c_{i_1,\dots,i_k} x_{i_1} \cdots x_{i_k}$$

is a k-th order multinomial in $x_1, \dots, x_n$. We say $f(x) \in \mathbb{R}[x_1, \dots, x_n]$.

The following Theorem is a clear consequence of the results we've thus far discussed in this Section:

Theorem 1.3.22.

Multinomials in the coordinates of a NLS V form continuous real-valued functions on V.

Example 1.3.23. *Define* $det : \mathbb{R}^{\,n\times n} \to \mathbb{R}$ *by*

$$det(A) = \sum_{i_1,\dots,i_n=1}^{n} \epsilon_{i_1,\dots,i_n} A_{1i_1} \cdots A_{ni_n}$$

hence $det(A) \in \mathbb{R}[A_{ij} \mid 1 \leq i,j \leq n]$ *is an* n*-th order multinomial in the coordinates* A_{ij} *with respect to the standard matrix basis for* $\mathbb{R}^{\,n\times n}$*. Thus the determinant is a continuous real-valued function of matrices.*

I'll let you explain why the complex-valued determinant function on $\mathbb{C}^{n\times n}$ is also continuous. Let's enjoy the application of these results:

Example 1.3.24. *The general linear group* $GL(n, \mathbb{R}) = \{A \in \mathbb{R}^{\,n\times n} \mid det(A) \neq 0\}$ *is an* **open** *subset of* $\mathbb{R}^{n\times n}$*. To see this notice that* $GL(n, \mathbb{R}) = det^{-1}((-\infty, 0) \cup (0, \infty))$*. But, the determinant is continuous and the inverse image of open sets is open. Clearly* $(-\infty, 0) \cup (0, \infty)$ *is open since each point is interior.*

To be picky, I have not shown the inverse image of open sets is open for a continuous map on an NLS, but, I will likely assign that as a homework, so, don't worry, you'll get a chance to ponder it.

Example 1.3.25. *Let* $T : \mathbb{R}^n \to \mathbb{R}^m$ *be a linear transformation. Then* $T(x)$ *has component functions which are formed from first order multinomials in* $x_1, \dots, x_n$*. Thus* T *is continuous on* $\mathbb{R}^n$*. It is likely I'll ask you to prove* T *is continuous by direct application of the definition of the limit. It's a good problem to work through.*

The squeeze theorem relies heavily on the order properties of $\mathbb{R}$. Generally a normed vector space has no natural ordering. For example, is $1 > i$ or is $1 < i$ in $\mathbb{C}$? That said, we can state a squeeze theorem for real-valued functions whose domain reside in a normed vector space. This is a generalization of what we learned in calculus I. That said, the proof offered below is very similar to the typical proof which is not given in calculus I[9]

[9] this is lifted word for word from my calculus I notes, however here the meaning of open ball is considerably more general and the linearity of the limit which is referenced is the one proven earlier in this section

Theorem 1.3.26. *Squeeze Theorem.*

Suppose $f : dom(f) \subseteq V \to \mathbb{R}$, $g : dom(g) \subseteq V \to \mathbb{R}$, $h : dom(h) \subseteq V \to \mathbb{R}$ where V is a normed vector space with norm $|| \cdot ||$. Let $f(x) \leq g(x) \leq h(x)$ for all x on some $\delta > 0$ ball of $a \in V$ and suppose the limits of $f(x), g(x), h(x)$ all exist at limit point a then

$$\lim_{x \to a} f(x) \leq \lim_{x \to a} g(x) \leq \lim_{x \to a} h(x).$$

Furthermore, if the limits of $f(x)$ and $h(x)$ exist with $\lim_{x\to a} f(x) = \lim_{x\to a} h(x) = L \in \mathbb{R}$ then the limit of $g(x)$ likewise exists and $\lim_{x\to a} g(x) = L$.

Proof: Suppose $f(x) \leq g(x)$ for all[10] $x \in B_{\delta_1}(a)_o$ for some $\delta_1 > 0$ and also suppose $\lim_{x\to a} f(x) = L_f \in \mathbb{R}$ and $\lim_{x\to a} g(x) = L_g \in \mathbb{R}$. We wish to prove that $L_f \leq L_g$. Suppose otherwise towards a contradiction. That is, suppose $L_f > L_g$. Note that $\lim_{x\to a}[g(x) - f(x)] = L_g - L_f$ by the linearity of the limit. It follows that for $\epsilon = \frac{1}{2}(L_f - L_g) > 0$ there exists $\delta_2 > 0$ such that $x \in B_{\delta_2}(a)_o$ implies $|g(x) - f(x) - (L_g - L_f)| < \epsilon = \frac{1}{2}(L_f - L_g)$. Expanding this inequality we have

$$-\frac{1}{2}(L_f - L_g) < g(x) - f(x) - (L_g - L_f) < \frac{1}{2}(L_f - L_g)$$

adding $L_g - L_f$ yields,

$$-\frac{3}{2}(L_f - L_g) < g(x) - f(x) < -\frac{1}{2}(L_f - L_g) < 0.$$

Thus, $f(x) > g(x)$ for all $x \in B_{\delta_2}(a)_o$. But, $f(x) \leq g(x)$ for all $x \in B_{\delta_1}(a)_o$ so we find a contradiction for each $x \in B_\delta(a)$ where $\delta = min(\delta_1, \delta_2)$. Hence $L_f \leq L_g$. The same proof can be applied to g and h thus the first part of the theorem follows.

Next, we suppose that $\lim_{x\to a} f(x) = \lim_{x\to a} h(x) = L \in \mathbb{R}$ and $f(x) \leq g(x) \leq h(x)$ for all $x \in B_{\delta_1}(a)$ for some $\delta_1 > 0$. We seek to show that $\lim_{x\to a} f(x) = L$. Let $\epsilon > 0$ and choose $\delta_2 > 0$ such that $|f(x) - L| < \epsilon$ and $|h(x) - L| < \epsilon$ for all $x \in B_\delta(a)_o$. We are free to choose such a $\delta_2 > 0$ because the limits of f and h are given at $x = a$. Choose $\delta = min(\delta_1, \delta_2)$ and note that if $x \in B_\delta(a)_o$ then

$$f(x) \leq g(x) \leq h(x)$$

hence,

$$f(x) - L \leq g(x) - L \leq h(x) - L$$

but $|f(x) - L| < \epsilon$ and $|h(x) - L| < \epsilon$ imply $-\epsilon < f(x) - L$ and $h(x) - L < \epsilon$ thus

$$-\epsilon < f(x) - L \leq g(x) - L \leq h(x) - L < \epsilon.$$

Therefore, for each $\epsilon > 0$ there exists $\delta > 0$ such that $x \in B_\delta(a)_o$ implies $|g(x) - L| < \epsilon$ so $\lim_{x\to a} g(x) = L$. □

Our typical use of the theorem above applies to equations of norms from a normed vector space. The norm takes us from V to $\mathbb{R}$ so the theorem above is essential to analyze interesting limits. We shall make use of it in future analysis.

[10] I use the notation $B_{\delta_1}(a)_o$ to denote the deleted open ball of radius δ_1 centered at a; $B_{\delta_1}(a)_o = B_{\delta_1}(a) - \{a\}$.

1.4 sequential analysis

Let $(V, ||\cdot||_V)$ be a normed vector space, a function from $\mathbb{N}$ to V is a called a **sequence**. Limits of sequences play an important role in analysis in normed linear spaces. The real analysis course makes great use of sequences to tackle questions which are more difficult with only $\epsilon - \delta$ arguments. In fact, we can reformulate limits in terms of sequences and subsequences. Perhaps one interesting feature of abstract topological spaces is the appearance of spaces in which sequential convergence is insufficient to capture the concept of limits. In general, one needs *nets* and *filters*. I digress. More important to our context, the criteria of **completeness**. Let us settle a few definitions to make the words meaningful.

Definition 1.4.1.

Suppose $\{a_n\}$ is a sequence then we say $\lim_{n\to\infty} a_n = L \in V$ iff for each $\epsilon > 0$ there exists $M \in \mathbb{N}$ such that $||a_n - L||_V < \epsilon$ for all $n \in \mathbb{N}$ with $n > M$. If $\lim_{n\to\infty} a_n = L \in V$ then we say $\{a_n\}$ is a **convergent sequence**.

We spent some effort attempting to understand the definition above and its application to the problem of infinite summations in calculus II. It is less likely you have thought much about the following:

Definition 1.4.2.

We say $\{a_n\}$ is a **Cauchy sequence** iff for each $\epsilon > 0$ there exists $M \in \mathbb{N}$ such that $||a_m - a_n||_V < \epsilon$ for all $m, n \in \mathbb{N}$ with $m, n > M$.

In other words, a sequence is Cauchy if the terms in the sequence get arbitarily close as we go sufficiently far out in the list. Many concepts we cover in calculus II are made clear with proofs built around the concept of a Cauchy sequence. The interesting thing about Cauchy is that for some spaces of numbers we can have a sequence which converges but is not Cauchy. For example, if you think about the rational numbers $\mathbb{Q}$ we can construct a sequence of truncated decimal expansions of π:

$$\{a_n\} = \{3, 3.1, 3.14, 3.141, 3.1415\dots\}$$

note that $a_n \in \mathbb{Q}$ for all $n \in \mathbb{N}$ and yet the $a_n \to \pi \notin \mathbb{Q}$. When spaces are missing their limit points they are in some sense incomplete.

Definition 1.4.3.

If every Cauchy sequence in a metric space converges to a point within the space then we say the metric space is **complete**. If a normed vector space V is complete then we say V is a **Banach space**.

A metric space need not be a vector space. In fact, we can take any open set of a normed vector space and construct a metric space. Metric spaces require less structure.

Fortunately all the main examples of this course are built on the real numbers which are complete, this induces completeness for $\mathbb{C}, \mathbb{R}^n$ and $\mathbb{R}^{m\times n}$. The proof that $\mathbb{R}$, $\mathbb{C}$, $\mathbb{R}^n$ and $\mathbb{R}^{m\times n}$ are Banach spaces follow from arguments similar to those given in the example below.

Example 1.4.4. Claim: $\mathbb{R}$ *complete implies* $\mathbb{R}^2$ *is complete.*

Proof: *suppose (x_n, y_n) is a Cauchy sequence in $\mathbb{R}^2$. Therefore, for each $\epsilon > 0$ there exists $N \in \mathbb{N}$ such that $m, n \in \mathbb{N}$ with $N < m < n$ implies $||(x_m, y_m) - (x_n, y_n)|| < \epsilon$. Consider that:*

$$||(x_m, y_m) - (x_n, y_n)|| = \sqrt{(x_m - x_n)^2 + (y_m - y_n)^2}$$

Therefore, as $|x_m - x_n| = \sqrt{(x_m - x_n)^2}$, it is clear that:

$$|x_m - x_n| \leq ||(x_m, y_m) - (x_n, y_n)||$$

But, this proves that $\{x_n\}$ is a Cauchy sequence of real numbers since for each $\epsilon > 0$ we can choose $N > 0$ such that $N < m < n$ implies $|x_m - x_n| < \epsilon$. The same holds true for the sequence $\{y_n\}$. By completeness of $\mathbb{R}$ we have $x_n \to x$ and $y_n \to y$ as $n \to \infty$. We propose that $(x_n, y_n) \to (x, y)$. Let $\epsilon > 0$ once more and choose $N_x > 0$ such that $n > N_x$ implies $|x_n - x| < \epsilon/2$ and $N_y > 0$ such that $n > N_y$ implies $|y_n - y| < \epsilon/2$. Let $N = max(N_x, N_y)$ and suppose $n > N$:

$$||(x_n, y_n) - (x, y)|| = ||x_n - x, 0) + (0, y_n - y)|| \leq |x_n - x| + |y_n - y| < \epsilon/2 + \epsilon/2 = \epsilon.$$

The key point here is that components of a Cauchy sequence form Cauchy sequences in $\mathbb{R}$. That will also be true for sets of matrices and complex numbers.

Finally, I close with an appliction to the matrix exponential. We define:

$$e^A = I + A + \frac{1}{2}A^2 + \frac{1}{3!}A^3 + \cdots = \sum_{k=0}^{\infty} \frac{1}{k!}A^k. \tag{1.24}$$

for such $A \in \mathbb{R}^{\ n \times n}$ as the series above converges. Convergence of a series of matrices is measured by the convergence of the sequence of partial sums. For e^A the n-th partial sum is simply:

$$S_n = \sum_{k=0}^{n-1} \frac{1}{k!}A^k = I + A + \cdots + \frac{1}{(n-1)!}A^{n-1} \tag{1.25}$$

Thus, assuming $m > n$,

$$S_m - S_n = \sum_{k=n}^{m-1} \frac{1}{k!}A^k = \frac{1}{n!}A^n + \cdots + \frac{1}{(m-1)!}A^{m-1} \tag{1.26}$$

The identity $\|AB\| \leq \|A\|\|B\|$ inductively extends to $\|A^k\| \leq \|A\|^k$ for $k \in \mathbb{N}$. With this identity and the triangle inequality we find:

$$\|S_m - S_n\| \leq \sum_{k=n}^{m-1} \frac{1}{k!}\|A\|^k = s_m - s_n \tag{1.27}$$

where $s_n = \sum_{k=0}^{n-1} \frac{1}{k!}\|a\|^k$ is the n-th partial sum of $\sum_{k=0}^{\infty} e^{\|A\|}$. Note s_n is convergence sequence in $\mathbb{R}$ hence it is Cauchy so as $m, n \to \infty$ we find $s_m - s_n \to 0$ and so by the squeeze theorem for sequences we deduce $\|S_m - S_n\| \to 0$ as $m, n \to \infty$. In other words, S_n forms a Cauchy sequence of matrices and thus by the completeness of $\mathbb{R}^{\ n \times n}$ we deduce the series defining the matrix exponential converges. Notice this argument holds for any matrix A.

I'm fond of the argument above, it was shown to me in some course I took with R.O Fulp, maybe a few courses. There is another argument from linear algebra which uses the real Jordan form. Since $A = P^{-1}JP$ for some $P \in GL(n, \mathbb{R})$ and e^J is easily calculated we obtain existence of e^A from the fact that $e^J = e^{PAP^{-1}} = Pe^AP^{-1}$. But, admittedly, it does take a little work to prove the existence of the real Jordan form for any $A \in \mathbb{R}^{\ n \times n}$. I bet there are many other arguments to show e^A is well-defined. The abstract concept of the exponential is much more useful than you might first expect. The past two summers I learned an exponential on the appropriate algebra solves any constant coefficient ODE, even when the coefficients are taken from algebras with all sorts of weird features.

Chapter 2

differentiation

Our goal in this chapter is to describe differentiation for functions to and from normed linear spaces. It turns out this is actually quite simple given the background of the preceding chapter. The differential at a point is a linear transformation which best approximates the change in a function at a particular point. We can quantify "best" by a limiting process which is naturally defined in view of the fact there is a norm on the spaces we consider.

The most important example is of course the case $f : \mathbb{R}^n \to \mathbb{R}^m$. In this context it is natural to write the differential as a matrix multiplication. The matrix of the differential is known as the Jacobian matrix. Partial derivatives are also defined in terms of directional derivatives. The directional derivative is sometimes defined where the differential fails to exist. We will discuss how the criteria of continuous differentiability allows us to build the differential from the directional derivatives. We study how the general concept of Frechet differentiation recovers all the derivatives you've seen previously in calculus and much more.

The general theory of differentiation is a bit of an adjustment from our previous experience differentiating. Dieudonne said it best: this is the introduction to his chapter on differentiation in *Modern Analysis* Chapter VIII.

> The subject matter of this Chapter is nothing else but the elementary theorems of Calculus, which however are presented in a way which will probably be new to most students. That presentation, which throughout adheres strictly to our general "geometric" outlook on Analysis, aims at keeping as close as possible to the fundamental idea of Calculus, namely the "local" approximation of functions by *linear* functions. **In the classical teaching of Calculus, the idea is immediately obscured by the accidental fact that, on a one-dimensional vector space, there is a one-to-one correspondence between *linear* forms and numbers, and therefore the derivative at a point is defined as a *number* instead of a *linear form*. This slavish subservience to the shibboleth[1] of numerical interpretation at any cost becomes much worse when dealing with functions of several variables...**

Dieudonne's then spends the next half page continuing this thought with explicit examples of how this custom of our calculus presentation injures the conceptual generalization. If you want to see

[1] from wikipedia: is a word, sound, or custom that a person unfamiliar with its significance may not pronounce or perform correctly relative to those who are familiar with it. It is used to identify foreigners or those who do not belong to a particular class or group of people. It also refers to features of language, and particularly to a word or phrase whose pronunciation identifies a speaker as belonging to a particular group.

differentiation written for mathematicians, that is the place to look. He proves many results for infinite dimensions because, well, why not?

In this chapter I define the Frechet differential and exhibit a number of abstract examples. Then we turn to proving the basic properties of the Frechet derivative including linearity and the chain rule. My proof of the chain rule has a bit of a gap, but, I hope the argument gives you some intuition as to why we should expect a chain rule. Next we explore partial derivatives in an NLS with respect to a given abstract basis. After that we focus on $\mathbb{R}^n$. Many many examples of Jacobians are given. We study a few perverse examples which fail to be continuously differentiable. We show continuous differentiability implies differentiability by a standard, but interesting, argument. I prove a quite general product rule, discuss the problem of higher derivatives in the abstract (I punt details to Zorich for now, sorry). Finally, I share some insights I've recently come to to understand about $\mathcal{A}$-Calculus. In particular, I discuss some of the rudiments of differentiating with respect to algebra variables.

2.1 the Frechet differential

The definition[2] below says that $\triangle F = F(a+h) - F(a) \cong dF_a(h)$ when h is close to zero.

Definition 2.1.1.

Let $(V, ||\cdot||_V)$ and $(W, ||\cdot||_W)$ be normed vector spaces. Suppose that U is open and $F : U \subseteq V \rightarrow W$ is a function the we say that F is **differentiable** at $a \in U$ iff there exists a linear mapping $L : V \rightarrow W$ such that

$$\lim_{h\rightarrow 0}\left[\frac{F(a+h) - F(a) - L(h)}{||h||_V}\right] = 0.$$

In such a case we call the linear mapping L the **differential at** a and we denote $L = dF_a$. In the case $V = \mathbb{R}^m$ and $W = \mathbb{R}^n$ the matrix of the differential is called the **derivative of** F **at** a or the **Jacobian matrix of** F **at** a and we denote $[dF_a] = F'(a) \in \mathbb{R}^{m\times n}$ which means that $dF_a(v) = F'(a)v$ for all $v \in \mathbb{R}^n$.

Notice this definition gives an equation which implicitly defines dF_a. For the moment the only way we have to calculate dF_a is educated guessing. We simply use brute-force calculation to suggest a guess for L which forces the Frechet quotient to vanish. In the next section we'll discover a systematic calculational method for functions on euclidean spaces. The purpose of this section is to understand the definition of the differential and to connect it to basic calculus. I'll begin with basic calculus as you probably are itching to understand where your beloved difference quotient has gone:

[2] Some authors might put a norm in the numerator of the quotient. That is an equivalent condition since a function $g : V \rightarrow W$ has $\lim_{h\rightarrow 0} g(h) = 0$ iff $\lim_{h\rightarrow 0} ||g(h)||_W = 0$

Example 2.1.2. *Suppose $f : dom(f) \subseteq \mathbb{R} \to \mathbb{R}$ is differentiable at x. It follows that there exists a linear function $df_x : \mathbb{R} \to \mathbb{R}$ such that*[3]

$$\lim_{h\to 0} \frac{f(x+h) - f(x) - df_x(h)}{|h|} = 0.$$

Note that

$$\lim_{h\to 0} \frac{f(x+h) - f(x) - df_x(h)}{|h|} = 0 \quad \Leftrightarrow \quad \lim_{h\to 0^\pm} \frac{f(x+h) - f(x) - df_x(h)}{|h|} = 0.$$

In the left limit $h \to 0^-$ we have $h < 0$ hence $|h| = -h$. On the other hand, in the right limit $h \to 0^+$ we have $h > 0$ hence $|h| = h$. Thus, differentiability suggests that $\lim_{h\to 0^\pm} \frac{f(x+h)-f(x)-df_x(h)}{\pm h} = 0$. But we can pull the minus out of the left limit to obtain $\lim_{h\to 0^-} \frac{f(x+h)-f(x)-df_x(h)}{h} = 0$. Therefore, after an algebra step, we find:

$$\lim_{h\to 0} \left[\frac{f(x+h) - f(x)}{h} - \frac{df_x(h)}{h} \right] = 0.$$

Linearity of $df_x : \mathbb{R} \to \mathbb{R}$ implies there exists $m \in \mathbb{R}^{1\times 1} = \mathbb{R}$ such that $df_x(h) = mh$. Observe that

$$\lim_{h\to 0} \frac{df_x(h)}{h} = \lim_{h\to 0} \frac{mh}{h} = m.$$

It is a simple exercise to show that if $\lim(A - B) = 0$ and $\lim(B)$ exists then $\lim(A)$ exists and $\lim(A) = \lim(B)$. Identify $A = \frac{f(x+h)-f(x)}{h}$ and $B = \frac{df_x(h)}{h}$. Therefore,

$$m = \lim_{h\to 0} \frac{f(x+h) - f(x)}{h}.$$

Consequently, we find the 1×1 matrix m of the differential is precisely $f'(x)$ as we defined it via a difference quotient in first semester calculus. In summary, we find $\boxed{df_x(h) = f'(x)h}$. In other words, if a function is differentiable in the sense we defined at the beginning of this chapter then it is differentiable in the terminology we used in calculus I. Moreover, the derivative at x is precisely the matrix of the differential.

Remark 2.1.3.

Incidentally, I should mention that df_x is the differential of f at the point x. The differential of f would be the mapping $x \mapsto df_x$. Technically, the differential df is a function from $\mathbb{R}$ to the set of linear transformations on $\mathbb{R}$. You can contrast this view with that of first semester calculus. There we say the mapping $x \mapsto f'(x)$ defines the derivative f' as a function from $\mathbb{R}$ to $\mathbb{R}$. This simplification in perspective is only possible because calculus in one-dimension is so special. More on this later. This distinction is especially important to understand if you begin to look at questions of higher derivatives.

Example 2.1.4. *Suppose $T : V \to W$ is a linear transformation of normed vector spaces V and W. I propose $L = T$. In other words, I think we can show the best linear approximation to the change in a linear function is simply the function itself. Clearly L is linear since T is linear. Consider the difference quotient:*

$$\frac{T(a+h) - T(a) - L(h)}{||h||_V} = \frac{T(a) + T(h) - T(a) - T(h)}{||h||_V} = \frac{0}{||h||_V}.$$

Note $h \neq 0$ implies $||h||_V \neq 0$ by the definition of the norm. Hence the limit of the difference quotient vanishes since it is identically zero for every nonzero value of h. We conclude that $dT_a = T$.

[3] unless we state otherwise, $\mathbb{R}^n$ is assumed to have the euclidean norm, in this case $||x||_{\mathbb{R}} = \sqrt{x^2} = |x|$.

Example 2.1.5. *Let $T : V \to W$ where V and W are normed vector spaces and define $T(v) = w_o$ for all $v \in V$. I claim the differential is the zero transformation. Linearity of $L(v) = 0$ is trivially verified. Consider the difference quotient:*

$$\frac{T(a+h) - T(a) - L(h)}{||h||_V} = \frac{w_o - w_o - 0}{||h||_V} = \frac{0}{||h||_V}.$$

Using the arguments to the preceding example, we find $dT_a = 0$.

Typically the difference quotient is not identically zero. The pair of examples above are very special cases. Our next example requires a bit more thought:

Example 2.1.6. *Suppose $F : \mathbb{R}^2 \to \mathbb{R}^3$ is defined by $F(x, y) = (xy,\ x^2,\ x + 3y)$ for all $(x, y) \in \mathbb{R}^2$. Consider the difference function $\triangle F$ at (x, y):*

$$\triangle F = F((x, y) + (h, k)) - F(x, y) = F(x + h, y + k) - F(x, y)$$

Calculate,

$$\triangle F = \big((x + h)(y + k),\ (x + h)^2,\ x + h + 3(y + k)\big) - \big(xy,\ x^2,\ x + 3y\big)$$

Simplify by cancelling terms which cancel with $F(x, y)$:

$$\triangle F = \big(xk + hy + hk,\ 2xh + h^2,\ h + 3k)\big)$$

Identify the linear part of $\triangle F$ as a good candidate for the differential. I claim that:

$$L(h, k) = \big(xk + hy,\ 2xh,\ h + 3k\big).$$

is the differential for f at (x,y). Observe first that we can write

$$L(h, k) = \begin{bmatrix} y & x \\ 2x & 0 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} h \\ k \end{bmatrix}.$$

therefore $L : \mathbb{R}^2 \to \mathbb{R}^3$ is manifestly linear. Use the algebra above to simplify the difference quotient below:

$$\lim_{(h,k)\to(0,0)} \left[\frac{\triangle F - L(h, k)}{||(h, k)||} \right] = \lim_{(h,k)\to(0,0)} \left[\frac{(hk, h^2, 0)}{||(h, k)||} \right]$$

Note $||(h, k)|| = \sqrt{h^2 + k^2}$ therefore we fact the task of showing that $\frac{1}{\sqrt{h^2+k^2}}(hk,\ h^2,\ 0) \to (0, 0, 0)$ as $(h, k) \to (0, 0)$. Notice that: $||(hk,\ h^2,\ 0)|| = |h|\sqrt{h^2 + k^2}$. Therefore, as $(h, k) \to 0$ we find

$$\left|\left| \frac{1}{\sqrt{h^2 + k^2}}(hk,\ h^2,\ 0) \right|\right| = |h| \to 0.$$

However, if $||v|| \to 0$ it follows $v \to 0$ so we derive the desired limit. Therefore,

$$df_{(x,y)}(h, k) = \begin{bmatrix} y & x \\ 2x & 0 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} h \\ k \end{bmatrix}.$$

Computation of less trivial multivariate limits is an art we'd like to avoid if possible. It turns out that we can actually avoid these calculations by computing partial derivatives. However, we still need a certain multivariate limit to exist for the partial derivative functions so in some sense it's unavoidable. The limits are there whether we like to calculate them or not.

Definition 2.1.7. *Linearization of a differentiable map.*

Let $(V, ||\cdot||_V)$ and $(W, ||\cdot||_W)$ be normed vector spaces and suppose $F : dom(F) \subseteq V \to W$ is differentiable at p then the **linearization of F at p** is given by $L_F^p(x) = F(p)+dF_p(x-p)$ for all $x \in V$. We also say $L_F^p : V \to W$ is the **affinization** of F at p.

Perhaps the term **linearization** is a holdover from the terminology *linear function* of the form $f(x) = mx+b$. Of course, this is an offense to the student of pure linear algebra. Unless $b = 0$ such a map is not technically **linear**. What is it? It's an affine function. So, I added the terminology **affinization of** F to the definition above. However, I must admit, I don't think that terminology is standard. Much can be said about affine maps of normed linear spaces, I probably fail to paint the big picture of affine maps in these notes. Maybe I should make it homework...

Example 2.1.8. *Suppose $F : \mathbb{R}^2 \to \mathbb{R}^3$ is defined by $F(x, y) = (xy,\ x^2,\ x + 3y)$ for all $(x, y) \in \mathbb{R}^2$ then calculate the linearization of f at $(1, -2)$. Following Example 2.1.6 we find*

$$df_{(x,y)}(h,k) = \begin{bmatrix} y & x \\ 2x & 0 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} h \\ k \end{bmatrix} \quad \Rightarrow \quad df_{(1,-2)}(h,k) = \begin{bmatrix} -2 & 1 \\ 2 & 0 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} h \\ k \end{bmatrix}.$$

The linearization of f at $(1, -2)$ is constructed as follows:

$$\begin{aligned} L_f^{(1,-2)}(x,y) &= f(1,-2) + df_{(1,-2)}(x-1, y+2) \qquad (2.1) \\ &= (-2, 1, -5) + \begin{bmatrix} -2 & 1 \\ 2 & 0 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} x-1 \\ y+2 \end{bmatrix} \\ &= (-2 - 2(x-1) + (y+2), 1 + 2(x-1), -5 + (x-1) + 3(y+2)) \\ &= (-2x + y + 2, 2x - 1, x + 3y). \end{aligned}$$

Calculation of the differential simplifies considerably when the domain is one-dimensional. We already worked out the case of $f : \mathbb{R} \to \mathbb{R}$ in Example 2.1.2 and the following pair of examples work out the concrete case of $F : \mathbb{R} \to \mathbb{C}$ and then the general case $F : \mathbb{R} \to V$ for an arbitrary finite dimensional normed linear space V.

Example 2.1.9. *Suppose $F(t) = U(t)+iV(t)$ for all $t \in dom(f)$ and both U and V are differentiable functions on $dom(F)$. By the arguments given in Example 2.1.2 it suffices to find $L : \mathbb{R} \to \mathbb{C}$ such that*

$$\lim_{h\to 0} \left[\frac{F(t+h) - F(t) - L(h)}{h} \right] = 0.$$

I propose that on the basis of analogy to Example 2.1.2 we ought to have $dF_t(h) = (U'(t)+iV'(t))h$. Let $L(h) = (U'(t) + iV'(t))h$. Observe that, using properties of $\mathbb{C}$:

$$\begin{aligned} L(h_1 + ch_2) &= (U'(t) + iV'(t))(h_1 + ch_2) \\ &= (U'(t) + iV'(t))h_1 + c(U'(t) + iV'(t))h_2 \\ &= L(h_1) + cL(h_2). \end{aligned}$$

for all $h_1, h_2 \in \mathbb{R}$ and $c \in \mathbb{R}$. Hence $L : \mathbb{R} \to \mathbb{C}$ is linear. Moreover,

$$\begin{aligned} \tfrac{F(t+h)-F(t)-L(h)}{h} &= \tfrac{1}{h}\Big(U(t+h) + iV(t+h) - U(t) - iV(t) - (U'(t) + iV'(t))h \Big) \\ &= \tfrac{1}{h}\Big(U(t+h) - U(t) - U'(t)h \Big) + i\tfrac{1}{h}\Big(V(t+h) - V(t) - V'(t)h \Big) \end{aligned}$$

Consider the problem of calculating $\lim_{h\to 0}\frac{F(t+h)-F(t)-L(h)}{h}$. *We observe that a complex function converges to zero iff the real and imaginary parts of the function separately converge to zero (this is covered by Theorem 1.3.22). By differentiability of* U *and* V *we find again using Example 2.1.2*

$$\lim_{h\to 0}\frac{1}{h}\Big(U(t+h)-U(t)-U'(t)h\Big)=0 \qquad \lim_{h\to 0}\frac{1}{h}\Big(V(t+h)-V(t)-V'(t)h\Big)=0.$$

Therefore, $dF_t(h)=(U'(t)+iV'(t))h$. *Note that the quantity* $U'(t)+iV'(t)$ *is not a real matrix in this case. To write the derivative in terms of a real matrix multiplication we need to construct some further notation which makes use of the isomorphism between* $\mathbb{C}$ *and* $\mathbb{R}^2$. *Actually, it's pretty easy if you agree that* $a+ib=(a,b)$ *then* $dF_t(h)=(U'(t),V'(t))h$ *so the matrix of the differential is* $(U'(t),V'(t))\in\mathbb{R}^{2\times 1}$ *which makes since as* $F:\mathbb{R}\to\mathbb{C}\approx\mathbb{R}^2$.

Example 2.1.10. *Suppose* V *is a normed vector space with basis* $\beta=\{f_1,f_2,\dots,f_n\}$. *Futhermore, let* $G:I\subseteq\mathbb{R}\to V$ *be defined by*

$$G(t)=\sum_{i=1}^{n}G_i(t)f_i$$

where $G_i:I\to\mathbb{R}$ *is differentiable on* I *for* $i=1,2,\dots,n$. *Recall Theorem 1.3.22 revealed that* $T=\sum_{j=1}^{n}T_jf_j:\mathbb{R}\to V$ *then* $\lim_{t\to 0}T(t)=\sum_{j=1}^{n}l_jf_j$ *iff* $\lim_{t\to 0}T_j(t)=l_j$ *for all* $j=1,2,\dots,n$. *In words, the limit of a vector-valued function can be parsed into a vector of limits. With this in mind consider (again we can trade* $|h|$ *for* h *as we explained in-depth in Example 2.1.2) the difference quotient* $\lim_{h\to 0}\left[\frac{G(t+h)-G(t)-h\sum_{i=1}^{n}\frac{dG_i}{dt}f_i}{h}\right]$, *factoring out the basis yields:*

$$\lim_{h\to 0}\left[\frac{\sum_{i=1}^{n}[G_i(t+h)-G_i(t)-h\frac{dG_i}{dt}]f_i}{h}\right]=\sum_{i=1}^{n}\left[\lim_{h\to 0}\frac{G_i(t+h)-G_i(t)-h\frac{dG_i}{dt}}{h}\right]f_i=0$$

where the zero above follows from the supposed differentiability of each component function. It follows that:

$$\boxed{dG_t(h)=h\sum_{i=1}^{n}\frac{dG_i}{dt}f_i}$$

The example above encompasses a number of cases at once:

(1.) $V=\mathbb{R}$, functions on $\mathbb{R}$, $f:\mathbb{R}\to\mathbb{R}$

(2.) $V=\mathbb{R}^n$, space curves in $\mathbb{R}$, $\vec{r}:\mathbb{R}\to\mathbb{R}^n$

(3.) $V=\mathbb{C}$, complex-valued functions of a real variable, $f=u+iv:\mathbb{R}\to\mathbb{C}$

(4.) $V=\mathbb{R}^{m\times n}$, matrix-valued functions of a real variable, $F:\mathbb{R}\to\mathbb{R}^{m\times n}$.

In short, when we differentiate a function which has a real domain then we can define the derivative of such a function by component-wise differentiation. It gets more interesting when the domain has several independent variables as Examples 2.1.6 and 2.1.11 illustrate.

Example 2.1.11. *Suppose $F : \mathbb{R}^{\,n\times n} \to \mathbb{R}^{\,n\times n}$ is defined by $F(A) = A^2$. Notice*

$$\triangle F = F(A+H) - F(A) = (A+H)(A+H) - A^2 = AH + HA + H^2$$

I propose that F is differentiable at A and $L(H) = AH + HA$. Let's check linearity,

$$L(H_1 + cH_2) = A(H_1 + cH_2) + (H_1 + cH_2)A = AH_1 + H_1A + c(AH_2 + H_2A)$$

Hence $L : \mathbb{R}^{\,n\times n} \to \mathbb{R}^{\,n\times n}$ is a linear transformation. By construction of L the linear terms in the numerator cancel leaving just the quadratic term,

$$\lim_{H\to 0} \frac{F(A+H) - F(A) - L(H)}{||H||} \;=\; \lim_{H\to 0} \frac{H^2}{||H||}.$$

It suffices to show that $\lim_{H\to 0} \frac{||H^2||}{||H||} = 0$ since $\lim(||g||) = 0$ iff $\lim(g) = 0$ in a normed vector space. Fortunately the normed vector space $\mathbb{R}^{\,n\times n}$ is actually a **Banach algebra***. A vector space with a multiplication operation is called an algebra. In the current context the multiplication is simply matrix multiplication. A Banach algebra is a normed vector space with a multiplication that satisfies $||XY|| \leq ||X||\,||Y||$. Thanks to this inequality*[4] *we can calculate our limit via the squeeze theorem. Observe $0 \leq \frac{||H^2||}{||H||} \leq ||H||$. As $H \to 0$ it follows $||H|| \to 0$ hence $\lim_{H\to 0} \frac{||H^2||}{||H||} = 0$. We find $dF_A(H) = AH + HA$.*

Generally constructing the derivative matrix for a function $f : V \to W$ where $V, W \neq \mathbb{R}$ involves a fair number of relatively ad-hoc conventions because the constructions necessarily involving choosing coordinates. The situation is similar in linear algebra. Writing abstract linear transformations in terms of matrix multiplication takes a little thinking. If you look back you'll notice that I did not bother to try to write a matrix for the differential in Examples 2.1.4 or 2.1.5.

Example 2.1.12. *Find the linearization of $F(X) = X^2$ at $X = I$. In Example 2.1.11 we proved $dF_A(H) = AH + HA$. Hence, for $A = I$ we find $dF_I(H) = IH + HI = 2H$. Thus the linearization is fairly simple to assemble,*

$$\begin{aligned} L_F^I(X) &= F(I) + dF_I(X - I) \qquad (2.2)\\ &= I + 2(X - I)\\ &= 2X - I. \end{aligned}$$

[4] it does take a bit of effort to prove this inequality holds for the matrix norm, I omit it since it would be distracting here

2.2 properties of the Frechet derivative

Linearity and the chain rule naturally generalize for Frechet derivatives on normed linear spaces. It is helpful for me to introduce some additional notation to analyze the convergence of the Frechet quotient: supposing that $F : dom(F) \subset V \to W$ is differentiable at a we set:

$$\eta_F(h) = F(a+h) - F(a) - dF_a(h) \tag{2.3}$$

hence the **Frechet quotient** can be written as:

$$\frac{\eta_F(h)}{\|h\|} = \frac{F(a+h) - F(a) - dF_a(h)}{\|h\|}. \tag{2.4}$$

Thus differentiability of F at a requires $\frac{\eta_F(h)}{\|h\|} \to 0$ as $h \to 0$. For $h \neq 0$ and $\|h\| < 1$ we have:

$$0 \leq \|\eta_F(h)\| < \frac{\|\eta_F(h)\|}{\|h\|}. \tag{2.5}$$

Thus $\|\eta_F(h)\| \to 0$ as $h \to 0$ by the squeeze theorem. Consequently,

$$\lim_{h \to 0} \eta_F(h) = 0. \tag{2.6}$$

Therefore, $\eta_F : V \to W$ is continuous at $h = 0$ since $\eta_F(0) = F(a) - F(a) - dF_a(0) = 0$ (I remind the reader that the linear transformation dF_a must map zero to zero). Continuity of η_F at $h = 0$ allows us to use theorems for continuous functions on η_F.

Theorem 2.2.1. *Linearity of the Frechet derivatives.*

Suppose V and W are normed linear spaces. If $F : dom(F) \subseteq V \to W$ and $G : dom(G) \subseteq V \to W$ are differentiable at a and $c \in \mathbb{R}$ then $cF + G$ is differentiable at a and

$$d(cF + G)_a = cdF_a + dG_a$$

Proof: Let $\eta_F(h) = F(a+h) - F(a) - dF_a(h)$ and $\eta_G(h) = G(a+h) - G(a) - dG_a(h)$ for all $h \in V$. Assume F and G differentiable at a hence $\lim_{h\to 0} \frac{\eta_F(h)}{\|h\|} = 0$ and $\lim_{h \to 0} \frac{\eta_G(h)}{\|h\|} = 0$. Moreover, $dF_a, dG_a : V \to W$ are linear hence $cdF_a + dG_a : V \to W$ is linear. Hence calculate,

$$\begin{aligned} \eta_{cF+G}(h) &= (cF+G)(a+h) - (cF+G)(a) - (cdF_a + dG_a)(h) \\ &= c\,(F(a+h) - F(a) - dF_a(h)) + G(a+h) - G(a) - dG_a(h) \\ &= c\eta_F(h) + \eta_G(h) \end{aligned} \tag{2.7}$$

Therefore, by Proposition 1.3.8, we complete the proof:

$$\lim_{h\to 0} \frac{\eta_{cF+G}(h)}{\|h\|} = \lim_{h \to 0} \frac{c\eta_F(h) + \eta_G(h)}{\|h\|} = c \lim_{h\to 0}\left(\frac{\eta_F(h)}{\|h\|}\right) + \lim_{h\to 0} \frac{\eta_G(h)}{\|h\|} = 0. \quad \square$$

Setting $c = 1$ or $G = 0$ we obtain important special cases:

$$d(F+G)_a = dF_a + dG_a \qquad \& \qquad d(cF)_a = cdF_a. \tag{2.8}$$

The chain rule is also a general rule of calculus on a NLS[5]. This single chain rule produces all the chain rules you saw in calculus I, II and III and much more. To appreciate this we need to understand partial differentiation for normed linear spaces.

[5] I state the rule with domains of the entire NLS, but, this can easily be stated for smaller domains like $F : U \subseteq V_1 \to V_2$ and $G : dom(G) \subseteq V_2 \to V_3$ where $F(U) \subset dom(G)$ so $F \circ G$ is well-defined, but, this has nothing to do with the theorem so I just made the domains uninteresting

Theorem 2.2.2. *chain rule for Frechet derivatives.*

> Suppose $G : V_1 \to V_2$ is differentiable at a and $F : V_2 \to V_3$ is differentiable at $G(a)$ then $F \circ G$ is differentiable at a and $d(F \circ G)_a = dF_{G(a)} \circ dG_a$.

The proof I offer here is not quite complete. The main ideas are here, but, there is a pesky term at the end which I have not quite pinned down to my liking. I found these notes by J. C. M. Grajales on page 40 have a proof which appears complete.

Proof: since G is differentiable at a we have the existence of η_G continuous at $h = 0$ defined by:

$$\eta_G(h) = G(a+h) - G(a) - dG_a(h) \tag{2.9}$$

Also, by differentiability of F at $G(a)$ we have the existence of η_F continuous at $k = 0$ given by:

$$\eta_F(k) = F(G(a)+k) - F(G(a)) - dF_{G(a)}(k) \tag{2.10}$$

Furthermore, the differentials are linear transformations and thus their composite $dF_{G(a)} \circ dG_a$ is likewise linear. It remains to show $\eta_{F \circ G}$ formed with $dF_{G(a)} \circ dG_a$ has the needed limiting property. Thus consider,

$$\begin{aligned} \eta_{F\circ G}(h) &= (F \circ G)(a+h) - (F \circ G)(a) - (dF_{G(a)} \circ dG_a)(h) \\ &= F(G(a+h)) - F(G(a)) - dF_{G(a)}(dG_a(h)) \\ &= F\left(G(a) + dG_a(h) + \eta_G(h)\right) - F(G(a)) - dF_{G(a)}(dG_a(h)) \\ &= F(G(a)) + dF_{G(a)}(dG_a(h) + \eta_G(h)) + \eta_F(dG_a(h) + \eta_G(h)) \\ &\qquad - F(G(a)) - dF_{G(a)}(dG_a(h)) \\ &= dF_{G(a)}(\eta_G(h)) + \eta_F(dG_a(h) + \eta_G(h)) \end{aligned} \tag{2.11}$$

where I used Equation 2.10 to make the expansion marked in blue. I need a bit of notation to help guide the remainder of the proof:

$$\frac{\eta_{F\circ G}(h)}{\|h\|} = \underbrace{\frac{1}{\|h\|} dF_{G(a)}(\eta_G(h))}_{(I.)} + \underbrace{\frac{1}{\|h\|}\eta_F(dG_a(h) + \eta_G(h))}_{(II.)} \tag{2.12}$$

We can understand (I.) using linearity and continuity of the linear map $dF_{G(a)}$:

$$\lim_{h\to 0}\left(\frac{1}{\|h\|} dF_{G(a)}(\eta_G(h))\right) = \lim_{h\to 0} dF_{G(a)}\left(\frac{\eta_G(h)}{\|h\|}\right) = dF_{G(a)}\left(\lim_{h\to 0}\frac{\eta_G(h)}{\|h\|}\right) = dF_{G(a)}(0) = 0. \tag{2.13}$$

To understand (II.) a substitution is helpful. Notice $dG_a(h) + \eta_G(h) \to 0$ as $h \to 0$. Let $k = dG_a(h) + \eta_G(h)$ and note $\frac{\eta_F(k)}{\|k\|} \to 0$ as $k \to 0$. Unfortunately, (II.) is not quite $\frac{\eta_F(k)}{\|k\|}$ since it has a denominator $\|h\|$ not $\|k\|$. We need to find a relation which binds $\|h\|$ and $\|k\|$. In particular, if we can find $m > 0$ for which $\|k\| < m\|h\|$ then

$$0 < \frac{\|\eta_F(k)\|}{m\|h\|} < \frac{\|\eta_F(k)\|}{\|k\|} \tag{2.14}$$

and we could argue (II.) vanishes as $h \to 0$ by the squeeze theorem. I leave this gap as an exercise for the reader. □

Remark 2.2.3.

> Other authors use the big and little O notation to help with the analysis of the proof above. It may be that if I adopted such notation is would help me will in the gap. For now I stick with my somewhat unusual η_F notation.

2.3 partial derivatives of differentiable maps

In the preceding sections we calculated the differential at a point via educated guessing. We should like to find better method to derive differentials. It turns out that we can systematically calculate the differential from partial derivatives of the component functions. However, certain topological conditions are required for us to properly paste together the partial derivatives of the component functions. We discuss the perils of proving differentiability from partial derivatives in Section 2.4. The purpose of the current section is to define partial differentiation and to explain how partial derivatives relate to the differential of a given differentiable map. To understand partial derivatives we begin with a study of directional derivatives. Once more we generalize the usual calculus III.

Remark 2.3.1.

Certainly parts of what is done in this section naturally generalize to an infinite dimensional context. You can read more about the Gateaux derivative in your future studies. However, here I limit our attention in this section to finite dimensional normed linear spaces.

2.3.1 partial differentiation in a finite dimensional real vector space

The directional derivative of a mapping F at a point $a \in dom(F)$ along v is defined to be the derivative of the curve $\gamma(t) = F(a + tv)$. In other words, the directional derivative gives you the instantaneous vector-rate of change in the mapping F at the point a along v.

Definition 2.3.2.

Suppose V and W are real normed linear spaces. Let $F : dom(F) \subseteq V \to W$ and suppose the limit below exists for $a \in dom(F)$ and $v \in V$ then we define the **directional derivative of F at a along v** to be $D_vF(a) \in W$ where

$$D_vF(a) = \lim_{h \to 0} \frac{F(a + hv) - F(a)}{h}$$

One great contrast we should pause to note is that the definition of the directional derivative is explicit whereas the definition of the differential was implicit. Naturally, if we take $V = W = \mathbb{R}$ then we recover the first semester difference quotient definition of the derivative at a point. This also reproduces the directional derivatives you were shown in multivariate calculus, except, we do not insist v have $\|v\| = 1$. Don't be fooled by the proof of the next Theorem, it's easier than it looks. Summary: since differentiability at a point controls the change of the map in all directions at a point in terms of the differential we can control the change in the map in a particular direction at the given point via the differential.

Theorem 2.3.3. *Differentiability implies directional differentiability.*

Let V, W be real normed linear spaces. If $F : U \subseteq V \to W$ is differentiable at $a \in U$ then the directional derivative $D_vF(a)$ exists for each $v \in V$ and $D_vF(a) = dF_a(v)$.

Proof: Suppose $a \in U$ such that dF_a is well-defined then we are given that

$$\lim_{h \to 0} \frac{F(a + h) - F(a) - dF_a(h)}{||h||} = 0.$$

This is a limit in V, when it exists it follows that the limits that approach the origin along particular paths also exist and are zero. Consider the path $t \mapsto tv$ for $v \neq 0$ and $t > 0$, we find

$$\lim_{tv \to 0,\ t>0} \frac{F(a+tv) - F(a) - dF_a(tv)}{||tv||} = \frac{1}{||v||} \lim_{t \to 0^+} \frac{F(a+tv) - F(a) - tdF_a(v)}{|t|} = 0.$$

Hence, as $|t| = t$ for $t > 0$ we find

$$\lim_{t \to 0^+} \frac{F(a+tv) - F(a)}{t} = \lim_{t \to 0} \frac{tdF_a(v)}{t} = dF_a(v).$$

Likewise we can consider the path $t \mapsto tv$ for $v \neq 0$ and $t < 0$

$$\lim_{tv \to 0,\ t<0} \frac{F(a+tv) - F(a) - dF_a(tv)}{||tv||} = \frac{1}{||v||} \lim_{t \to 0^-} \frac{F(a+tv) - F(a) - tdF_a(v)}{|t|} = 0.$$

Note $|t| = -t$ thus the limit above yields

$$\lim_{t \to 0^-} \frac{F(a+tv) - F(a)}{-t} = \lim_{t \to 0^-} \frac{tdF_a(v)}{-t} \quad \Rightarrow \quad \lim_{t \to 0^-} \frac{F(a+tv) - F(a)}{t} = dF_a(v).$$

Therefore,

$$\lim_{t \to 0} \frac{F(a+tv) - F(a)}{t} = dF_a(v)$$

and we conclude that $D_vF(a) = dF_a(v)$ for all $v \in V$ since the $v = 0$ case follows trivially. □

Partial derivatives are just directional derivatives in standard directions. In particular, given a basis $\beta = \{v_1, \dots, v_n\}$ with coordinate maps $x_1, \dots, x_n$ there is a standard concept of partial differentiation on an NLS:

Definition 2.3.4. *partial derivative with respect to coordinate on an NLS.*

Let V be a NLS with basis $\beta = \{v_1, \dots, v_n\}$ and coordinates $\Phi_\beta = (x_1, \dots, x_n)$. Then if $F : dom(F) \subseteq V \to W$ we define, for such points $a \in dom(F)$ as the limit exists,

$$\frac{\partial F}{\partial x_i}(a) = D_{v_i}F(a) = \lim_{h \to 0} \frac{F(a + hv_i) - F(a)}{h}.$$

Alternatively, we can present the partial derivative in terms of an ordinary derivative:

$$\frac{\partial F}{\partial x_i}(a) = \frac{d}{dt}\left[F(a + tv_i)\right]\bigg|_{t=0} \tag{2.15}$$

Let's revisit the map from Example 2.1.11 and see if we can recover the differential in terms of partial derivatives.

Example 2.3.5. *Let $F(X) = X^2$ for each $X \in \mathbb{R}^{n \times n}$. Let X_{ij} be the usual coordinates with respect to the standard matrix basis $\{E_{ij}\}$. Calculate the partial derivative of F with respect to X_{ij} at A: using Equation 2.15 with v_i replaced with E_{ij} and a with A,*

$$\begin{aligned} \frac{\partial F}{\partial X_{ij}}(A) &= \frac{d}{dt}\left[(A + tE_{ij})^2\right]\bigg|_{t=0} \qquad (2.16) \\ &= \frac{d}{dt}\left[A^2 + tAE_{ij} + tE_{ij}A + t^2E_{ij}^2\right]\bigg|_{t=0} \\ &= (AE_{ij} + E_{ij}A + 2tE_{ij}^2)\big|_{t=0} \\ &= AE_{ij} + E_{ij}A. \end{aligned}$$

If we know a map of normed linear spaces is differentiable then we can express the differential in terms of partial derivatives.

Theorem 2.3.6. *differentials can be built from partial derivatives.*

Let V, W be real normed linear spaces where V has basis $\beta = \{v_1, \dots, v_n\}$ with coordinates $x_1, \dots, x_n$. If $F : dom(F) \subseteq V \to W$ is differentiable at a and $h = \sum_{i=1}^n h_i v_i$ then

$$dF_a(h) = \sum_{i=1}^n h_i \frac{\partial F}{\partial x_i}(a).$$

Proof: observe that

$$dF_a\left(\sum_{i=1}^n h_i v_i\right) = \sum_{i=1}^n h_i dF_a(v_i) = \sum_{i=1}^n h_i D_{v_i}F(a) = \sum_{i=1}^n h_i \frac{\partial F}{\partial x_i}(a). \tag{2.17}$$

follows immediately from linearity of differential paired with Theorem 2.3.3. □

Let's apply the above result to Example 2.3.5.

Example 2.3.7. *Consider $F(X) = X^2$ for $X \in \mathbb{R}^{\,n\times n}$. Construct the differential from the partial derivatives with respect to the standard basis matrix $\{E_{ij}\}$. Let $H = \sum_{i,j} H_{ij}E_{ij}$ and calculate using Equation 2.16*

$$dF_A(H) = \sum_{i,j} H_{ij}(AE_{ij} + E_{ij}A) = A\left(\sum_{i,j} H_{ij}E_{ij}\right) + \left(\sum_{i,j} H_{ij}E_{ij}\right)A = AH + HA.$$

I should emphasize, at this point in our development, we cannot conclude the differential exists merely from partial derivatives existing[6]. The example above is reasonable because we have already shown differentiability of the $F(A) = A^2$ map in Example 2.1.11.

Remark 2.3.8.

I have deliberately defined the derivative in slightly more generality than we need for this course. It's probably not much trouble to continue to develop the theory of differentiation for a normed vector space, however I will for the most part stop here modulo an example here or there. If you understand many of the theorems that follow from here on out for $\mathbb{R}^n$ then it is a simple matter to transfer arguments to the setting of a Banach space by using an appropriate isomorphism. Traditionally this type of course only covers continuous differentiability, inverse and implicit function theorems in the context of mappings from $\mathbb{R}^n$ to $\mathbb{R}^m$.

For the reader interested in generalizing these results to the context of an abstract normed vector space feel free to discuss it with me sometime. Also, if you want to read a master on these topics you could look at the text by Shlomo Sternberg on Advanced Calculus. He develops many things for normed spaces. Or, take a look at Dieudonne's Modern Analysis which pays special attention to reaping infinite dimensional results from our finite-dimensional arguments. I also find Zorich's two volume set on Mathematical Analysis is quite helpful. I'm hoping to borrow some arguments from Zorich in this update to my notes. Any of these texts would be good to read to follow-up my course with something deeper.

[6] we study this in depth in Section 2.4.

2.3.2 partial differentiation for real

Consider $F : dom(F) \subseteq \mathbb{R}^n \to \mathbb{R}^n$, in this case the differential dF_a is a linear transformation from $\mathbb{R}^n \to \mathbb{R}^n$ and we can calculate the standard matrix for the differential using the preceding proposition. Recall that if $L : \mathbb{R}^n \to \mathbb{R}^m$ then the standard matrix was simply

$$[L] = [L(e_1)|L(e_2)|\cdots|L(e_n)]$$

and thus the action of L is expressed nicely as a matrix multiplication; $L(v) = [L]v$. Similarly, $dF_a : \mathbb{R}^n \to \mathbb{R}^m$ is linear transformation and thus $dF_a(v) = [dF_a]v$ where

$$[dF_a] = [dF_a(e_1)|dF_a(e_2)|\cdots|dF_a(e_n)].$$

Moreover, by the preceding proposition we can calculate $dF_a(e_j) = D_{e_j}F(a)$ for $j = 1, 2, \dots, n$. Clearly the directional derivatives in the coordinate directions are of great importance. For this reason we make the following definition:

Definition 2.3.9. *Partial derivatives are directional derivatives in coordinate directions.*

Suppose that $F : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ is a mapping the we say that F is **has partial derivative** $\frac{\partial F}{\partial x_i}(a)$ at $a \in U$ iff the directional derivative in the e_i direction exists at a. In this case we denote,

$$\frac{\partial F}{\partial x_i}(a) = D_{e_i}F(a).$$

Also, the notation $D_{e_i}F(a) = D_iF(a)$ or $\partial_i F = \frac{\partial F}{\partial x_i}$ is convenient. We construct the **partial derivative function** $\partial_i F : V \subseteq \mathbb{R}^n \to \mathbb{R}^m$ as the function defined pointwise for each $v \in V$ where $\partial_i F(v)$ exists.

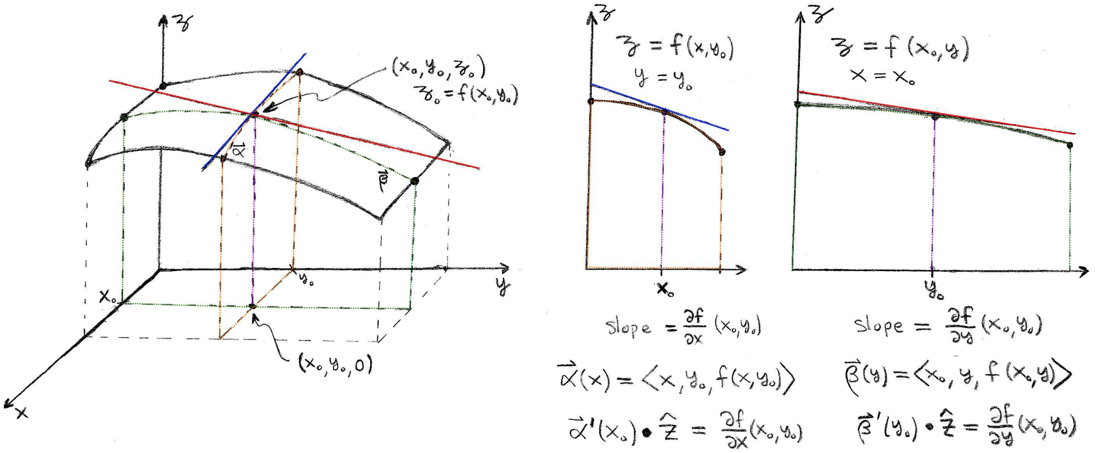
Let's expand this definition a bit. Note that if $F = (F_1, F_2, \dots, F_m)$ then

$$D_{e_i}F(a) = \lim_{h\to 0} \frac{F(a+he_i) - F(a)}{h} \quad \Rightarrow \quad [D_{e_i}F(a)] \cdot e_j = \lim_{h\to 0} \frac{F_j(a+he_i) - F_j(a)}{h}$$

for each $j = 1, 2, \dots m$. But then the limit of the component function F_j is precisely the directional derivative at a along e_i hence we find the result

$$\frac{\partial F}{\partial x_i} \cdot e_j = \frac{\partial F_j}{\partial x_i} \qquad \text{in other words,} \qquad \boxed{\partial_i F = (\partial_i F_1, \partial_i F_2, \dots, \partial_i F_m).}$$

In the particular case $f : \mathbb{R}^2 \to \mathbb{R}$ the partial derivatives with respect to x and y at (x_o, y_o) are related to the graph $z = f(x, y)$ as illustrated below:



z
(x_o, y_o, z_o)
$z_o = f(x_o, y_o)$
$\vec{\alpha}$
$\vec{\beta}$
x_o
y_o
y
x
$(x_o, y_o, 0)$
z
$z = f(x, y_o)$
$y = y_o$
x_o
slope $= \frac{\partial f}{\partial x}(x_o, y_o)$
$\vec{\alpha}(x) = \langle x, y_o, f(x, y_o)\rangle$
$\vec{\alpha}'(x_o) \bullet \hat{z} = \frac{\partial f}{\partial x}(x_o, y_o)$
z
$z = f(x_o, y)$
$x = x_o$
y_o
slope $= \frac{\partial f}{\partial y}(x_o, y_o)$
$\vec{\beta}(y) = \langle x_o, y, f(x_o, y)\rangle$
$\vec{\beta}'(y_o) \bullet \hat{z} = \frac{\partial f}{\partial y}(x_o, y_o)$

Similar pictures can be imagined for partial derivatives of more variables, even for vector-valued maps, but direct visualization is not possible (at least for me).

The proposition below shows how the differential of a m-vector-valued function of n-real variables is connected to a matrix of partial derivatives.

Proposition 2.3.10.

If $F : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at $a \in U$ then the differential dF_a has derivative matrix $F'(a)$ and it has components which are expressed in terms of partial derivatives of the component functions:

$$[dF_a]_{ij} = \partial_j F_i$$

for $1 \leq i \leq m$ and $1 \leq j \leq n$.

Perhaps it is helpful to expand the derivative matrix explicitly for future reference:

$$F'(a) = \begin{bmatrix} \partial_1 F_1(a) & \partial_2 F_1(a) & \cdots & \partial_n F_1(a) \\ \partial_1 F_2(a) & \partial_2 F_2(a) & \cdots & \partial_n F_2(a) \\ \vdots & \vdots & \vdots & \vdots \\ \partial_1 F_m(a) & \partial_2 F_m(a) & \cdots & \partial_n F_m(a) \end{bmatrix}$$

Let's write the operation of the differential for a differentiable mapping at some point $a \in \mathbb{R}$ in terms of the explicit matrix multiplication by $F'(a)$. Let $v = (v_1, v_2, \dots v_n) \in \mathbb{R}^n$,

$$dF_a(v) = F'(a)v = \begin{bmatrix} \partial_1 F_1(a) & \partial_2 F_1(a) & \cdots & \partial_n F_1(a) \\ \partial_1 F_2(a) & \partial_2 F_2(a) & \cdots & \partial_n F_2(a) \\ \vdots & \vdots & \vdots & \vdots \\ \partial_1 F_m(a) & \partial_2 F_m(a) & \cdots & \partial_n F_m(a) \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

You may recall the notation from calculus III at this point, omitting the a-dependence,

$$\boxed{\nabla F_j = grad(F_j) = \left[\ \partial_1 F_j,\ \partial_2 F_j,\ \cdots,\ \partial_n F_j\ \right]^T}$$

So if the derivative exists we can write it in terms of a stack of gradient vectors of the component functions: (I used a transpose to write the stack side-ways),

$$F' = \big[\nabla F_1 | \nabla F_2 | \cdots | \nabla F_m\big]^T$$

Finally, just to collect everything together,

$$\boxed{F' = \begin{bmatrix} \partial_1 F_1 & \partial_2 F_1 & \cdots & \partial_n F_1 \\ \partial_1 F_2 & \partial_2 F_2 & \cdots & \partial_n F_2 \\ \vdots & \vdots & \vdots & \vdots \\ \partial_1 F_m & \partial_2 F_m & \cdots & \partial_n F_m \end{bmatrix} = \left[\partial_1 F \mid \partial_2 F \mid \cdots \mid \partial_n F\ \right] = \begin{bmatrix} (\nabla F_1)^T \\ \hline (\nabla F_2)^T \\ \hline \vdots \\ \hline (\nabla F_m)^T \end{bmatrix}}$$

Example 2.3.11. *Recall that in Example 2.1.6 we showed that $F : \mathbb{R}^2 \to \mathbb{R}^3$ defined by $F(x, y) = (xy,\ x^2,\ x + 3y)$ for all $(x, y) \in \mathbb{R}^2$ was differentiable. In fact we calculated that*

$$dF_{(x,y)}(h, k) = \begin{bmatrix} y & x \\ 2x & 0 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} h \\ k \end{bmatrix}.$$

If you recall from calculus III the mechanics of partial differentiation it's simple to see that

$$\frac{\partial F}{\partial x} = \frac{\partial}{\partial x}(xy,\ x^2,\ x+3y) = (y,\ 2x, 1) = \begin{bmatrix} y \\ 2x \\ 1 \end{bmatrix}$$

$$\frac{\partial F}{\partial y} = \frac{\partial}{\partial y}(xy,\ x^2,\ x+3y) = (x,\ 0, 3) = \begin{bmatrix} x \\ 0 \\ 3 \end{bmatrix}$$

Thus $[dF] = [\partial_x F | \partial_y F]$ *(as we expect given the derivations in this section!)*

Directional derivatives and partial derivatives are of secondary importance in this course. They are merely the substructure of what is truly of interest: the differential. That said, it is useful to know how to construct directional derivatives via partial derivative formulas. In fact, in careless calculus texts it sometimes presented as the definition.

Proposition 2.3.12.

If $F : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at $a \in U$ then the directional derivative $D_vF(a)$ can be expressed as a sum of partial derivative maps for each $v =< v_1, v_2, \dots, v_n > \in \mathbb{R}^n$:

$$D_vF(a) = \sum_{j=1}^{n} v_j \partial_j F(a)$$

Proof: since F is differentiable at a the differential dF_a exists and $D_vF(a) = dF_a(v)$ for all $v \in \mathbb{R}^n$. Use linearity of the differential to calculate that

$$D_vF(a) = dF_a(v_1e_1 + \cdots + v_ne_n) = v_1 dF_a(e_1) + \cdots + v_n dF_a(e_n).$$

Note $dF_a(e_j) = D_{e_j}F(a) = \partial_j F(a)$ and the prop. follows. $\square$

Example 2.3.13. *Suppose* $f : \mathbb{R}^3 \to \mathbb{R}$ *then* $\nabla f = [\partial_x f, \partial_y f, \partial_z f]^T$ *and we can write the directional derivative in terms of*

$$D_v f = [\partial_x f, \partial_y f, \partial_z f]^T v = \nabla f \cdot v$$

if we insist that $||v|| = 1$ *then we recover the standard directional derivative we discuss in calculus III. Naturally the* $||\nabla f(a)||$ *yields the maximum value for the directional derivative at* a *if we limit the inputs to vectors of unit-length. If we did not limit the vectors to unit length then the directional derivative at* a *can become arbitrarily large as* $D_v f(a)$ *is proportional to the magnitude of* v*. Since our primary motivation in calculus III was describing rates of change along certain directions for some multivariate function it made sense to specialize the directional derivative to vectors of unit-length. The definition used in these notes better serves the theoretical discussion.*[7]

2.3.3 examples of Jacobian matrices

Our goal here is simply to exhibit the Jacobian matrix and partial derivatives for a few mappings. At the base of all these calculations is the observation that partial differentiation is just ordinary differentiation where we treat all the independent variable not being differentiated as constants. The criteria of independence is important. We'll study the case where variables are not independent in a later section (see implicit differentiation).

[7] If you read my calculus III notes you'll find a derivation of how the directional derivative in Stewart's calculus arises from the general definition of the derivative as a linear mapping. Look up page 305g.

Example 2.3.14. *Let $f(t) = (t, t^2, t^3)$ then $f'(t) = (1, 2t, 3t^2)$. In this case we have*

$$f'(t) = [df_t] = \begin{bmatrix} 1 \\ 2t \\ 3t^2 \end{bmatrix}$$

Example 2.3.15. *Let $f(\vec{x}, \vec{y}) = \vec{x} \cdot \vec{y}$ be a mapping from $\mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$. I'll denote the coordinates in the domain by $(x_1, x_2, x_3, y_1, y_2, y_3)$ thus $f(\vec{x}, \vec{y}) = x_1y_1 + x_2y_2 + x_3y_3$. Calculate,*

$$[df_{(\vec{x},\vec{y})}] = \nabla f(\vec{x}, \vec{y})^T = [y_1, y_2, y_3, x_1, x_2, x_3]$$

Example 2.3.16. *Let $f(\vec{x}, \vec{y}) = \vec{x} \cdot \vec{y}$ be a mapping from $\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$. I'll denote the coordinates in the domain by $(x_1, \dots, x_n, y_1, \dots, y_n)$ thus $f(\vec{x}, \vec{y}) = \sum_{i=1}^n x_iy_i$. Calculate,*

$$\frac{\partial}{x_j}\left[\sum_{i=1}^n x_iy_i\right] = \sum_{i=1}^n \frac{\partial x_i}{x_j} y_i = \sum_{i=1}^n \delta_{ij}y_i = y_j$$

Likewise,

$$\frac{\partial}{y_j}\left[\sum_{i=1}^n x_iy_i\right] = \sum_{i=1}^n x_i\frac{\partial y_i}{y_j} = \sum_{i=1}^n x_i\delta_{ij} = x_j$$

Therefore, noting that $\nabla f = (\partial_{x_1} f, \dots, \partial_{x_n} f, \partial_{y_1} f, \dots, \partial_{y_n} f)$,

$$[df_{(\vec{x},\vec{y})}]^T = (\nabla f)(\vec{x}, \vec{y}) = \vec{y} \times \vec{x} = (y_1, \dots, y_n, x_1, \dots, x_n)$$

Example 2.3.17. *Suppose $F(x, y, z) = (xyz, y, z)$ we calculate,*

$$\frac{\partial F}{\partial x} = (yz, 0, 0) \qquad \frac{\partial F}{\partial y} = (xz, 1, 0) \qquad \frac{\partial F}{\partial z} = (xy, 0, 1)$$

Remember these are actually column vectors in my sneaky notation; $(v_1, \dots, v_n) = [v_1, \dots, v_n]^T$. This means the **derivative** *or* **Jacobian matrix** *of F at (x, y, z) is*

$$F'(x, y, z) = [dF_{(x,y,z)}] = \begin{bmatrix} yz & xz & xy \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Example 2.3.18. *Suppose $F(x, y, z) = (x^2 + z^2, yz)$ we calculate,*

$$\frac{\partial F}{\partial x} = (2x, 0) \qquad \frac{\partial F}{\partial y} = (0, z) \qquad \frac{\partial F}{\partial z} = (2z, y)$$

The derivative is a 2×3 matrix in this example,

$$F'(x, y, z) = [dF_{(x,y,z)}] = \begin{bmatrix} 2x & 0 & 2z \\ 0 & z & y \end{bmatrix}$$

Example 2.3.19. *Suppose $F(x, y) = (x^2 + y^2, xy, x + y)$ we calculate,*

$$\frac{\partial F}{\partial x} = (2x, y, 1) \qquad \frac{\partial F}{\partial y} = (2y, x, 1)$$

The derivative is a 3×2 matrix in this example,

$$F'(x, y) = [dF_{(x,y)}] = \begin{bmatrix} 2x & 2y \\ y & x \\ 1 & 1 \end{bmatrix}$$

Example 2.3.20. *Suppose $P(x,v,m) = (P_o, P_1) = (\frac{1}{2}mv^2 + \frac{1}{2}kx^2, mv)$ for some constant k. Let's calculate the derivative via gradients this time,*

$$\nabla P_o = (\partial P_o/\partial x, \partial P_o/\partial v, \partial P_o/\partial m) = (kx, mv, \tfrac{1}{2}v^2)$$

$$\nabla P_1 = (\partial P_1/\partial x, \partial P_1/\partial v, \partial P_1/\partial m) = (0, m, v)$$

Therefore,

$$P'(x,v,m) = \left[\begin{array}{ccc} kx & mv & \frac{1}{2}v^2 \\ 0 & m & v \end{array}\right]$$

Example 2.3.21. *Let $F(r,\theta) = (r\cos\theta, r\sin\theta)$. We calculate,*

$$\partial_r F = (\cos\theta, \sin\theta) \qquad \textit{and} \qquad \partial_\theta F = (-r\sin\theta, r\cos\theta)$$

Hence,

$$F'(r,\theta) = \left[\begin{array}{cc} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{array}\right]$$

Example 2.3.22. *Let $G(x,y) = (\sqrt{x^2+y^2}, \tan^{-1}(y/x))$. We calculate,*

$$\partial_x G = \big(\tfrac{x}{\sqrt{x^2+y^2}}, \tfrac{-y}{x^2+y^2}\big) \qquad \textit{and} \qquad \partial_y G = \big(\tfrac{y}{\sqrt{x^2+y^2}}, \tfrac{x}{x^2+y^2}\big)$$

Hence,

$$G'(x,y) = \left[\begin{array}{cc} \frac{x}{\sqrt{x^2+y^2}} & \frac{y}{\sqrt{x^2+y^2}} \\ \frac{-y}{x^2+y^2} & \frac{x}{x^2+y^2} \end{array}\right] = \left[\begin{array}{cc} \frac{x}{r} & \frac{y}{r} \\ \frac{-y}{r^2} & \frac{x}{r^2} \end{array}\right] \qquad \big(\textit{ using } r = \sqrt{x^2+y^2}\ \big)$$

Example 2.3.23. *Let $F(x,y) = (x, y, \sqrt{R^2-x^2-y^2})$ for a constant R. We calculate,*

$$\nabla\sqrt{R^2-x^2-y^2} = \left(\frac{-x}{\sqrt{R^2-x^2-y^2}}, \frac{-y}{\sqrt{R^2-x^2-y^2}} \right)$$

Also, $\nabla x = (1,0)$ and $\nabla y = (0,1)$ thus

$$F'(x,y) = \left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \\ \frac{-x}{\sqrt{R^2-x^2-y^2}} & \frac{-y}{\sqrt{R^2-x^2-y^2}} \end{array}\right]$$

Example 2.3.24. *Let $F(x,y,z) = (x,y,z,\sqrt{R^2-x^2-y^2-z^2})$ for a constant R. We calculate,*

$$\nabla\sqrt{R^2-x^2-y^2-z^2} = \left(\frac{-x}{\sqrt{R^2-x^2-y^2-z^2}}, \frac{-y}{\sqrt{R^2-x^2-y^2-z^2}}, \frac{-z}{\sqrt{R^2-x^2-y^2-z^2}} \right)$$

Also, $\nabla x = (1,0,0)$, $\nabla y = (0,1,0)$ and $\nabla z = (0,0,1)$ thus

$$F'(x,y,z) = \left[\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ \frac{-x}{\sqrt{R^2-x^2-y^2-z^2}} & \frac{-y}{\sqrt{R^2-x^2-y^2-z^2}} & \frac{-z}{\sqrt{R^2-x^2-y^2-z^2}} \end{array}\right]$$

Example 2.3.25. *Let $f(x,y,z) = (x+y, y+z, x+z, xyz)$. You can calculate,*

$$[df_{(x,y,z)}] = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \\ yz & xz & xy \end{bmatrix}$$

Example 2.3.26. *Let $f(x,y,z) = xyz$. You can calculate,*

$$[df_{(x,y,z)}] = \begin{bmatrix} yz & xz & xy \end{bmatrix}$$

Example 2.3.27. *Let $f(x,y,z) = (xyz, 1-x-y)$. You can calculate,*

$$[df_{(x,y,z)}] = \begin{bmatrix} yz & xz & xy \\ -1 & -1 & 0 \end{bmatrix}$$

Example 2.3.28. *Let $f : \mathbb{R}^3 \times \mathbb{R}^3$ be defined by $f(x) = x \times v$ for a fixed vector $v \neq 0$. We denote $x = (x_1, x_2, x_3)$ and calculate,*

$$\frac{\partial}{\partial x_a}(x \times v) = \frac{\partial}{\partial x_a}\Big(\sum_{i,j,k} \epsilon_{ijk} x_i v_j e_k \Big) = \sum_{i,j,k} \epsilon_{ijk} \frac{\partial x_i}{\partial x_a} v_j e_k = \sum_{i,j,k} \epsilon_{ijk} \delta_{ia} v_j e_k = \sum_{j,k} \epsilon_{ajk} v_j e_k$$

It follows,

$$\frac{\partial}{\partial x_1}(x \times v) = \sum_{j,k} \epsilon_{1jk} v_j e_k = v_2 e_3 - v_3 e_2 = (0, -v_3, v_2)$$

$$\frac{\partial}{\partial x_2}(x \times v) = \sum_{j,k} \epsilon_{2jk} v_j e_k = v_3 e_1 - v_1 e_3 = (v_3, 0, -v_1)$$

$$\frac{\partial}{\partial x_3}(x \times v) = \sum_{j,k} \epsilon_{3jk} v_j e_k = v_1 e_2 - v_2 e_1 = (-v_2, v_1, 0)$$

Thus the Jacobian is simply,

$$[df_{(x,y)}] = \begin{bmatrix} 0 & v_3 & -v_2 \\ -v_3 & 0 & -v_1 \\ v_2 & v_1 & 0 \end{bmatrix}$$

In fact, $df_p(h) = f(h) = h \times v$ for each $p \in \mathbb{R}^3$. The given mapping is linear so the differential of the mapping is precisely the mapping itself (we could short-cut much of this calculation and simply quote Example 2.1.4 where we proved $dT = T$ for linear T).

Example 2.3.29. *Let $f(x,y) = (x, y, 1-x-y)$. You can calculate,*

$$[df_{(x,y,z)}] = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -1 & -1 \end{bmatrix}$$

Example 2.3.30. *Let $X(u,v) = (x,y,z)$ where x, y, z denote functions of u, v and I prefer to omit the explicit depedendence to reduce clutter in the equations to follow.*

$$\frac{\partial X}{\partial u} = X_u = (x_u, y_u, z_u) \quad \text{and} \quad \frac{\partial X}{\partial v} = X_v = (x_v, y_v, z_v)$$

Then the Jacobian is the 3×2 matrix

$$\big[dX_{(u,v)}\big] = \begin{bmatrix} x_u & x_v \\ y_u & y_v \\ z_u & z_v \end{bmatrix}$$

Remark 2.3.31.

I return to these examples in the next chapter and we'll explore the geometric content of these formulas as they support the application of certain theorems. More on that later, for the remainder of this chapter we continue to focus on properties of differentiation.

2.3.4 on chain rule and Jacobian matrix multiplication

In calculus III you may have learned how to calculate partial derivatives in terms of tree-diagrams and intermediate variable etc... We now have a way of understanding those rules and all the other chain rules in terms of one over-arching calculation: matrix multiplication of the constituent Jacobians in the composite function. Of course once we have this rule for the composite of two functions we can generalize to n-functions by a simple induction argument. For example, for three suitably defined mappings F, G, H,

$$(F \circ G \circ H)'(a) = F'(G(H(a)))G'(H(a))H'(a)$$

Example 2.3.32. .

$$f = f(x,y) \longrightarrow f' = (\nabla f)^T = \overbrace{\left[\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right]}^{1\times 2 \text{ Jacobian}}$$

$$f : \mathbb{R}^2 \longrightarrow \mathbb{R}$$

$$\vec{r}(u,v,w) = [x(u,v,w), y(u,v,w)]^T \longrightarrow \vec{r}\,' = \left[\frac{\partial \vec{r}}{\partial u} \Big| \frac{\partial \vec{r}}{\partial v} \Big| \frac{\partial \vec{r}}{\partial w} \right]$$

$$\vec{r} : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$$

$$\vec{r}\,' = \underbrace{\begin{bmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} & \frac{\partial x}{\partial w} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} & \frac{\partial y}{\partial w} \end{bmatrix}}_{(2\times 3) \text{ Jacobian}}$$

$$f \circ \vec{r} : \mathbb{R}^3 \longrightarrow \mathbb{R}^2 \longrightarrow \mathbb{R}$$

$$\begin{aligned}
(f \circ \vec{r})' &= f' \vec{r}\,' \\
&= \left[\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right] \begin{bmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} & \frac{\partial x}{\partial w} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} & \frac{\partial y}{\partial w} \end{bmatrix} \\
&= \Big[\underbrace{\frac{\partial f}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial u}}_{\frac{\partial f}{\partial u}}, \underbrace{\frac{\partial f}{\partial x}\frac{\partial x}{\partial v} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial v}}_{\frac{\partial f}{\partial v}}, \underbrace{\frac{\partial f}{\partial x}\frac{\partial x}{\partial w} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial w}}_{\frac{\partial f}{\partial w}} \Big]
\end{aligned}$$

Example 2.3.33. .

$$\underbrace{f : \mathbb{R} \longrightarrow \mathbb{R}^n} \quad \text{and} \quad g : \mathbb{R}^n \longrightarrow \mathbb{R}$$

$$\frac{df}{dt} = \left[\frac{df_1}{dt}, \frac{df_2}{dt}, \dots, \frac{df_n}{dt} \right]^T \quad \text{and} \quad g'(\vec{x}) = (\nabla g)(\vec{x})^T = \left[\frac{\partial g}{\partial x_1}, \dots, \frac{\partial g}{\partial x_n} \right]$$

$$(g \circ f)'(t) = g'(f(t))\, f'(t) = \left[\frac{\partial g}{\partial x_1}, \dots, \frac{\partial g}{\partial x_n} \right] \begin{bmatrix} df_1/dt \\ \vdots \\ df_n/dt \end{bmatrix}$$

$$\underline{\frac{d}{dt}\Big[g(f(t))\Big] = \frac{\partial g}{\partial x_1}\frac{df_1}{dt} + \frac{\partial g}{\partial x_2}\frac{df_2}{dt} + \dots + \frac{\partial g}{\partial x_n}\frac{df_n}{dt}}.$$

Example 2.3.34. .

Let $f(x,y) = x^2y^2$ and $g(t) = (t, t^2)$

We have $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ and $g : \mathbb{R} \longrightarrow \mathbb{R}^2$ note,

$$f'(x,y) = [2xy^2, \ 2x^2y] \quad \text{and} \quad g'(t) = \begin{bmatrix} 1 \\ 2t \end{bmatrix}$$

Note $f \circ g : \mathbb{R} \to \mathbb{R}^2 \to \mathbb{R}$ has

$$(f \circ g)'(t) = f'(g(t))g'(t) = f'(t,t^2)g'(t) = [2t^5, 2t^4]\begin{bmatrix} 1 \\ 2t \end{bmatrix} = 6t^5.$$

Note that $(f \circ g)(t) = f(t, t^2) = t^2t^4 = t^6$ so this result is not surprising!

Example 2.3.35. .

$$T(r,\theta) = (r\cos\theta, r\sin\theta) = (x,y)$$

$$T' = \begin{bmatrix} \frac{\partial x}{\partial r} & \frac{\partial y}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{bmatrix} = \begin{bmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{bmatrix}$$

If $w = f(x,y)$ then $w = \underbrace{g(r,\theta) = f(T(r,\theta))}_{f \text{ rewritten in polars.}}$

$$\left[\frac{\partial g}{\partial r}, \frac{\partial g}{\partial \theta} \right] = \left[\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right] \begin{bmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{bmatrix} = \left[\cos\theta \frac{\partial f}{\partial x} + \sin\theta \frac{\partial f}{\partial y}, -r\sin\theta \frac{\partial f}{\partial x} + r\cos\theta \frac{\partial f}{\partial y} \right]$$

With the proper understanding we have derived,

$$\frac{\partial}{\partial r} = \cos\theta \frac{\partial}{\partial x} + \sin\theta \frac{\partial}{\partial y}$$

$$\frac{\partial}{\partial \theta} = -r\sin\theta \frac{\partial}{\partial x} + r\cos\theta \frac{\partial}{\partial y}$$

You can invert these, $r = \sqrt{x^2+y^2}$ & $\theta = \tan^{-1}(y/x)$

$$\frac{\partial}{\partial x} = \frac{\partial r}{\partial x}\frac{\partial}{\partial r} + \frac{\partial \theta}{\partial x}\frac{\partial}{\partial \theta} = \frac{x}{r}\frac{\partial}{\partial r} - \frac{y}{r^2}\frac{\partial}{\partial \theta} = \cos\theta \frac{\partial}{\partial r} - \frac{\sin\theta}{r}\frac{\partial}{\partial \theta}$$

$$\frac{\partial}{\partial y} = \frac{\partial r}{\partial y}\frac{\partial}{\partial r} + \frac{\partial \theta}{\partial y}\frac{\partial}{\partial \theta} = \frac{y}{r}\frac{\partial}{\partial r} + \frac{x}{r^2}\frac{\partial}{\partial \theta} = \sin\theta \frac{\partial}{\partial r} + \frac{\cos\theta}{r}\frac{\partial}{\partial \theta}$$

You can use these to change coordinates. For example

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = \left(\cos\theta \frac{\partial}{\partial r} - \frac{\sin\theta}{r}\frac{\partial}{\partial \theta}\right)\left(\cos\theta \frac{\partial f}{\partial r} - \frac{\sin\theta}{r}\frac{\partial f}{\partial \theta}\right)$$
$$+ \left(\sin\theta \frac{\partial}{\partial r} + \frac{\cos\theta}{r}\frac{\partial}{\partial \theta}\right)\left(\sin\theta \frac{\partial f}{\partial r} + \frac{\cos\theta}{r}\frac{\partial f}{\partial \theta}\right)$$

$$= \cos^2\theta \frac{\partial^2 f}{\partial r^2} - \cancel{\cos\theta\sin\theta \frac{\partial}{\partial r}\left[\frac{1}{r}\frac{\partial f}{\partial \theta}\right]} - \frac{\sin\theta}{r}\frac{\partial}{\partial \theta}\left[\cos\theta \frac{\partial f}{\partial r}\right] + \frac{\sin\theta}{r^2}\frac{\partial}{\partial \theta}\left[\sin\theta \frac{\partial f}{\partial \theta}\right]$$
$$+ \sin^2\theta \frac{\partial^2 f}{\partial r^2} + \cancel{\sin\theta\cos\theta \frac{\partial}{\partial r}\left[\frac{1}{r}\frac{\partial f}{\partial \theta}\right]} + \frac{\cos\theta}{r}\frac{\partial}{\partial \theta}\left[\sin\theta \frac{\partial f}{\partial r}\right] + \frac{\cos\theta}{r^2}\frac{\partial}{\partial \theta}\left[\cos\theta \frac{\partial f}{\partial \theta}\right]$$

$$= \frac{\partial^2 f}{\partial r^2} + \frac{\sin^2\theta}{r}\frac{\partial f}{\partial r} + \frac{\cos^2\theta}{r}\frac{\partial f}{\partial r} - \cancel{\frac{\sin\theta\cos\theta}{r}\frac{\partial^2 f}{\partial \theta \partial r}} + \cancel{\frac{\cos\theta\sin\theta}{r}\frac{\partial^2 f}{\partial \theta \partial r}} + $$
$$+ \cancel{\frac{\sin\theta\cos\theta}{r^2}\frac{\partial f}{\partial \theta}} - \cancel{\frac{\cos\theta\sin\theta}{r^2}\frac{\partial f}{\partial \theta}} + \frac{\sin^2\theta}{r^2}\frac{\partial^2 f}{\partial \theta^2} + \frac{\cos^2\theta}{r^2}\frac{\partial^2 f}{\partial \theta^2}$$

$$\therefore \boxed{\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = \frac{\partial^2 f}{\partial r^2} + \frac{1}{r}\frac{\partial f}{\partial r} + \frac{1}{r^2}\frac{\partial^2 f}{\partial \theta^2}}$$

2.4 continuous differentiability

We have noted that differentiablility on some set U implies all sorts of nice formulas in terms of the partial derivatives. Curiously the converse is not quite so simple. It is possible for the partial derivatives to exist on some set and yet the mapping may fail to be differentiable. We need an extra topological condition on the partial derivatives if we are to avoid certain pathological[8] examples.

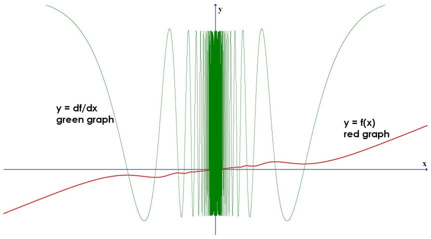
Example 2.4.1. *I found this example in Hubbard's advanced calculus text(see Ex. 1.9.4, pg. 123). It is a source of endless odd examples, notation and bizarre quotes. Let $f(x) = 0$ and*

$$f(x) = \frac{x}{2} + x^2 \sin\frac{1}{x}$$

for all $x \neq 0$. I can be shown that the derivative $f'(0) = 1/2$. Moreover, we can show that $f'(x)$ exists for all $x \neq 0$, we can calculate:

$$f'(x) = \frac{1}{2} + 2x\sin\frac{1}{x} - \cos\frac{1}{x}$$

Notice that $dom(f') = \mathbb{R}$. Note then that the tangent line at $(0,0)$ is $y = x/2$.



You might be tempted to say then that this function is increasing at a rate of 1/2 for x near zero. But this claim would be false since you can see that $f'(x)$ oscillates wildly without end near zero. We have a tangent line at $(0,0)$ with positive slope for a function which is not increasing at $(0,0)$ (recall that increasing is a concept we must define in a open interval to be careful). This sort of thing cannot happen if the derivative is continuous near the point in question.

The one-dimensional case is really quite special, even though we had discontinuity of the derivative we still had a well-defined tangent line to the point. However, many interesting theorems in calculus of one-variable require the function to be continuously differentiable near the point of interest. For example, to apply the 2nd-derivative test we need to find a point where the first derivative is zero and the second derivative exists. We cannot hope to compute $f''(x_o)$ unless f' is continuous at x_o. The next example is *sick*.

Example 2.4.2. *Let us define $f(0,0) = 0$ and*

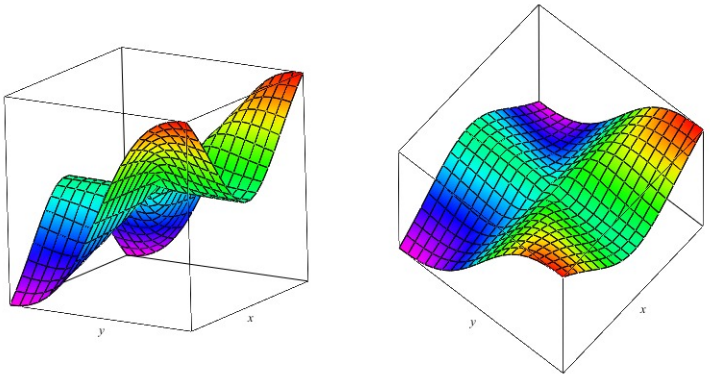
$$f(x,y) = \frac{x^2 y}{x^2 + y^2}$$

for all $(x,y) \neq (0,0)$ in $\mathbb{R}^2$. It can be shown that f is continuous at $(0,0)$. Moreover, since $f(x,0) = f(0,y) = 0$ for all x and all y it follows that f vanishes identically along the coordinate axis. Thus the rate of change in the e_1 or e_2 directions is zero. We can calculate that

$$\frac{\partial f}{\partial x} = \frac{2xy^3}{(x^2+y^2)^2} \qquad \textit{and} \qquad \frac{\partial f}{\partial y} = \frac{x^4 - x^2y^2}{(x^2+y^2)^2}$$

If you examine the plot of $z = f(x,y)$ you can see why the tangent plane does not exist at $(0,0,0)$.

[8] "pathological" as in, "your clothes are so pathological, where'd you get them?"



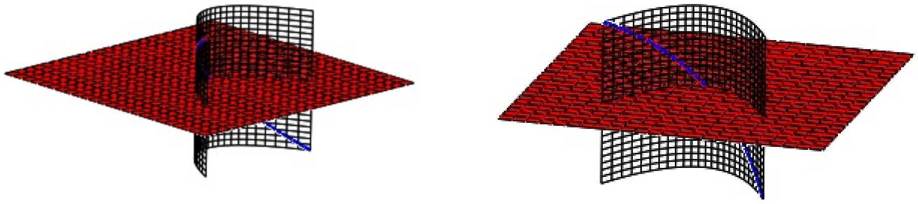
Notice the sides of the box in the picture are parallel to the x and y axes so the path considered below would fall on a diagonal slice of these boxes[9]*. Consider the path to the origin $t \mapsto (t,t)$ gives $f_x(t,t) = 2t^4/(t^2+t^2)^2 = 1/2$ hence $f_x(x,y) \to 1/2$ along the path $t \mapsto (t,t)$, but $f_x(0,0) = 0$ hence the partial derivative f_x is not continuous at $(0,0)$. In this example, the discontinuity of the partial derivatives makes the tangent plane fail to exist.*

One might be tempted to suppose that if a function is continuous at a given point and if all the possible directional derivatives exist then differentiability should follow. It turns out this is not sufficient since continuity of the function does not imply some continuity along the partial derivatives. For example:

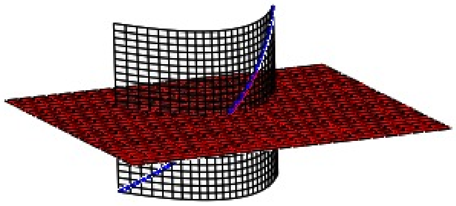
Example 2.4.3. *Let us define $f : \mathbb{R}^2 \to \mathbb{R}$ by $f(x,y) = 0$ for $y \neq x^2$ and $f(x,x^2) = x$. I invite the reader to verify that this function is continuous at the origin. Moreover, consider the directional derivatives at $(0,0)$. We calculate, if $v = \langle a,b \rangle$*

$$D_v f(0,0) = \lim_{h \to 0} \frac{f(0+hv) - f(0)}{h} = \lim_{h \to 0} \frac{f(ah,bh)}{h} = \lim_{h \to 0} \frac{0}{h} = 0.$$

To see why $f(ah,bh) = 0$, consider the intersection of $\vec{r}(h) = (ha,hb)$ and $y = x^2$ the intersection is found at $hb = (ha)^2$ hence, noting $h = 0$ is not of interest in the limit, $b = ha^2$. If $a = 0$ then clearly (ah,bh) only falls on $y = x^2$ at $(0,0)$. If $a \neq 0$ then the solution $h = b/a^2$ gives $f(ha,hb) = ha$ a nontrivial value. However, as $h \to 0$ we eventually reach values close enough to $(0,0)$ that $f(ah,bh) = 0$. Hence we find **all** *directional derivatives exist and are zero at $(0,0)$. Let's examine the graph of this example to see how this happened. The pictures below graph the xy-plane as red and the nontrivial values of f as a blue curve. The union of these forms the graph $z = f(x,y)$.*



[9] the argument to follow stands alone, you don't need to understand the picture to understand the math here, but it's nice if you do



Clearly, f is continuous at $(0,0)$ as I invited you to prove. Moreover, clearly $z = f(x,y)$ cannot be well-approximated by a tangent plane at $(0,0,0)$. If we capture the xy-plane then we lose the blue curve of the graph. On the other hand, if we use a tilted plane then we lose the xy-plane part of the graph.

The moral of the story in the last two examples is simply that derivatives at a point, or even all directional derivatives at a point do not necessarily tell you much about the function near the point. This much is clear: something else is required if the differential is to have meaning which extends beyond one point in a nice way. Therefore, we consider the following:

It would seem the trouble has something to do with discontinuity in the derivative. Continuity of the derivative requires the assignment $a \mapsto dF_a$ is continuous. Or,

$$\lim_{x \to a} dF_x = dF_a. \tag{2.18}$$

But, this is a limit of operators. Let us study this limit in view of the operator norm we discussed in the previous chapter. Let $\epsilon > 0$ then we must be able to find $\delta > 0$ such that $0 < \|x - a\| < \delta$ implies $\|dF_x - dF_a\| < \epsilon$. So, we need to control $\|dF_x - dF_a\|$ to be sure the derivative is continuous. Consider,

$$\begin{aligned}
\|dF_x - dF_a\| &= \sup\{\|(dF_x - dF_a)(u)\| \ : \ \|u\| = 1\} \qquad (2.19)\\
&= \sup\{\|dF_x(u) - dF_a(u)\| \ : \ \|u\| = 1\}\\
&= \sup\left\{\left\|\sum_{i=1}^n u_i \frac{\partial F}{\partial x_i}(x) - \sum_{i=1}^n u_i \frac{\partial F}{\partial x_i}(a)\right\| \ : \ \|u\| = 1\right\}\\
&\leq \sum_{i=1}^n \left\|\frac{\partial F}{\partial x_i}(x) - \frac{\partial F}{\partial x_i}(a)\right\|
\end{aligned}$$

Therefore, the data $\lim_{x \to a} \frac{\partial F}{\partial x_i}(x) = \frac{\partial F}{\partial x_i}(a)$ for $i = 1, \dots, n$ allows us to prove $\lim_{x \to a} dF_x = dF_a$. Naturally, when we teach multivariate calculus the preferred concept does not involve operator norms. Therefore, to be nice to the non-math majors we define:

Definition 2.4.4.

A mapping $F : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ is **continuously differentiable** at $a \in U$ iff the partial derivative mappings D_jF exist on an open set containing a and are continuous at a.

Equation 2.19 shows maps **continuously differentiable** at $x = a$ are those for which the mapping $x \to dF_x$ is a continuous mapping at $x = a$.

The import of the theorem below is that we can build the tangent plane from the Jacobian matrix provided the partial derivatives exist near the point of tangency and are continuous at the point of tangency. This is a very nice result because the concept of the linear mapping is quite abstract but partial differentiation of a given mapping is often easy. The proof that follows here is found in many texts, in particular see C.H. Edwards *Advanced Calculus of Several Variables* on pages 72-73.

Theorem 2.4.5.

> If $F : \mathbb{R}^n \rightarrow \mathbb{R}$ is continuously differentiable at a then F is differentiable at a

Proof: Consider $a+h$ sufficiently close to a that all the partial derivatives of F exist. Furthermore, consider going from a to $a+h$ by traversing a hyper-parallel-piped travelling n-perpendicular paths:

$$\underbrace{a}_{p_o} \rightarrow \underbrace{a+h_1e_1}_{p_1} \rightarrow \underbrace{a+h_1e_1+h_2e_2}_{p_2} \rightarrow \cdots \underbrace{a+h_1e_1+\cdots+h_ne_n}_{p_n} = a+h.$$

Let us denote $p_j = a+b_j$ where clearly b_j ranges from $b_o = 0$ to $b_n = h$ and $b_j = \sum_{i=1}^{j} h_ie_i$. Notice that the difference between p_j and p_{j-1} is given by:

$$p_j - p_{j-1} = a + \sum_{i=1}^{j} h_ie_i - a - \sum_{i=1}^{j-1} h_ie_i = h_je_j$$

Consider then the following identity,

$$F(a+h) - F(a) = F(p_n) - F(p_{n-1}) + F(p_{n-1}) - F(p_{n-2}) + \cdots + F(p_1) - F(p_o)$$

This is to say the change in F from $p_o = a$ to $p_n = a+h$ can be expressed as a sum of the changes along the n-steps. Furthermore, if we consider the difference $F(p_j) - F(p_{j-1})$ you can see that only the j-th component of the argument of F changes. Since the j-th partial derivative exists on the interval for h_j considered by construction we can apply the mean value theorem to locate c_j such that:

$$h_j\partial_jF(p_{j-1}+c_je_j) = F(p_j) - F(p_{j-1})$$

Therefore, using the mean value theorem for each interval, we select $c_1, \dots c_n$ with:

$$F(a+h) - F(a) = \sum_{j=1}^{n} h_j\partial_jF(p_{j-1}+c_je_j)$$

It follows we should propose L to satisfy the definition of Frechet differentation as follows:

$$L(h) = \sum_{j=1}^{n} h_j\partial_jF(a)$$

It is clear that L is linear (in fact, perhaps you recognize this as $L(h) = (\nabla F)(a) \bullet h$). Let us prepare to study the Frechet quotient,

$$\begin{aligned} F(a+h) - F(a) - L(h) &= \sum_{j=1}^{n} h_j\partial_jF(p_{j-1}+c_je_j) - \sum_{j=1}^{n} h_j\partial_jF(a) \\ &= \sum_{j=1}^{n} h_j \underbrace{\left[\partial_jF(p_{j-1}+c_je_j) - \partial_jF(a)\right]}_{g_j(h)} \end{aligned}$$

Observe that $p_{j-1}+c_je_j \rightarrow a$ as $h \rightarrow 0$. Thus, $g_j(h) \rightarrow 0$ by the continuity of the partial derivatives at $x = a$. Finally, consider the Frechet quotient:

$$\lim_{h\rightarrow 0} \frac{F(a+h) - F(a) - L(h)}{||h||} = \lim_{h\rightarrow 0} \frac{\sum_j h_jg_j(h)}{||h||} = \lim_{h\rightarrow 0} \sum_j \frac{h_j}{||h||} g_j(h)$$

Notice $|h_j| \leq ||h||$ hence $\left|\frac{h_j}{||h||}\right| \leq 1$ and

$$0 \leq \left|\frac{h_j}{||h||}g_j(h)\right| \leq |g_j(h)|$$

Apply the squeeze theorem to deduce each term in the sum $\star$ limits to zero. Consquently, $L(h)$ satisfies the Frechet quotient and we have shown that F is differentiable at $x = a$ and the differential is expressed in terms of partial derivatives as expected; $dF_x(h) = \sum_{j=1}^{n} h_j \partial_j F(a)$ □.

Given the result above it is a simple matter to extend the proof to $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$.

Theorem 2.4.6.

If $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is continuously differentiable at a then F is differentiable at a

Proof: If F is continuously differentiable at a then clearly each component function $F^j : \mathbb{R}^n \rightarrow \mathbb{R}$ is continuously differentiable at a. Thus, by Theorem 2.4.5 we have F^j differentiable at a hence

$$\lim_{h\rightarrow 0} \frac{F^j(a+h) - F^j(a) - dF_a^j(h)}{||h||} = 0 \ \text{ for all } j \in \mathbb{N}_m \quad \Rightarrow \quad \lim_{h\rightarrow 0} \frac{F(a+h) - F(a) - dF_a(h)}{||h||} = 0$$

by Theorem 1.3.11. This proves F is differentiable at a □.

2.5 the product rule

When I first wrote notes for advanced calculus I realized I was writing the same argument over and over. The result below is a result. This argument simultaneously covers derivatives of scalar multiplications, matrix multiplications, dot and cross products.

Theorem 2.5.1.

Let W_1, W_2, W_3, V be finite dimensional real normed linear spaces and suppose $U \subseteq V$ is open. Let $\beta = \{r_1, \dots, r_n\}$ be a basis for V with coordinates $x_1, \dots, x_n$. Let $\gamma_1 = \{w_1, \dots, w_{m_1}\}$ be the basis for W_1. Let $\gamma_2 = \{v_1, \dots, v_{m_2}\}$ be the basis for W_2. Let $\gamma_3 = \{\varepsilon_1, \dots, \varepsilon_{m_3}\}$ be the basis for W_3. Assume there exists a product $\star : W_1 \times W_2 \rightarrow W_3$ such that

$$(cx + y) \star z = c(x \star z) + y \star z \qquad \& \qquad x \star (cz + w) = c(x \star z) + x \star w$$

for all $c \in \mathbb{R}$ and $x, y \in W_1$ and $z, w \in W_2$. Then, if $F : U \rightarrow W_1$ and $G : U \rightarrow W_2$ are continuously differentiable at $a \in U$ then $F \star G$ is continuously differentiable at $a \in U$ where $(F \star G)(a) = F(a) \star G(a)$. Moreover, denoting $\partial/\partial x_j$ by ∂_j we have

$$\partial_j(F \star G)(a) = (\partial_j F)(a) \star G(a) + F(a) \star (\partial_j G)(a).$$

Hence, for each $h \in V$,

$$d(F \star G)_a(h) = dF_a(h) \star G(a) + F(a) \star dG_a(h).$$

Proof: assume the notation given in the Theorem and define structure constants $c_{ijk} \in \mathbb{R}$ such that:

$$v_i \star w_j = \sum_{k=1}^{m_3} c_{ijk}\varepsilon_k. \tag{2.20}$$

These constants characterize the nature of the multiplication $\star$. Interestingly, they have little to do with the proof, essentially the play the role of bystanders. Assuming $F : U \to W_1$ and $G : U \to W_2$ are continuously differentiable at a means their component functions $F_1, \dots, F_{m_1} : U \to \mathbb{R}$ with respect to γ_1 and $G_1, \dots, G_{m_2} : U \to \mathbb{R}$ with respect to γ_2 are continuous at a. The component functions of $F \star G$ are naturally related to those of F and G as follows:

$$\begin{aligned} F \star G &= \left(\sum_{i=1}^{m_1} F_i v_i\right) \star \left(\sum_{j=1}^{m_2} G_j w_j\right) \\ &= \sum_{i=1}^{m_1}\sum_{j=1}^{m_2} F_i G_j \left(v_i \star w_j\right) \\ &= \sum_{i=1}^{m_1}\sum_{j=1}^{m_2} F_i G_j \left(\sum_{k=1}^{m_3} c_{ijk}\varepsilon_k\right) \\ &= \sum_{k=1}^{m_3}\left(\sum_{i=1}^{m_1}\sum_{j=1}^{m_2} F_i G_j c_{ijk}\right)\varepsilon_k \end{aligned} \tag{2.21}$$

Thus $F \star G$ has component function $\sum_{i=1}^{m_1}\sum_{j=1}^{m_2} F_i G_j c_{ijk}$. Observe this is the sum of products of continuously differentiable functions at a which is once again continuously differentiable a. Thus $F \star G$ is continuously differentiable at a as it has component functions whose partial derivative functions are continous at a. This becomes explicitly clear if we calculate the partial derivative of $F \star G$ with respect to x_l for points near a,

$$\begin{aligned} \partial_l(F \star G) &= \sum_{k=1}^{m_3} \partial_l\left(\sum_{i=1}^{m_1}\sum_{j=1}^{m_2} F_i G_j c_{ijk}\right)\varepsilon_k && : \partial_l \text{ done componentwise} \\ &= \sum_{k=1}^{m_3}\left(\sum_{i=1}^{m_1}\sum_{j=1}^{m_2} c_{ijk}\partial_l(F_i G_j)\right)\varepsilon_k && : \text{linearity of } \partial_l \\ &= \sum_{k=1}^{m_3}\left(\sum_{i=1}^{m_1}\sum_{j=1}^{m_2} c_{ijk}\left[(\partial_l F_i)G_j + F_i\partial_l G_j\right]\right)\varepsilon_k && : \text{ordinary product rule} \\ &= \sum_{i=1}^{m_1}\sum_{j=1}^{m_2}\sum_{k=1}^{m_3} c_{ijk}(\partial_l F_i)G_j\varepsilon_k + \sum_{i=1}^{m_1}\sum_{j=1}^{m_2}\sum_{k=1}^{m_3} c_{ijk}F_i(\partial_l G_j)\varepsilon_k \\ &= (\partial_l F) \star G + F \star (\partial_l G). \end{aligned} \tag{2.22}$$

where I used the calculation of Equation 2.21 in reverse in order to make the final step. The calculation makes it explicitly clear that the partial derivatives of $F \star G$ are sums and products of continuous functions hence $F \star G$ is continuously differentiable as claimed. Finally, we can construct the differential from partial derivatives: for $h = \sum_{l=1}^n h_l r_l$ calculate:

$$\begin{aligned} d(F \star G)_a(h) &= \sum_{l=1}^n h_l \partial_l (F \star G)(a) \qquad (2.23) \\ &= \sum_{l=1}^n h_l \left[(\partial_l F)(a) \star G(a) + F(a) \star (\partial_l G)(a) \right] \\ &= \left[\sum_{l=1}^n h_l (\partial_l F)(a) \right] \star G(a) + F(a) \star \left[\sum_{l=1}^n h_l (\partial_l G)(a) \right]. \\ &= dF_a(h) \star G(a) + F(a) \star dG_a(h). \end{aligned}$$

This completes the proof. □

Let's unwrap a few common cases of this general product rule. I'll continue to use the W_1, W_2, W_3 and V notation to connect directly to Theorem 2.5.1.

(1.) Set $W_1 = W_2 = W_3 = \mathbb{R}$ and $V = \mathbb{R}$ to produce the usual first semester calculus product rule:

$$\frac{d}{dt}(fg) = \frac{df}{dt}g + f\frac{dg}{dt}.$$

Of course, this was the heart of the proof.

(2.) Set $W_1 = W_2 = W_3 = \mathbb{R}$ and $V = \mathbb{R}^n$ to produce the usual product rule for real-valued functions of several variables:

$$\frac{\partial}{\partial x_i}(fg) = \frac{\partial f}{\partial x_i}g + f\frac{\partial g}{\partial x_i}.$$

(3.) Set $W_1 = \mathbb{R}$ and $W_2 = W_3$ and $V = \mathbb{R}^n$ to produce the usual product rule for a scalar function multiplied on a vector-valued function:

$$\frac{\partial}{\partial x_i}(f\vec{v}) = \frac{\partial f}{\partial x_i}\vec{v} + f\frac{\partial \vec{v}}{\partial x_i}.$$

(4.) Set $W_1 = W_2 = \mathbb{R}^n$ and $W_3 = \mathbb{R}$ and $V = \mathbb{R}$ to produce the product rule for dot-products of paths:

$$\frac{d}{dt}(\vec{v} \bullet \vec{w}) = \frac{d\vec{v}}{dt} \bullet \vec{w} + \vec{v} \bullet \frac{d\vec{w}}{dt}.$$

(5.) Set $W_1 = W_2 = \mathbb{R}^3$ and $W_3 = \mathbb{R}^3$ and $V = \mathbb{R}$ to produce the product rule for cross-products of paths:

$$\frac{d}{dt}(\vec{v} \times \vec{w}) = \frac{d\vec{v}}{dt} \times \vec{w} + \vec{v} \times \frac{d\vec{w}}{dt}.$$

(6.) Set $W_1 = W_2 = W_3 = \mathbb{R}^{\ n \times n}$ and $V = \mathbb{R}$ to produce the product rule for matrix-valued functions of a real variable: $t \mapsto A(t)$, $t \mapsto B(t)$,

$$\frac{d}{dt}(AB) = \frac{dA}{dt}B + A\frac{dB}{dt}.$$

(7.) Set $W_1 = W_2 = W_3 = \mathbb{C}$ and $V = \mathbb{C}$ with $z = x + iy$ we find for $f_1 = u_1 + iv_1$ and $f_2 = u_2 + iv_2$

$$\frac{\partial}{\partial x}(f_1 f_2) = \frac{\partial f_1}{\partial x} f_2 + f_1 \frac{\partial f_2}{\partial x} \qquad \& \qquad \frac{\partial}{\partial y}(f_1 f_2) = \frac{\partial f_1}{\partial y} f_2 + f_1 \frac{\partial f_2}{\partial y}.$$

Of course, there is much more. I simply wish to impress on you that these product rules are **all** simply the standard product rule married to the algebraic structure of the given product. So long as the product has the needed linearity properties, there will be a corresponding product rule for functions.

2.6 higher derivatives

Given normed linear spaces V, W and $U \subseteq V$ open and a differentiable map $F : U \to W$ we find a linear transformation $dF_a : V \to W$ for each $a \in U$. Therefore, we can define the map $f' : U \to \mathcal{L}(V; W)$ by the natural map $a \mapsto dF_a$. That is, $f'(a) = df_a$. Furthermore, since $\mathcal{L}(V; W)$ is itself a normed linear space we may study derivatives of f'. In particular, if $df'_a : V \to \mathcal{L}(V; \mathcal{L}(V; W))$ is linear for each $a \in U$ and satisfies the needed Frechet quotient then we may likewise define $f'' : U \to \mathcal{L}(V; \mathcal{L}(V; W))$ by $f''(a) = (f')'(a) = (df')_a \in \mathcal{L}(V; \mathcal{L}(V; W))$ for each $a \in U$. This all gets a bit meta, so, its helpful to make use of an isomorphism $\Psi : \mathcal{L}(V; \mathcal{L}(V; W)) \to \mathcal{L}(V, V; W)$ defined by:

$$\Psi(T)(x, y) = (T(x))(y) \tag{2.24}$$

for all $x, y \in V$ and $T \in \mathcal{L}(V, W)$. Typically the Ψ is not written. With this abuse of language, we have $f''(a) : V \times V \to W$ given by

$$f''(a)(h, k) = df'_a(h, k) = d(h \mapsto df_h)_a(k) \tag{2.25}$$

Thus, in stark contrast to first semester calculus, each added derivative brings out a new object. Using the isomorphism and its extension to higher derivatives, we find the n-th derivative of $f : V \to W$ is naturally understood as an n-linear map from V to W. What is beautiful is that we can capture this simply in terms of iterated partial derivatives provided a certain continuity is given. I'll attempt to explain this for the case of second derivatives this semester. For the sake of time, I'll let Zorich provide the many details I omit here. If I find time to prepare and Lecture, we may examine the proof that partial derivatives commute. Whether or not we have time for the proof, the fact that partial derivatives commute is a cornerstone of abstract calculus.

2.7 differentiation in an algebra variable

Here I share with you the rudiments of what I have come to call $\mathcal{A}$-calculus. We say $\mathcal{A}$ is an **algebra** if there is a multiplication $\star : \mathcal{A} \times \mathcal{A} \to \mathcal{A}$ which behaves like ordinary multiplication:

(1.) $(x + y) \star z = x \star z + y \star z$ and $x \star (y + z) = x \star y + x \star z$

(2.) $(cx) \star y = x \star (cy) = c(x \star y)$

(3.) $(x \star y) \star z = x \star (y \star z)$

(4.) there exists $1_{\mathcal{A}} \in \mathcal{A}$ for which $1_{\mathcal{A}} \star x = x = x \star 1_{\mathcal{A}}$ for each $x \in \mathcal{A}$

I usually think about **real algebras** which means there is essentially a copy of $\mathbb{R}$ in the center of the algebra. In item (2.) I assume $c \in \mathbb{R}$. However, the $1_{\mathcal{A}}$ may not appear manifestly as $1 \in \mathbb{R}$. Let me give a couple simple examples and forego the general theory.

Example 2.7.1. *$\mathcal{A} = \mathbb{C}$ is a nice example. If $a + ib, c + id \in \mathbb{C}$ then we define $(a + ib)(c + id) = ac - bd + i(ad + bc)$. Equivalently, we could just proclaim $i^2 = -1$ and otherwise, calculate like usual. Here $1_{\mathcal{A}} = 1$ as you might expect*[10]

[10] Using $\mathbb{C} = \mathbb{R}^2$ as a point set we note $1 = (1, 0)$ and $i = (0, 1)$ hence $c_{111} = 1$ and $c_{221} = -1$ and $c_{122} = c_{212} = 1$ whereas all other structure constants are zero. This has not much to do with anything, but, I thought it might be fun given the proof of the previous section

Example 2.7.2. *The* **direct product algebra** *of $\mathcal{A} = \mathbb{R} \times \mathbb{R}$ is defined by $(a, b)(x, y) = (ax, by)$. Here $(1, 1)(x, y) = (x, y)$ for all $(x, y) \in \mathcal{A}$ and in fact $1_{\mathcal{A}} = (1, 1)$.*

Example 2.7.3. *The* **hyperbolic numbers** *are of the form $a + bj$ where $j^2 = 1$. In particular, define $(a + bj)(c + jd) = ac + bd + j(ad + bc)$.*

Example 2.7.4. *The* **3-hyperbolic numbers** *are of the form $a + bj + cj^2$ where $j^3 = 1$. In particular, define*

$$(a + bj + cj^2)(x + jy + j^2z) = ax + by + cz + j(bx + ay + cz) + j^2(cx + by + az).$$

All the algebras I've listed thus far are **commutative**. There are also many noncommutative algebras like the quaternions or matrix algebras. Notice $\mathbb{R}^{\ n\times n}$ forms an algebra. Basically, I think of algebras as **generalized number systems**. So, given that, it is interesting to ask what it means to differentiate with respect to a variable which takes values in $\mathcal{A}$. In fact, we have a whole course devoted to studying what happens when you do calculus with respect to a complex variable. Many schools have such a course. What is less known, which is a shame since it's really pretty simple, is that you can differentiate with respect to an algebra variable in much the same way.

Definition 2.7.5. *Let $U \subseteq \mathcal{A}$ be an open set containing p. If $f : U \to \mathcal{A}$ is a function then we say f is* **$\mathcal{A}$-differentiable at** *p if there exists a linear function $d_pf \in \mathcal{R}_{\mathcal{A}}$ such that*

$$\lim_{h\to 0} \frac{f(p + h) - f(p) - d_pf(h)}{||h||} = 0. \tag{2.26}$$

When I say $d_pf \in \mathcal{R}_{\mathcal{A}}$ this simply means that $d_pf : \mathcal{A} \to \mathcal{A}$ is $\mathbb{R}$-linear mapping on $\mathcal{A}$ and $d_pf(v \star w) = d_pf(v) \star w$ for all $v, w \in \mathcal{A}$. In other words, $\mathcal{A}$-differentiability amounts to differentiability at p with an extra condition. Furthermore, we define the derivative at p as follows:

$$(d_pf)(h) = f'(p)h \tag{2.27}$$

But, since $(d_pf)(h) = d_pf(1 \star h) = d_pf(1) \star h = f'(p)h$ we have $f'(p) = d_pf(1)$. In contrast to the differential of an arbitrary real differentiable map on $\mathcal{A}$, the formula for d_pf is equivalent to the selection of a number in $\mathcal{A}$ for p. In other words, there is a natural manner to interpret the derivative of a function as a function once more. Furthermore, it can be shown for higher derivatives of an $\mathcal{A}$-differentiable function we have

$$d^nf(v_1, v_2, \ldots, v_n) = d^nf(1, 1, \ldots, 1) \star v_1 \star v_2 \star \cdots \star v_n \tag{2.28}$$

So the n-th derivative is also uniquely fixed by the value of $d^nf(1, 1, \ldots, 1)$. In fact, we can naturally identify the n-th derivative of a function as a function once more. In general, the n-th derivative is a symmetric n-linear functon. Finally, I must tell you a beautiful formula which makes $\mathcal{A}$-Calculus so very interesting: provided the basis for $\mathcal{A}$ has $1_{\mathcal{A}} = 1$ paired with coordinate x_1:

$$\frac{\partial^n f}{\partial x_{i_1}\partial x_{i_2}\cdots\partial x_{i_n}} = \frac{\partial^n f}{\partial x_1^n} \star v_{i_1} \star v_{i_2} \star \cdots \star v_{i_n} \tag{2.29}$$

If $\mathcal{A} = \mathbb{R}^n$ as a point set and $e_1 = 1$ then the formulas describing $\mathcal{A}$-calculus are quite nice.

Example 2.7.6. *Consider $f = u + iv$ which is complex differentiable at $p \in \mathbb{C}$. Use $z = x + iy$ as the typical variable in $\mathbb{C}$. Notice, $d_pf(i) = d_pf(1)i$ implies that $\frac{\partial f}{\partial y} = \frac{\partial f}{\partial x}i$. These are the famed* **Cauchy Riemann** *equations. To help the reader make the connection, note $f_y = u_y + iv_y$ and $f_x = u_x + iv_x$ hence $f_y = if_x$ amounts to $(u_y + iv_y) = i(u_x + iv_x)$ hence $u_y = -v_x$ and $v_y = u_x$. Jumping ahead a bit, with no intention of explaining why here, it is fun to note since $i^2 + 1 = 0$ it follows $f_{yy} + f_{xx} = 0$ hence the component functions of a complex differentiable function are solutions to Laplace's equation.*

Example 2.7.7. *Consider $f = u + jv$ which is hyperbolic differentiable at $p \in \mathcal{H} = \mathbb{R} \oplus j\mathbb{R}$ (this is just notation for the hyperbolic numbers). Use $z = x + jy$ as the typical variable in $\mathcal{H}$. Notice, $d_pf(j) = d_pf(1)j$ implies that $\frac{\partial f}{\partial y} = \frac{\partial f}{\partial x}j$. These are the no so well-known hyperbolic Cauchy Riemann equations. To help the reader make the connection, note $f_y = u_y + jv_y$ and $f_x = u_x + jv_x$ hence $f_y = jf_x$ amounts to $(u_y + jv_y) = j(u_x + jv_x)$ hence $u_y = v_x$ and $v_y = u_x$. Jumping ahead a bit, with no intention of explaining why here, it is fun to note since $1 - j^2 = 0$ it follows $f_{xx} - f_{yy} = 0$ hence the component functions of a hyperbolic differentiable function are solutions to the one-dimensional wave equation.*

Basically, any identity which appears amongst the basis elements of an algebra will be mirrored in a PDE which is solve by each function differentiable over the algebra. Most familar case is with $\mathbb{C}$ where harmonic functions are a standard and beatiful topic. But, this is just one of many function theories. In ordinary real analysis essentially $\mathcal{A} = \mathbb{R}$ itself so this feature cannot be seen. However, once $\mathcal{A}$ is two or more dimensional, the differentiability with respect to $\mathcal{A}$ binds real variables together in such a way that the change in one real variable is necessarily coupled to the rest.

Ok, so, let's return to our uber product rule once more, assume f, g are $\mathcal{A}$-differentiable at p in a commutative algebra then note:

$$\begin{aligned} d_p(f \star g)(v \star w) &= (d_pf)(v \star w) \star g(p) + f(p) \star d_pg(v \star w) \qquad (2.30) \\ &= d_pf(v) \star w \star g(p) + f(p) \star d_pg(v) \star w \\ &= (d_pf(v) \star g(p) + f(p) \star d_pg(v)) \star w \end{aligned}$$

We can argue $f \star g$ is real differentiable and $d_p(f \star g) \in \mathcal{R}_{\mathcal{A}}$ thus $f \star g$ is $\mathcal{A}$-differentiable at p. Moreover, as $(f \star g)'(p) = d_p(f \star g)(1)$ we derive from the result above that

$$(f \star g)'(p) = f'(p) \star g(p) + f(p) \star g'(p).$$

Many further results about the calculus over an algebra are known and many resemble closely the calculus you've already seen. However, I've also found a few suprises, mostly thanks to the students who've helped me study $\mathcal{A}$-calculus the past few years. If this section was a bit too terse, my apologies, I have much more to say in my primer on $\mathcal{A}$-calculus: Introduction to $\mathcal{A}$-Calculus and my $\mathcal{A}$-Calculus II paper with Daniel Freese and my differential equations on an algebra paper with Nathan BeDell. I will probably share some tidbits about these papers when the time seems right in this course. But, our main focus is elsewhere.

Chapter 3

inverse and implicit function theorems

It is tempting to give a complete and rigourous proof of these theorems at the outset, but I will resist the temptation in lecture[1]. I'm actually more interested that the student understand what the theorem claims before I show the real proof. I will sketch the proof and show many applications. A nearly complete proof is found in Edwards where he uses an iterative approximation technique founded on the contraction mapping principle, we will go through that a bit later in the course. I probably will not have typed notes on that material this semester, but Edward's is fairly readable and I think we'll profit from working through those sections. That said, we develop an intuition for just what these theorems are all about to start. That is the point of this chapter: to grasp what the linear algebra of the Jacobian suggests about the local behaviour of functions and equations.

3.1 inverse function theorem

Consider the problem of finding a **local** inverse for $f : dom(f) \subseteq \mathbb{R} \rightarrow \mathbb{R}$. If we are given a point $p \in dom(f)$ such that there exists an open interval I containing p with $f|_I$ a one-one function then we can reasonably construct an inverse function by the simple rule $f^{-1}(y) = x$ iff $f(x) = y$ for $x \in I$ and $y \in f(I)$. A sufficient condition to insure the existence of a local inverse is that the derivative function is either strictly positive or strictly negative on some neighborhood of p. If we are give a continuously differentiable function at p then it has a derivative which is continuous on some neighborhood of p. For such a function if $f'(p) \neq 0$ then there exists some interval centered at p for which the derivative is strictly positive or negative. It follows that such a function is strictly monotonic and is hence one-one thus there is a local inverse at p. We should all learn in calculus I that the derivative informs us about the local invertibility of a function. Natural question to ask for us here: does this extend to higher dimensions? If so, how?

The arguments I just made are supported by theorems that are developed in calculus I. Let me shift gears a bit and give a direct calculational explanation based on the linearization approximation.

If $x \approx p$ then $f(x) \approx f(p) + f'(p)(x - p)$. To find the formula for the inverse we solve $y = f(x)$ for x:

$$y \approx f(p) + f'(p)(x - p) \quad \Rightarrow \quad x \approx p + \frac{1}{f'(p)}\big[y - f(p)\big]$$

[1] I have written careful notes which are based largely on Edward's Advanced Calculus text, see Chapter 10 if you wish to see the details, that material assumes deeper mathematical maturity, ideally completion of Math 431 and a healthy dose of good old fashion curiosity

Therefore, $\boxed{f^{-1}(y) \approx p + \frac{1}{f'(p)}\big[y - f(p)\big]}$ for y near $f(p)$.

Example 3.1.1. *Just to help you believe me, consider $f(x) = 3x - 2$ then $f'(x) = 3$ for all x. Suppose we want to find the inverse function near $p = 2$ then the discussion preceding this example suggests,*

$$f^{-1}(y) = 2 + \frac{1}{3}(y - 4).$$

I invite the reader to check that $f(f^{-1}(y)) = y$ and $f^{-1}(f(x)) = x$ for all $x, y \in \mathbb{R}$.

In the example above we found a global inverse exactly, but this is thanks to the linearity of the function in the example. Generally, inverting the linearization just gives the first approximation to the inverse.

Consider $F : dom(F) \subseteq \mathbb{R}^n \to \mathbb{R}^n$. If F is differentiable at $p \in \mathbb{R}^n$ then we can write $F(x) \approx F(p) + F'(p)(x - p)$ for $x \approx p$. Set $y = F(x)$ and solve for x via matrix algebra. This time we need to assume $F'(p)$ is an invertible matrix in order to isolate x,

$$y \approx F(p) + F'(p)(x - p) \quad \Rightarrow \quad x \approx p + (F'(p))^{-1}\big[y - f(p)\big]$$

Therefore,

$$\boxed{F^{-1}(y) \approx p + (F'(p))^{-1}\big[y - f(p)\big]}$$

for y near $F(p)$. Apparently the condition to find a local inverse for a mapping on $\mathbb{R}^n$ is that the derivative matrix is nonsingular[2] in some neighborhood of the point. Experience has taught us from the one-dimensional case that we must insist the derivative is continuous near the point in order to maintain the validity of the approximation.

Recall from calculus II that as we attempt to approximate a function with a power series it takes an infinite series of power functions to recapture the formula exactly. Well, something similar is true here. However, the method of approximation is through an iterative approximation procedure which is built off the idea of Newton's method. The product of this iteration is a nested sequence of composite functions. To prove the theorem below one must actually provide proof the recursively generated sequence of functions converges. See pages 160-187 of Edwards for an in-depth exposition of the iterative approximation procedure. Then see pages 404-411 of Edwards for some material on uniform convergence[3] The main analytical tool which is used to prove the convergence is called the **contraction mapping principle**. The proof of the principle is relatively easy to follow and interestingly the main non-trivial step is an application of the geometric series. For the student of analysis this is an important topic which you should spend considerable time really trying to absorb as deeply as possible. The contraction mapping is at the base of a number of interesting and nontrivial theorems. Read Rosenlicht's *Introduction to Analysis* for a broader and better organized exposition of this analysis. In contrast, Edwards' uses analysis as a tool to obtain results for advanced calculus but his central goal is not a broad or well-framed treatment of analysis. Consequently, if analysis is your interest then you really need to read something else in parallel to get a better ideas about sequences of functions and uniform convergence. I have some notes from a series of conversations with a student about Rosenlicht, I'll post those for the interested student.

[2] nonsingular matrices are also called invertible matrices and a convenient test is that A is invertible iff $det(A) \neq 0$.

[3] actually that later chapter is part of why I chose Edwards' text, he makes a point of proving things in $\mathbb{R}^n$ in such a way that the proof naturally generalizes to function space. This is done by arguing with properties rather than formulas. The properties offen extend to infinite dimensions whereas the formulas usually do not.

These notes focus on the part of the material I require for this course. This is Theorem 3.3 on page 185 of Edwards' text:

Theorem 3.1.2. *(inverse function theorem)*

> Suppose $F : \mathbb{R}^n \to \mathbb{R}^n$ is continuously differentiable in an open set W containing a and the derivative matrix $F'(a)$ is invertible. Then F is locally invertible at a. This means that there exists an open set $U \subseteq W$ containing a and V a open set containing $b = F(a)$ and a one-one, continuously differentiable mapping $G : V \to W$ such that $G(F(x)) = x$ for all $x \in U$ and $F(G(y)) = y$ for all $y \in V$. Moreover, the local inverse G can be obtained as the limit of the sequence of successive approximations defined by
>
> $$G_o(y) = a \;\text{ and }\; G_{n+1}(y) = G_n(y) - [F'(a)]^{-1}[F(G_n(y)) - y] \qquad \text{for all } y \in V.$$

The qualifier local is important to note. If we seek a global inverse then other ideas are needed. If the function is everywhere injective then logically $F(x) = y$ defines $F^{-1}(y) = x$ and F^{-1} so constructed in single-valued by virtue of the injectivity of F. However, for differentiable mappings, one might wonder how can the criteria of global injectivity be tested via the differential. Even in the one-dimensional case a vanishing derivative does not indicate a lack of injectivity; $f(x) = x^3$ has $f^{-1}(y) = \sqrt[3]{y}$ and yet $f'(0) = 0$ (therefore $f'(0)$ is not invertible). One the other hand, we'll see in the examples that follow that even if the derivative is invertible over a set it is possible for the values of the mapping to double-up and once that happens we cannot find a single-valued inverse function[4]

Remark 3.1.3. *James R. Munkres' Analysis on Manifolds good for a different proof.*

> Another good place to read the inverse function theorem is in James R. Munkres *Analysis on Manifolds*. That text is careful and has rather complete arguments which are not entirely the same as the ones given in Edwards. Munkres' text does not use the contraction mapping principle, instead the arguments are more topological in nature.

To give some idea of what I mean by topological let be give an example of such an argument. Suppose $F : \mathbb{R}^n \to \mathbb{R}^n$ is continuously differentiable and $F'(p)$ is invertible. Here's a sketch of the argument that $F'(x)$ is invertible for all x near p as follows:

1. the function $g : \mathbb{R}^n \to \mathbb{R}$ defined by $g(x) = det(F'(x))$ is formed by a multinomial in the component functions of $F'(x)$. This function is clearly continuous since we are given that the partial derivatives of the component functions of F are all continuous.

2. note we are given $F'(p)$ is invertible and hence $det(F'(p)) \neq 0$ thus the continuous function g is nonzero at p. It follows there is some open set U containing p for which $0 \notin g(U)$

3. we have $det(F'(x)) \neq 0$ for all $x \in U$ hence $F'(x)$ is invertible on U.

I would argue this is a topological argument because the key idea here is the continuity of g. Topology is the study of continuity in general.

Example 3.1.4. *Suppose* $F(x, y) = \big(\, \sin(y) + 1,\ \sin(x) + 2 \,\big)$ *for* $(x, y) \in \mathbb{R}^2$. *Clearly* F *is continuously differentiable as all its component functions have continuous partial derivatives. Observe,*

$$F'(x, y) = [\ \partial_x F \mid \partial_y F\ ] = \begin{bmatrix} 0 & \cos(y) \\ \cos(x) & 0 \end{bmatrix}$$

[4] there are scientists and engineers who work with multiply-valued functions with great success, however, as a point of style if nothing else, we try to use functions in math.

Hence $F'(x,y)$ is invertible at points (x,y) such that det$(F'(x,y)) = -\cos(x)\cos(y) \neq 0$. This means we **may** *not be able to find local inverses at points (x,y) with $x = \frac{1}{2}(2n+1)\pi$ or $y = \frac{1}{2}(2m+1)\pi$ for some $m,n \in \mathbb{Z}$. Points where $F'(x,y)$ are singular are points where one or both of $\sin(y)$ and $\sin(x)$ reach extreme values thus the points where the Jacobian matrix are singular are in fact points where we cannot find a local inverse. Why? Because the function is clearly not 1-1 on any set which contains the points of singularity for dF. Continuing, recall from precalculus that sine has a standard inverse on $[-\pi/2,\pi/2]$. Suppose $(x,y) \in [-\pi/2,\pi/2]^2$ and seek to solve $F(x,y) = (a,b)$ for (x,y):*

$$F(x,y) = \left[\begin{array}{c} \sin(y)+1 \\ \sin(x)+2 \end{array}\right] = \left[\begin{array}{c} a \\ b \end{array}\right] \ \Rightarrow \ \left\{\begin{array}{c} \sin(y)+1 = a \\ \sin(x)+2 = b \end{array}\right\} \ \Rightarrow \ \left\{\begin{array}{c} y = \sin^{-1}(a-1) \\ x = \sin^{-1}(b-2) \end{array}\right\}$$

It follows that $F^{-1}(a,b) = \big(\sin^{-1}(b-2),\ \sin^{-1}(a-1)\big)$ for $(a,b) \in [0,2]\times[1,3]$ where you should note $F([-\pi/2,\pi/2]^2) = [0,2]\times[1,3]$. We've found a local inverse for F on the region $[-\pi/2,\pi/2]^2$. In other words, we just found a global inverse for the restriction of F to $[-\pi/2,\pi/2]^2$. Technically we ought not write F^{-1}, to be more precise we should write:

$$(F|_{[-\pi/2,\pi/2]^2})^{-1}(a,b) = \big(\sin^{-1}(b-2),\ \sin^{-1}(a-1)\big).$$

It is customary to avoid such detail in many contexts. Inverse functions for sine, cosine, tangent etc... are good examples of this slight of langauge.

A **coordinate system** on $\mathbb{R}^n$ is an invertible mapping of $\mathbb{R}^n$ to $\mathbb{R}^n$. However, in practice the term coordinate system is used with less rigor. Often a coordinate system has various degeneracies. For example, in polar coordinates you could say $\theta = \pi/4$ or $\theta = 9\pi/4$ or generally $\theta = 2\pi k + \pi/4$ for any $k \in \mathbb{Z}$. Let's examine polar coordinates in view of the inverse function theorem.

Example 3.1.5. *Let $T(r,\theta) = \big(\ r\cos(\theta),\ r\sin(\theta)\ \big)$ for $(r,\theta) \in [0,\infty)\times(-\pi/2,\pi/2)$. Clearly T is continuously differentiable as all its component functions have continuous partial derivatives. To find the inverse we seek to solve $T(r,\theta) = (x,y)$ for (r,θ). Hence, consider $x = r\cos(\theta)$ and $y = r\sin(\theta)$. Note that*

$$x^2+y^2 = r^2\cos^2(\theta) + r^2\sin^2(\theta) = r^2(\cos^2(\theta)+\sin^2(\theta)) = r^2$$

and

$$\frac{y}{x} = \frac{r\sin(\theta)}{r\cos(\theta)} = \tan(\theta).$$

It follows that $r = \sqrt{x^2+y^2}$ and $\theta = \tan^{-1}(y/x)$ for $(x,y) \in (0,\infty)\times\mathbb{R}$. We find

$$T^{-1}(x,y) = \Big(\ \sqrt{x^2+y^2},\ \tan^{-1}(y/x)\ \Big).$$

Let's see how the derivative fits with our results. Calcuate,

$$T'(r,\theta) = [\ \partial_r T \mid \partial_\theta T\ ] = \left[\begin{array}{cc} \cos(\theta) & -r\sin(\theta) \\ \sin(\theta) & r\cos(\theta) \end{array}\right]$$

note that det$(T'(r,\theta)) = r$ hence we the inverse function theorem provides the existence of a local inverse around any point except the origin. Notice the derivative does not detect the defect in the angular coordinate. Challenge, find the inverse function for $T(r,\theta) = \big(\ r\cos(\theta),\ r\sin(\theta)\ \big)$ with dom$(T) = [0,\infty)\times(\pi/2,3\pi/2)$. Or, find the inverse for polar coordinates in a neighborhood of $(0,-1)$.

Example 3.1.6. *Suppose $T : \mathbb{R}^3 \to \mathbb{R}^3$ is defined by $T(x,y,z) = (ax, by, cz)$ for constants $a,b,c \in \mathbb{R}$ where $abc \neq 0$. Clearly T is continuously differentiable as all its component functions have continuous partial derivatives. We calculate $T'(x,y,z) = [\partial_x T|\partial_y T|\partial_z T] = [ae_1|be_2|ce_3]$. Thus $det(T'(x,y,z)) = abc \neq 0$ for all $(x,y,z) \in \mathbb{R}^3$ hence this function is locally invertible everywhere. Moreover, we calculate the inverse mapping by solving $T(x,y,z) = (u,v,w)$ for (x,y,z):*

$$(ax, by, cz) = (u,v,w) \quad \Rightarrow \quad (x,y,z) = (u/a, v/b, w/c) \quad \Rightarrow \quad \boxed{T^{-1}(u,v,w) = (u/a, v/b, w/c).}$$

Example 3.1.7. *Suppose $F : \mathbb{R}^n \to \mathbb{R}^n$ is defined by $F(x) = Ax+b$ for some matrix $A \in \mathbb{R}^{\ n\times n}$ and vector $b \in \mathbb{R}^n$.* **Under what conditions is such a function invertible ?***. Since the formula for this function gives each component function as a polynomial in the n-variables we can conclude the function is continuously differentiable. You can calculate that $F'(x) = A$. It follows that a sufficient condition for local inversion is $det(A) \neq 0$. It turns out that this is also a necessary condition as $det(A) = 0$ implies the matrix A has nontrivial solutions for $Av = 0$. We say $v \in Null(A)$ iff $Av = 0$. Note if $v \in Null(A)$ then $F(v) = Av + b = b$. This is not a problem when $det(A) \neq 0$ for in that case the null space is contains just zero; $Null(A) = \{0\}$. However, when $det(A) = 0$ we learn in linear algebra that $Null(A)$ contains infinitely many vectors so F is far from injective. For example, suppose $Null(A) = span\{e_1\}$ then you can show that $F(a_1, a_2, \dots, a_n) = F(x, a_2, \dots, a_n)$ for all $x \in \mathbb{R}$. Hence any point will have other points nearby which output the same value under F. Suppose $det(A) \neq 0$, to calculate the inverse mapping formula we should solve $F(x) = y$ for x,*

$$y = Ax + b \quad \Rightarrow \quad x = A^{-1}(y-b) \quad \Rightarrow \quad \boxed{F^{-1}(y) = A^{-1}(y-b).}$$

Remark 3.1.8. *inverse function theorem holds for higher derivatives.*

In Munkres the inverse function theorem is given for r-times differentiable functions. In short, a C^r function with invertible differential at a point has a C^r inverse function local to the point. Edwards also has arguments for $r > 1$, see page 202 and arguments and surrounding arguments.

3.2 implicit function theorem

Consider the problem of solving $x^2 + y^2 = 1$ for y as a function of x.

$$x^2 + y^2 = 1 \quad \Rightarrow \quad y^2 = 1 - x^2 \quad \Rightarrow \quad y = \pm\sqrt{1-x^2}.$$

A function cannot have two outputs for a single input, when we write ± in the expression above it simply indicates our ignorance as to which is chosen. Once further information is given then we may be able to choose a $+$ or a $-$. For example:

1. if $x^2 + y^2 = 1$ and we want to solve for y near $(0,1)$ then $y = \sqrt{1-x^2}$ is the correct choice since $y > 0$ at the point of interest.

2. if $x^2 + y^2 = 1$ and we want to solve for y near $(0,-1)$ then $y = -\sqrt{1-x^2}$ is the correct choice since $y < 0$ at the point of interest.

3. if $x^2 + y^2 = 1$ and we want to solve for y near $(1,0)$ then it's impossible to find a single function which reproduces $x^2 + y^2 = 1$ on an **open disk** centered at $(1,0)$.

What is the defect of case (3.) ? The trouble is that no matter how close we zoom in to the point there are always two y-values for each given x-value. Geometrically, this suggests either we have a discontinuity, a kink, or a vertical tangent in the graph. The given problem has a vertical tangent and hopefully you can picture this with ease since its just the unit-circle. In calculus I we studied implicit differentiation, our starting point was to assume $y = y(x)$ and then we differentiated equations to work out implicit formulas for dy/dx. Take the unit-circle and differentiate both sides,

$$x^2 + y^2 = 1 \quad \Rightarrow \quad 2x + 2y\frac{dy}{dx} = 0 \quad \Rightarrow \quad \frac{dy}{dx} = -\frac{x}{y}.$$

Note $\frac{dy}{dx}$ is not defined for $y = 0$. It's no accident that those two points $(-1, 0)$ and $(1, 0)$ are precisely the points at which we cannot solve for y as a function of x. Apparently, the singularity in the derivative indicates where we may have trouble solving an equation for one variable as a function of the remaining variable.

We wish to study this problem in general. Given n-equations in $(m+n)$-unknowns when can we solve for the last n-variables as functions of the first m-variables ? Given a continuously differentiable mapping $G = (G_1, G_2, \dots, G_n) : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^n$ study the level set: (here $k_1, k_2, \dots, k_n$ are constants)

$$\begin{aligned} G_1(x_1, \dots, x_m, y_1, \dots, y_n) &= k_1 \\ G_2(x_1, \dots, x_m, y_1, \dots, y_n) &= k_2 \\ &\vdots \\ G_n(x_1, \dots, x_m, y_1, \dots, y_n) &= k_n \end{aligned}$$

We wish to locally solve for $y_1, \dots, y_n$ as functions of $x_1, \dots x_m$. That is, find a mapping $h : \mathbb{R}^m \to \mathbb{R}^n$ such that $G(x, y) = k$ iff $y = h(x)$ near some point $(a, b) \in \mathbb{R}^m \times \mathbb{R}^n$ such that $G(a, b) = k$. In this section we use the notation $x = (x_1, x_2, \dots x_m)$ and $y = (y_1, y_2, \dots, y_n)$.

Before we turn to the general problem let's analyze the unit-circle problem in this notation. We are given $G(x, y) = x^2 + y^2$ and we wish to find $f(x)$ such that $y = f(x)$ solves $G(x, y) = 1$. Differentiate with respect to x and use the chain-rule:

$$\frac{\partial G}{\partial x}\frac{dx}{dx} + \frac{\partial G}{\partial y}\frac{dy}{dx} = 0$$

We find that $\boxed{dy/dx = -G_x/G_y} = -x/y$. Given this analysis we should suspect that if we are given some level curve $G(x, y) = k$ then we may be able to solve for y as a function of x near p if $G(p) = k$ and $G_y(p) \neq 0$. This suspicion is valid and it is one of the many consequences of the implicit function theorem.

We again turn to the linearization approximation. Suppose $G(x, y) = k$ where $x \in \mathbb{R}^m$ and $y \in \mathbb{R}^n$ and suppose $G : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^n$ is continuously differentiable. Suppose $(a, b) \in \mathbb{R}^m \times \mathbb{R}^n$ has $G(a, b) = k$. Replace G with its linearization based at (a, b):

$$G(x, y) \approx k + G'(a, b)(x - a, y - b)$$

here we have the matrix multiplication of the $n \times (m + n)$ matrix $G'(a, b)$ with the $(m + n) \times 1$ column vector $(x - a, y - b)$ to yield an n-component column vector. It is convenient to define partial derivatives with respect to a whole vector of variables,

$$\frac{\partial G}{\partial x} = \begin{bmatrix} \frac{\partial G_1}{\partial x_1} & \cdots & \frac{\partial G_1}{\partial x_m} \\ \vdots & & \vdots \\ \frac{\partial G_n}{\partial x_1} & \cdots & \frac{\partial G_n}{\partial x_m} \end{bmatrix} \qquad \frac{\partial G}{\partial y} = \begin{bmatrix} \frac{\partial G_1}{\partial y_1} & \cdots & \frac{\partial G_1}{\partial y_n} \\ \vdots & & \vdots \\ \frac{\partial G_n}{\partial y_1} & \cdots & \frac{\partial G_n}{\partial y_n} \end{bmatrix}$$

In this notation we can write the $n \times (m+n)$ matrix $G'(a,b)$ as the concatenation of the $n \times m$ matrix $\frac{\partial G}{\partial x}(a,b)$ and the $n \times n$ matrix $\frac{\partial G}{\partial y}(a,b)$

$$G'(a,b) = \left[\frac{\partial G}{\partial x}(a,b) \Big| \frac{\partial G}{\partial y}(a,b) \right]$$

Therefore, for points close to (a,b) we have:

$$G(x,y) \approx k + \frac{\partial G}{\partial x}(a,b)(x-a) + \frac{\partial G}{\partial y}(a,b)(y-b)$$

The nonlinear problem $G(x,y) = k$ has been (locally) replaced by the linear problem of solving what follows for y:

$$k \approx k + \frac{\partial G}{\partial x}(a,b)(x-a) + \frac{\partial G}{\partial y}(a,b)(y-b) \tag{3.1}$$

Suppose the square matrix $\frac{\partial G}{\partial y}(a,b)$ is invertible at (a,b) then we find the following approximation for the implicit solution of $G(x,y) = k$ for y as a function of x:

$$y = b - \left[\frac{\partial G}{\partial y}(a,b) \right]^{-1} \left[\frac{\partial G}{\partial x}(a,b)(x-a) \right].$$

Of course this is not a formal proof, but it does suggest that $det\big[\frac{\partial G}{\partial y}(a,b)\big] \neq 0$ is a necessary condition for solving for the y variables.

As before suppose $G : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^n$. Suppose we have a continuously differentiable function $h : \mathbb{R}^m \to \mathbb{R}^n$ such that $h(a) = b$ and $G(x, h(x)) = k$. We seek to find the derivative of h in terms of the derivative of G. This is a generalization of the implicit differentiation calculation we perform in calculus I. I'm including this to help you understand the notation a bit more before I state the implicit function theorem. Differentiate with respect to x_l for $l \in \mathbb{N}_m$:

$$\frac{\partial}{\partial x_l}\Big[G(x,h(x)) \Big] = \sum_{i=1}^{m} \frac{\partial G}{\partial x_i}\frac{\partial x_i}{\partial x_l} + \sum_{j=1}^{n} \frac{\partial G}{\partial y_j}\frac{\partial h_j}{\partial x_l} = \frac{\partial G}{\partial x_l} + \sum_{j=1}^{n} \frac{\partial G}{\partial y_j}\frac{\partial h_j}{\partial x_l} = 0$$

we made use of the identity $\frac{\partial x_i}{\partial x_k} = \delta_{ik}$ to squash the sum of i to the single nontrivial term and the zero on the r.h.s follows from the fact that $\frac{\partial}{\partial x_l}(k) = 0$. Concatenate these derivatives from $k = 1$ up to $k = m$:

$$\left[\frac{\partial G}{\partial x_1} + \sum_{j=1}^{n} \frac{\partial G}{\partial y_j}\frac{\partial h_j}{\partial x_1} \Big| \frac{\partial G}{\partial x_2} + \sum_{j=1}^{n} \frac{\partial G}{\partial y_j}\frac{\partial h_j}{\partial x_2} \Big| \cdots \Big| \frac{\partial G}{\partial x_m} + \sum_{j=1}^{n} \frac{\partial G}{\partial y_j}\frac{\partial h_j}{\partial x_m} \right] = [0|0|\cdots|0]$$

Properties of matrix addition allow us to parse the expression above as follows:

$$\left[\frac{\partial G}{\partial x_1} \Big| \frac{\partial G}{\partial x_2} \Big| \cdots \Big| \frac{\partial G}{\partial x_m} \right] + \left[\sum_{j=1}^{n} \frac{\partial G}{\partial y_j}\frac{\partial h_j}{\partial x_1} \Big| \sum_{j=1}^{n} \frac{\partial G}{\partial y_j}\frac{\partial h_j}{\partial x_2} \Big| \cdots \Big| \sum_{j=1}^{n} \frac{\partial G}{\partial y_j}\frac{\partial h_j}{\partial x_m} \right] = [0|0|\cdots|0]$$

But, this reduces to

$$\frac{\partial G}{\partial x} + \left[\frac{\partial G}{\partial y}\frac{\partial h}{\partial x_1} \bigg| \frac{\partial G}{\partial y}\frac{\partial h}{\partial x_2} \bigg| \cdots \bigg| \frac{\partial G}{\partial y}\frac{\partial h}{\partial x_m} \right] = 0 \in \mathbb{R}^{m \times n}$$

The concatenation property of matrix multiplication states $[Ab_1|Ab_2|\cdots|Ab_m] = A[b_1|b_2|\cdots|b_m]$ we use this to write the expression once more,

$$\frac{\partial G}{\partial x} + \frac{\partial G}{\partial y}\left[\frac{\partial h}{\partial x_1} \bigg| \frac{\partial h}{\partial x_2} \bigg| \cdots \bigg| \frac{\partial h}{\partial x_m} \right] = 0 \quad \Rightarrow \quad \frac{\partial G}{\partial x} + \frac{\partial G}{\partial y}\frac{\partial h}{\partial x} = 0 \quad \Rightarrow \quad \boxed{\frac{\partial h}{\partial x} = -\frac{\partial G}{\partial y}^{-1}\frac{\partial G}{\partial x}}$$

where in the last implication we made use of the assumption that $\frac{\partial G}{\partial y}$ is invertible.

Theorem 3.2.1. *(Theorem 3.4 in Edwards's Text see pg 190)*

Let $G : dom(G) \subseteq \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^n$ be continuously differentiable in a open ball about the point (a, b) where $G(a, b) = k$ (a constant vector in $\mathbb{R}^n$). If the matrix $\frac{\partial G}{\partial y}(a, b)$ is invertible then there exists an open ball U containing a in $\mathbb{R}^m$ and an open ball W containing (a, b) in $\mathbb{R}^m \times \mathbb{R}^n$ and a continuously differentiable mapping $h : U \to \mathbb{R}^n$ such that $G(x, y) = k$ iff $y = h(x)$ for all $(x, y) \in W$. Moreover, the mapping h is the limit of the sequence of successive approximations defined inductively below

$$h_o(x) = b, \quad h_{n+1} = h_n(x) - [\tfrac{\partial G}{\partial y}(a, b)]^{-1}G(x, h_n(x)) \qquad \text{for all } x \in U.$$

We will not attempt a proof of the last sentence for the same reasons we did not pursue the details in the inverse function theorem. However, we have already derived the first step in the iteration in our study of the linearization solution.

Proof: Let $G : dom(G) \subseteq \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^n$ be continuously differentiable in a open ball B about the point (a, b) where $G(a, b) = k$ ($k \in \mathbb{R}^n$ a constant). Furthermore, assume the matrix $\frac{\partial G}{\partial y}(a, b)$ is invertible. We seek to use the inverse function theorem to prove the implicit function theorem. Towards that end consider $F : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^m \times \mathbb{R}^n$ defined by $F(x, y) = (x, G(x, y))$. To begin, observe that F is continuously differentiable in the open ball B which is centered at (a, b) since G and x have continuous partials of their components in B. Next, calculate the derivative of $F = (x, G)$,

$$F'(x, y) = [\partial_x F | \partial_y F] = \left[\begin{array}{c|c} \partial_x x & \partial_y x \\ \hline \partial_x G & \partial_y G \end{array} \right] = \left[\begin{array}{c|c} I_m & 0_{m \times n} \\ \hline \partial_x G & \partial_y G \end{array} \right]$$

The determinant of the matrix above is the product of the deteminant of the blocks I_m and $\partial_y G$; $det(F'(x, y) = det(I_m)det(\partial_y G) = \partial_y G$. We are given that $\frac{\partial G}{\partial y}(a, b)$ is invertible and hence $det(\frac{\partial G}{\partial y}(a, b)) \neq 0$ thus $det(F'(x, y) \neq 0$ and we find $F'(a, b)$ is invertible. Consequently, the inverse function theorem applies to the function F at (a, b). Therefore, there exists $F^{-1} : V \subseteq \mathbb{R}^m \times \mathbb{R}^n \to U \subseteq \mathbb{R}^m \times \mathbb{R}^n$ such that F^{-1} is continuously differentiable. Note $(a, b) \in U$ and V contains the point $F(a, b) = (a, G(a, b)) = (a, k)$.

Our goal is to find the implicit solution of $G(x, y) = k$. We know that

$$F^{-1}(F(x, y)) = (x, y) \qquad \text{and} \qquad F(F^{-1}(u, v)) = (u, v)$$

for all $(x, y) \in U$ and $(u, v) \in V$. As usual to find the formula for the inverse we can solve $F(x, y) = (u, v)$ for (x, y) this means we wish to solve $(x, G(x, y)) = (u, v)$ hence $x = u$. The formula for v is more elusive, but we know it exists by the inverse function theorem. Let's say $y = H(u,v)$ where $H : V \to \mathbb{R}^n$ and thus $F^{-1}(u,v) = (u, H(u,v))$. Consider then,

$$(u,v) = F(F^{-1}(u,v) = F(u, H(u,v)) = (u, G(u, H(u,v))$$

Let $v = k$ thus $(u,k) = (u, G(u, H(u,k))$ for all $(u,v) \in V$. Finally, define $h(u) = H(u,k)$ for all $(u,k) \in V$ and note that $k = G(u, h(u))$. In particular, $(a,k) \in V$ and at that point we find $h(a) = H(a,k) = b$ by construction. It follows that $y = h(x)$ provides a continuously differentiable solution of $G(x,y) = k$ near (a,b).

Uniqueness of the solution follows from the uniqueness for the limit of the sequence of functions described in Edwards' text on page 192. However, other arguments for uniqueness can be offered, independent of the iterative method, for instance: see page 75 of Munkres *Analysis on Manifolds.* □

Remark 3.2.2. *notation and the implementation of the implicit function theorem.*

We assumed the variables y were to be written as functions of x variables to make explicit a local solution to the equation $G(x,y) = k$. This ordering of the variables is convenient to argue the proof, however the real theorem is far more general. We can select any subset of n input variables to make up the "y" so long as $\frac{\partial G}{\partial y}$ is invertible. I will use this generalization of the formal theorem in the applications that follow. Moreover, the notations x and y are unlikely to maintain the same interpretation as in the previous pages. Finally, we will for convenience make use of the notation $y = y(x)$ to express the existence of a function f such that $y = f(x)$ when appropriate. Also, $z = z(x,y)$ means there is some function h for which $z = h(x,y)$. If this notation confuses then invent names for the functions in your problem.

Example 3.2.3. *Suppose $G(x,y,z) = x^2 + y^2 + z^2$. Suppose we are given a point (a,b,c) such that $G(a,b,c) = R^2$ for a constant R.* **Problem: For which variable can we solve? What, if any, influence does the given point have on our answer?** *Solution: to begin, we have one equation and three unknowns so we should expect to find one of the variables as functions of the remaining two variables. The implicit function theorem applies as G is continuously differentiable.*

1. *if we wish to solve $z = z(x,y)$ then we need $G_z(a,b,c) = 2c \neq 0$.*
2. *if we wish to solve $y = y(x,z)$ then we need $G_y(a,b,c) = 2b \neq 0$.*
3. *if we wish to solve $x = x(y,z)$ then we need $G_x(a,b,c) = 2a \neq 0$.*

The point has no local solution for z if it is a point on the intersection of the xy-plane and the sphere $G(x,y,z) = R^2$. Likewise, we cannot solve for $y = y(x,z)$ on the $y = 0$ slice of the sphere and we cannot solve for $x = x(y,z)$ on the $x = 0$ slice of the sphere.

Notice, algebra verifies the conclusions we reached via the implicit function theorem:

$$z = \pm\sqrt{R^2 - x^2 - y^2} \qquad y = \pm\sqrt{R^2 - x^2 - z^2} \qquad x = \pm\sqrt{R^2 - y^2 - z^2}$$

When we are at zero for one of the coordinates then we cannot choose + or − since we need both on an open ball intersected with the sphere centered at such a point[5]. Remember, when I talk about local solutions I mean solutions which exist over the intersection of the solution set and an open

[5] if you consider $G(x,y,z) = R^2$ as a space then the open sets on the space are taken to be the intersection with the space and open balls in $\mathbb{R}^3$. This is called the subspace topology in topology courses.

ball in the ambient space ($\mathbb{R}^3$ in this context). The preceding example is the natural extension of the unit-circle example to $\mathbb{R}^3$. A similar result is available for the n-sphere in $\mathbb{R}^n$. I hope you get the point of the example, if we have one equation then if we wish to solve for a particular variable in terms of the remaining variables then all we need is continuous differentiability of the level function and a nonzero partial derivative at the point where we wish to find the solution. Now, the implicit function theorem doesn't find the solution for us, but it does provide the existence. In the section on implicit differentiation, existence is really all we need since focus our attention on rates of change rather than actually solutions to the level set equation.

Example 3.2.4. *Consider the equation* $e^{xy} + z^3 - xyz = 2$. **Can we solve this equation for** $z = z(x, y)$ **near** $(0, 0, 1)$**?** *Let* $G(x, y, z) = e^{xy} + z^3 - xyz$ *and note* $G(0, 0, 1) = e^0 + 1 + 0 = 2$ *hence* $(0, 0, 1)$ *is a point on the solution set* $G(x, y, z) = 2$*. Note* G *is clearly continuously differentiable and*

$$G_z(x, y, z) = 3z^2 - xy \quad \Rightarrow \quad G_z(0, 0, 1) = 3 \neq 0$$

therefore, there exists a continuously differentiable function $h : dom(h) \subseteq \mathbb{R}^2 \to \mathbb{R}$ *which solves* $G(x, y, h(x, y)) = 2$ *for* (x, y) *near* $(0, 0)$ *and* $h(0, 0) = 1$.

I'll not attempt an explicit solution for the last example.

Example 3.2.5. *Let* $(x, y, z) \in S$ *iff* $x + y + z = 2$ *and* $y + z = 1$. **Problem: For which variable(s) can we solve?** *Solution: define* $G(x, y, z) = (x + y + z, y + z)$ *we wish to study* $G(x, y, z) = (2, 1)$*. Notice the solution set is not empty since* $G(1, 0, 1) = (1 + 0 + 1, 0 + 1) = (2, 1)$ *Moreover,* G *is continuously differentiable. In this case we have two equations and three unknowns so we expect two variables can be written in terms of the remaining free variable. Let's examine the derivative of* G*:*

$$G'(x, y, z) = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}$$

Suppose we wish to solve $x = x(z)$ *and* $y = y(z)$ *then we should check invertiblility of*[6]

$$\frac{\partial G}{\partial(x, y)} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}.$$

The matrix above is invertible hence the implicit function theorem applies and we can solve for x *and* y *as functions of* z*. On the other hand, if we tried to solve for* $y = y(x)$ *and* $z = z(x)$ *then we'll get no help from the implicit function theorem as the matrix*

$$\frac{\partial G}{\partial(y, z)} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}.$$

is not invertible. Geometrically, we can understand these results from noting that $G(x, y, z) = (2, 1)$ *is the intersection of the plane* $x + y + z = 2$ *and* $y + z = 1$*. Substituting* $y + z = 1$ *into* $x + y + z = 2$ *yields* $x + 1 = 2$ *hence* $x = 1$ *on the line of intersection. We can hardly use* x *as a free variable for the solution when the problem fixes* x *from the outset.*

The method I just used to analyze the equations in the preceding example was a bit adhoc. In linear algebra we do much better for systems of linear equations. A procedure called Gaussian elimination naturally reduces a system of equations to a form in which it is manifestly obvious how

[6] this notation should not be confused with $\frac{\partial(x,y)}{\partial(u,v)}$ which is used to denote a particular determinant associated with coordinate change of integrals, or pull-back of a differential form as explained on page 100 of H.M Edward's Advanced Calculus: A differential Forms Approach, we should discuss it in a later chapter.

to eliminate redundant variables in terms of a minimal set of basic free variables. The "y" of the implicit function proof discussions plays the role of the so-called **pivotal variables** whereas the "x" plays the role of the remaining **free variables**. These variables are generally intermingled in the list of total variables so to reproduce the pattern assumed for the implicit function theorem we would need to relabel variables from the outset of a calculation. In the following example, I show how reordering the variables allows us to solve for various pairs. In short, put the dependent variable first and the independent variables second so the Gaussian elimination shows the solution with minimal effort. Here's how:

Example 3.2.6. *Consider $G(x, y, u, v) = (3x + 2y - u,\ 2x + y - v) = (-1, 3)$. We have two equations with four variables. Let's investigate which pairs of variables can be taken as independent or dependent variables. The most efficient method to dispatch these questions is probably Gaussian elimination. I leave it to the reader to verify that:*

$$rref\left[\begin{array}{cccc|c} 3 & 2 & -1 & 0 & -1 \\ 2 & 1 & 0 & -1 & 3 \end{array}\right] = \left[\begin{array}{cccc|c} 1 & 0 & 1 & -2 & 7 \\ 0 & 1 & -2 & 3 & -11 \end{array}\right]$$

We can immediately read from the result above that x, y can be taken to depend on u, v via the formulas:

$$x = -u + 2v + 7, \qquad y = 2u - 3v - 11$$

On the other hand, if we order the variables (u, v, x, y) then Gaussian elimination gives:

$$rref\left[\begin{array}{cccc|c} -1 & 0 & 3 & 2 & -1 \\ 0 & -1 & 2 & 1 & 3 \end{array}\right] = \left[\begin{array}{cccc|c} 1 & 0 & -3 & -2 & 1 \\ 0 & 1 & -2 & -1 & -3 \end{array}\right]$$

Therefore, we find $u(x, y)$ and $v(x, y)$ as follows:

$$u = 3x + 2y + 1, \qquad v = 2x + y - 3.$$

To solve for x, u as functions of y, v consider:

$$rref\left[\begin{array}{cccc|c} 3 & -1 & 2 & 0 & -1 \\ 2 & 0 & 1 & -1 & 3 \end{array}\right] = \left[\begin{array}{cccc|c} 1 & 0 & 1/2 & -1/2 & 3/2 \\ 0 & 1 & -1/2 & -3/2 & 11/2 \end{array}\right]$$

From which we can read,

$$x = -y/2 + v/2 + 3/2, \qquad u = y/2 + 3v/2 + 11/2.$$

I could solve the problem below in the efficient style above, but I will instead follow the method in which we discussed in the paragraphs surrounding Equation 3.1. In contrast to the general case, because the problem is linear the solution of Equation 3.1 is also a solution of the actual problem.

Example 3.2.7. *Solve the following system of equations near (1,2,3,4,5).*

$$G(x, y, z, a, b) = \begin{bmatrix} x + y + z + 2a + 2b \\ x + 0 + 2z + 2a + 3b \\ 3x + 2y + z + 3a + 4b \end{bmatrix} = \begin{bmatrix} 24 \\ 30 \\ 42 \end{bmatrix}$$

Differentiate to find the Jacobian:

$$G'(x, y, z, a, b) = \begin{bmatrix} 1 & 1 & 1 & 2 & 2 \\ 1 & 0 & 2 & 2 & 3 \\ 3 & 2 & 1 & 3 & 4 \end{bmatrix}$$

Let us solve $G(x, y, z, a, b) = (24, 30, 42)$ for $x(a, b), y(a, b), z(a, b)$ by the method of Equation 3.1. I'll omit the point-dependence of the Jacobian since it clearly has none.

$$G(x, y, z, a, b) = \begin{bmatrix} 24 \\ 30 \\ 42 \end{bmatrix} + \frac{\partial G}{\partial(x, y, z)} \begin{bmatrix} x-1 \\ y-2 \\ z-3 \end{bmatrix} + \frac{\partial G}{\partial(a, b)} \begin{bmatrix} a-4 \\ b-5 \end{bmatrix}$$

Let me make the notational chimera above explicit:

$$G(x, y, z, a, b) = \begin{bmatrix} 24 \\ 30 \\ 42 \end{bmatrix} + \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 2 \\ 3 & 2 & 1 \end{bmatrix} \begin{bmatrix} x-1 \\ y-2 \\ z-3 \end{bmatrix} + \begin{bmatrix} 2 & 2 \\ 2 & 3 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} a-4 \\ b-5 \end{bmatrix}$$

To solve $G(x, y, z, a, b) = (24, 30, 42)$ for (x, y, z) we may use the expression above. After a little calculation one finds:

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 2 \\ 3 & 2 & 1 \end{bmatrix}^{-1} = \frac{1}{3} \begin{bmatrix} -4 & 1 & 2 \\ 5 & -2 & -1 \\ 2 & 1 & -1 \end{bmatrix}$$

The constant term cancels and we find:

$$\begin{bmatrix} x-1 \\ y-2 \\ z-3 \end{bmatrix} = -\frac{1}{3} \begin{bmatrix} -4 & 1 & 2 \\ 5 & -2 & -1 \\ 2 & 1 & -1 \end{bmatrix} \begin{bmatrix} 2 & 2 \\ 2 & 3 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} a-4 \\ b-5 \end{bmatrix}$$

Multiplying the matrices gives:

$$\begin{bmatrix} x-1 \\ y-2 \\ z-3 \end{bmatrix} = -\frac{1}{3} \begin{bmatrix} 0 & 3 \\ 3 & 0 \\ 3 & 3 \end{bmatrix} \begin{bmatrix} a-4 \\ b-5 \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ -1 & 0 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} a-4 \\ b-5 \end{bmatrix} = \begin{bmatrix} 5-b \\ 4-a \\ 9-a-b \end{bmatrix}$$

Therefore,

$$\boxed{x = 6-b, \quad y = 6-a, \quad z = 12-a-b.}$$

Is it possible to solve for any triple of the variables x, y, z, a, b for the given system? In fact, no. Let me explain by linear algebra. We can calculate: the augmented coefficient matrix for $G(x, y, z, a, b) = (24, 30, 42)$ Gaussian eliminates as follows:

$$rref \left[\begin{array}{ccccc|c} 1 & 1 & 1 & 2 & 2 & 24 \\ 1 & 0 & 2 & 2 & 3 & 30 \\ 3 & 2 & 1 & 3 & 4 & 42 \end{array}\right] = \left[\begin{array}{ccccc|c} 1 & 0 & 0 & 0 & 1 & 6 \\ 0 & 1 & 0 & 1 & 0 & 6 \\ 0 & 0 & 1 & 1 & 1 & 12 \end{array}\right].$$

First, note this is consistent with the answer we derived above. Second, examine the columns of $rref[G']$. You can ignore the 6-th column in the interest of this thought extending to nonlinear systems. The question of the suitability of a triple amounts to the invertibility of the submatrix of G' which corresponds to the triple. Examine:

$$\frac{\partial G}{\partial(y, z, a)} = \begin{bmatrix} 1 & 1 & 2 \\ 0 & 2 & 2 \\ 2 & 1 & 3 \end{bmatrix}, \qquad \frac{\partial G}{\partial(x, z, b)} = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 2 & 3 \\ 3 & 1 & 4 \end{bmatrix}$$

both of these are clearly singular since the third column is the sum of the first two columns. Alternatively, you can calculate the determinant of each of the matrices above is zero. In contrast,

$$\frac{\partial G}{\partial(z, a, b)} = \begin{bmatrix} 1 & 2 & 2 \\ 2 & 2 & 2 \\ 1 & 3 & 4 \end{bmatrix}$$

is non-singular. How to I know there is no linear dependence? Well, we could calculate the determinant is $1(8-6)-2(8-2)+2(6-2)=-2\neq 0$*. Or, we could examine the row reduction above. The column correspondance property*[7] *states that linear dependences amongst columns of a matrix are preserved under row reduction. This means we can easily deduce dependence (if there is any) from the reduced matrix. Observe that column* 4 *is clearly the sum of columns* 2 *and* 3*. Likewise, column* 5 *is the sum of columns* 1 *and* 3*. On the other hand, columns* $3,4,5$ *admit no linear dependence. In general, more calculation would be required to "see" the independence of the far right columns. One reorders the columns and performs a new reduction to ascertain dependence. No such calculation is needed here since the problem is not that complicated.*

I find calculating the determinant of sub-Jacobian matrices is the simplest way for most students to quickly understand. I'll showcase this method in a series of examples attached to a later section. I have made use of some matrix theory in this section. If you didn't learn it in linear (or haven't taken linear yet) it's worth learning. These are nice tools to keep for later problems in life.

Remark 3.2.8. *independent constraints*

Gaussian elimination on a system of linear equations may produce a row of zeros. For example, $x+y=0$ and $2x+2y=0$ gives rref$\begin{bmatrix} 1 & 1 & 0 \\ 2 & 2 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$. The reason for this is quite obvious: the equations consider are not indpendent. In fact the second equation is a scalar multiple of the first. Generally, if there is some linear-dependence in a set of equations then we can expect this will happen. Although, if the equations are inhomogenous the last column might not be trivial because the system could be inconsistent (for example $x+y=1$ and $2x+2y=5$).

Consider $G:\mathbb{R}^n\times\mathbb{R}^p\rightarrow\mathbb{R}^n$. As we linearize $G=k$ we arrive at a homogeneous system which can be written briefly as $G'\vec{r}=0$ (think about Equation 3.1 with the k's cancelled). We should **define** $G(\vec{r})=k$ is a **system of** n **independent equations at** $\vec{r}_o$ iff $G(\vec{r}_o)=k$ and rref$[G'(\vec{r}_o)]$ has zero row. In other terminology, we could say the system of (possibly nonlinear) equations $G(\vec{r})=k$ is built from n-independent equations near $\vec{r}_o$ iff the Jacobian matrix has **full-rank** at $\vec{r}_o$. If this full-rank condition is met then we can solve for n of the variables in terms of the remaining p variables. In general there will be many choices of how to do this, and some choices will be forbidden as we have seen in the examples already.

[7] I like to call it the CCP in my linear notes

3.3 implicit differentiation

Enough theory, let's calculate. In this section I apply previous theoretical constructions to specific problems. I also introduce standard notation for "constrained" partial differentiation which is also sometimes called "partial differentiation with a side condition". The typical problem is the following: given equations:

$$\begin{aligned} G_1(x_1, \dots, x_m, y_1, \dots, y_n) &= k_1 \\ G_2(x_1, \dots, x_m, y_1, \dots, y_n) &= k_2 \\ &\vdots \\ G_n(x_1, \dots, x_m, y_1, \dots, y_n) &= k_n \end{aligned}$$

calculate partial derivative of dependent variables with respect to independent variables. Continuing with the notation of the implicit function discussion we'll assume that y will be dependent on x. I want to recast some of our arguments via differentials[8]. Take the total differential of each equation above,

$$\begin{aligned} dG_1(x_1, \dots, x_m, y_1, \dots, y_n) &= 0 \\ dG_2(x_1, \dots, x_m, y_1, \dots, y_n) &= 0 \\ &\vdots \\ dG_n(x_1, \dots, x_m, y_1, \dots, y_n) &= 0 \end{aligned}$$

Hence,

$$\begin{aligned} \partial_{x_1} G_1 dx_1 + \cdots + \partial_{x_m} G_1 dx_m + \partial_{y_1} G_1 dy_1 + \cdots + \partial_{y_n} G_1 dy_n &= 0 \\ \partial_{x_1} G_2 dx_1 + \cdots + \partial_{x_m} G_2 dx_m + \partial_{y_1} G_2 dy_1 + \cdots + \partial_{y_n} G_2 dy_n &= 0 \\ &\vdots \\ \partial_{x_1} G_n dx_1 + \cdots + \partial_{x_m} G_n dx_m + \partial_{y_1} G_n dy_1 + \cdots + \partial_{y_n} G_n dy_n &= 0 \end{aligned}$$

Notice, this can be nicely written in column vector notation as:

$$\partial_{x_1} G dx_1 + \cdots + \partial_{x_m} G dx_m + \partial_{y_1} G dy_1 + \cdots + \partial_{y_n} G dy_n = 0$$

Or, in matrix notation:

$$[\partial_{x_1} G | \cdots | \partial_{x_m} G] \begin{bmatrix} dx_1 \\ \vdots \\ dx_m \end{bmatrix} + [\partial_{y_1} G | \cdots | \partial_{y_n} G] \begin{bmatrix} dy_1 \\ \vdots \\ dy_n \end{bmatrix} = 0$$

Finally, solve for dy, we assume $[\partial_{y_1} G | \cdots | \partial_{y_n} G]^{-1}$ exists,

$$\begin{bmatrix} dy_1 \\ \vdots \\ dy_n \end{bmatrix} = -[\partial_{y_1} G | \cdots | \partial_{y_n} G]^{-1} [\partial_{x_1} G | \cdots | \partial_{x_m} G] \begin{bmatrix} dx_1 \\ \vdots \\ dx_m \end{bmatrix}$$

Given all of this we can calculate $\frac{\partial y_i}{\partial x_j}$ by simply reading the coeffient dx_j in the i-th row. I will make this idea quite explicit in the examples that follow.

[8] in contrast, In the previous section we mostly used derivative notation

Example 3.3.1. *Let's return to a common calculus III problem. Suppose $F(x,y,z) = k$ for some constant k.* **Find partial derivatives of x, y or z with repsect to the remaining variables.** *Solution: I'll use the method of differentials once more:*

$$dF = F_x dx + F_y dy + F_z dz = 0$$

We can solve for dx, dy or dz provided F_x, F_y or F_z is nonzero respective and these differential expressions reveal various partial derivatives of interest:

$$dx = -\frac{F_y}{F_x}dy - \frac{F_z}{F_x}dz \qquad \Rightarrow \qquad \frac{\partial x}{\partial y} = -\frac{F_y}{F_x} \;\; \& \;\; \frac{\partial x}{\partial z} = -\frac{F_z}{F_x}$$

$$dy = -\frac{F_x}{F_y}dx - \frac{F_z}{F_y}dz \qquad \Rightarrow \qquad \frac{\partial y}{\partial x} = -\frac{F_x}{F_y} \;\; \& \;\; \frac{\partial y}{\partial z} = -\frac{F_z}{F_y}$$

$$dz = -\frac{F_x}{F_z}dx - \frac{F_y}{F_z}dy \qquad \Rightarrow \qquad \frac{\partial z}{\partial x} = -\frac{F_x}{F_z} \;\; \& \;\; \frac{\partial z}{\partial y} = -\frac{F_y}{F_z}$$

In each case above, the implicit function theorem allows us to solve for one variable in terms of the remaining two. If the partial derivative of F in the denominator are zero then the implicit function theorem does not apply and other thoughts are required. Often calculus text give the following as a homework problem:

$$\frac{\partial x}{\partial y}\frac{\partial y}{\partial z}\frac{\partial z}{\partial x} = -\frac{F_y}{F_x}\frac{F_z}{F_y}\frac{F_x}{F_z} = -1.$$

In the equation above we have x appear as a dependent variable on y, z and also as an independent variable for the dependent variable z. These mixed expressions are actually of interest to engineering and physics. The less mbiguous notation below helps better handle such expressions:

$$\left(\frac{\partial x}{\partial y}\right)_z \left(\frac{\partial y}{\partial z}\right)_x \left(\frac{\partial z}{\partial x}\right)_y = -1.$$

In each part of the expression we have clearly denoted which variables are taken to depend on the others and in turn what sort of partial derivative we mean to indicate. Partial derivatives are not taken alone, they must be done in concert with an understanding of the totality of the indpendent variables for the problem. We hold all the remaining indpendent variables fixed as we take a partial derivative.

The explicit independent variable notation is more important for problems where we can choose more than one set of indpendent variables for a given dependent variables. In the example that follows we study $w = w(x,y)$ but we could just as well consider $w = w(x,z)$. Generally it will not be the case that $\left(\frac{\partial w}{\partial x}\right)_y$ is the same as $\left(\frac{\partial w}{\partial x}\right)_z$. In calculation of $\left(\frac{\partial w}{\partial x}\right)_y$ we hold y constant as we vary x whereas in $\left(\frac{\partial w}{\partial x}\right)_z$ we hold z constant as we vary x. There is no reason these ought to be the same[9].

Example 3.3.2. *Suppose $x+y+z+w = 3$ and $x^2-2xyz+w^3 = 5$.* **Calculate partial derivatives of z and w with respect to the independent variables x, y.** *Solution: we begin by calculation of the differentials of both equations:*

$$\begin{aligned} &dx + dy + dz + dw = 0 \\ &(2x - 2yz)dx - 2xzdy - 2xydz + 3w^2dw = 0 \end{aligned}$$

[9] a good exercise would be to do the example over but instead aim to calculate partial derivatives for y, w with respect to independent variables x, z

We can solve for (dz, dw). In this calculation we can treat the differentials as formal variables.

$$\begin{aligned} dz + dw &= -dx - dy \\ -2xydz + 3w^2 dw &= -(2x - 2yz)dx + 2xzdy \end{aligned}$$

I find matrix notation is often helpful,

$$\begin{bmatrix} 1 & 1 \\ -2xy & 3w^2 \end{bmatrix} \begin{bmatrix} dz \\ dw \end{bmatrix} = \begin{bmatrix} -dx - dy \\ -(2x - 2yz)dx + 2xzdy \end{bmatrix}$$

Use Kramer's rule, multiplication by inverse, substitution, adding/subtracting equations etc... whatever technique of solving linear equations you prefer. Our goal is to solve for dz and dw in terms of dx and dy. I'll use Kramer's rule this time:

$$dz = \frac{det \left[\begin{array}{c|c} -dx - dy & 1 \\ -(2x - 2yz)dx + 2xzdy & 3w^2 \end{array} \right]}{det \begin{bmatrix} 1 & 1 \\ -2xy & 3w^2 \end{bmatrix}} = \frac{3w^2(-dx - dy) + (2x - 2yz)dx - 2xzdy}{3w^2 + 2xy}$$

Collecting terms,

$$dz = \left(\frac{-3w^2 + 2x - 2yz}{3w^2 + 2xy} \right) dx + \left(\frac{-3w^2 - 2xz}{3w^2 + 2xy} \right) dy$$

From the expression above we can read various implicit derivatives,

$$\boxed{ \left(\frac{\partial z}{\partial x} \right)_y = \frac{-3w^2 + 2x - 2yz}{3w^2 + 2xy} \qquad \& \qquad \left(\frac{\partial z}{\partial y} \right)_x = \frac{-3w^2 - 2xz}{3w^2 + 2xy} }$$

The notation above indicates that z is understood to be a function of independent variables x, y. $\left(\frac{\partial z}{\partial x}\right)_y$ means we take the derivative of z with respect to x while holding y fixed. The appearance of the dependent variable w can be removed by using the equations $G(x, y, z, w) = (3, 5)$. Similar ambiguities exist for implicit differentiation in calculus I. Apply Kramer's rule once more to solve for dw:

$$dw = \frac{det \left[\begin{array}{c|c} 1 & -dx - dy \\ -2xy & -(2x - 2yz)dx + 2xzdy \end{array} \right]}{det \begin{bmatrix} 1 & 1 \\ -2xy & 3w^2 \end{bmatrix}} = \frac{-(2x - 2yz)dx + 2xzdy - 2xy(dx + dy)}{3w^2 + 2xy}$$

Collecting terms,

$$dw = \left(\frac{-2x + 2yz - 2xy}{3w^2 + 2xy} \right) dx + \left(\frac{2xzdy - 2xydy}{3w^2 + 2xy} \right) dy$$

We can read the following from the differential above:

$$\boxed{ \left(\frac{\partial w}{\partial x} \right)_y = \frac{-2x + 2yz - 2xy}{3w^2 + 2xy} \qquad \& \qquad \left(\frac{\partial w}{\partial y} \right)_x = \frac{2xzdy - 2xydy}{3w^2 + 2xy} }$$

You should ask: where did we use the implicit function theorem in the preceding example? Notice our underlying hope is that we can solve for $z = z(x, y)$ and $w = w(x, y)$. The implicit function theorem states this is possible precisely when $\frac{\partial G}{\partial(z,w)} = \begin{bmatrix} 1 & 1 \\ -2xy & 3w^2 \end{bmatrix}$ is non singular. Interestingly this is the same matrix we must consider to isolate dz and dw. The calculations of the example are only meaningful if the $det \begin{bmatrix} 1 & 1 \\ -2xy & 3w^2 \end{bmatrix} \neq 0$. In such a case the implicit function theorem applies and it is reasonable to suppose z, w can be written as functions of x, y.

3.3.1 computational techniques for partial differentiation with side conditions

In this section I show you how I teach this to calculus III. In other words, we set-aside the explicit mention of the implicit function theorem and work out some typical calculations. If one desires rigor then the answer is found from the implicit function theorems careful application, that is how to justify what follows. These notes are taken from my calculus III notes, but I thought it wise to include them here since most calculus texts do not bother to show these calculations (which is sad since they actually matter to the application of multivariate analysis to many real world applications) To begin, we define[10] the total differential.

Definition 3.3.3.

If $f = f(x_1, x_2, \dots, x_n)$ then $df = \frac{\partial f}{\partial x_1}dx_1 + \frac{\partial f}{\partial x_2}dx_2 + \cdots + \frac{\partial f}{\partial x_n}dx_n$.

Example 3.3.4. *Suppose $E = pv + t^2$ then $dE = vdp + pdv + 2tdt$. In this example the dependent variable is E whereas the independent variables are p, v and t.*

Example 3.3.5. **Problem:** *what are $\partial F/\partial x$ and $\partial F/\partial y$ if we know that $F = F(x,y)$ and $dF = (x^2+y)dx - \cos(xy)dy$.*
Solution: *if $F = F(x,y)$ then the total differential has the form $dF = F_x dx + F_y dy$. We simply compare the general form to the given $dF = (x^2+y)dx - \cos(xy)dy$ to obtain:*

$$\frac{\partial F}{\partial x} = x^2 + y, \qquad \frac{\partial F}{\partial y} = -\cos(xy).$$

Example 3.3.6. *Suppose $w = xyz$ then $dw = yzdx + xzdy + xydz$. On the other hand, we can solve for $z = z(x,y,w)$*

$$z = \frac{w}{xy} \qquad \Rightarrow \qquad dz = -\frac{w}{x^2y}dx - \frac{w}{xy^2}dy + \frac{1}{xy}dw. \ \star$$

If we solve $dw = yzdx + xzdy + xydz$ directly for dz we obtain:

$$dz = -\frac{z}{x}dx - \frac{z}{y}dy + \frac{1}{xy}dw \quad \star\star.$$

Are $\star$ and $\star\star$ consistent? Well, yes. Note $\frac{w}{x^2y} = \frac{xyz}{x^2y} = \frac{z}{x}$ and $\frac{w}{xy^2} = \frac{xyz}{xy^2} = \frac{z}{y}$.

Which variables are independent/dependent in the example above? It depends. In this initial portion of the example we treated x, y, z as independent whereas w was dependent. But, in the last half we treated x, y, w as independent and z was the dependent variable. Consider this, if I ask you what the value of $\frac{\partial z}{\partial x}$ is in the example above then this question is ambiguous!

$$\underbrace{\frac{\partial z}{\partial x} = 0}_{z \ indpendent \ of \ x} \qquad \text{verses} \qquad \underbrace{\frac{\partial z}{\partial x} = \frac{-z}{x}}_{z \ depends \ on \ x}$$

Obviously this sort of ambiguity is rather unpleasant. A natural solution to this trouble is simply to write a bit more when variables are used in multiple contexts. In particular,

$$\underbrace{\frac{\partial z}{\partial x}\bigg|_{y,z} = 0}_{means \ x,y,z \ independent} \qquad \text{is different than} \qquad \underbrace{\frac{\partial z}{\partial x}\bigg|_{y,w} = \frac{-z}{x}}_{means \ x,y,w \ independent} \quad .$$

[10] I invite the reader to verify the notation ”defined” in this section is in fact totally sympatico with our previous definitions

The key concept is that all the other independent variables are held fixed as an indpendent variable is partial differentiated. Holding y, z fixed as x varies means z does not change hence $\frac{\partial z}{\partial x}\big|_{y,z} = 0$. On the other hand, if we hold y, w fixed as x varies then the change in z need not be trivial; $\frac{\partial z}{\partial x}\big|_{y,w} = \frac{-z}{x}$. Let me expand on how this notation interfaces with the total differential.

Definition 3.3.7.

If w, x, y, z are variables then

$$dw = \frac{\partial w}{\partial x}\bigg|_{y,z} dx + \frac{\partial w}{\partial y}\bigg|_{x,z} dy + \frac{\partial w}{\partial z}\bigg|_{x,y} dz.$$

Alternatively,

$$dx = \frac{\partial x}{\partial w}\bigg|_{y,z} dw + \frac{\partial x}{\partial y}\bigg|_{w,z} dy + \frac{\partial x}{\partial z}\bigg|_{w,y} dz.$$

The larger idea here is that we can identify partial derivatives from the coefficients in equations of differentials. I'd say a differential equation but you might get the wrong idea... Incidentally, there is a whole theory of solving differential equations by clever use of differentials, I have books if you are interested.

Example 3.3.8. *Suppose $w = x + y + z$ and $x + y = wz$ then calculate $\frac{\partial w}{\partial x}\big|_y$ and $\frac{\partial w}{\partial x}\big|_z$. Notice we must choose dependent and independent variables to make sense of partial derivatives in question.*

1. *suppose w, z both depend on x, y. Calculate,*

$$\frac{\partial w}{\partial x}\bigg|_y = \frac{\partial}{\partial x}\bigg|_y (x+y+z) = \frac{\partial x}{\partial x}\bigg|_y + \frac{\partial y}{\partial x}\bigg|_y + \frac{\partial z}{\partial x}\bigg|_y = 1 + 0 + \frac{\partial z}{\partial x}\bigg|_y \quad \star$$

To calculate further we need to eliminate w by substituting $w = x + y + z$ into $x + y = wz$; thus $x + y = (x + y + z)z$ hence $dx + dy = (dx + dy + dz)z + (x + y + z)dz$

$$(2z + x + y)dz = (1 - z)dx + (1 - z)dy \quad \star\star$$

Therefore,

$$dz = \frac{1-z}{2z+x+y}dx + \frac{1-z}{2z+x+y}dy = \frac{\partial z}{\partial x}\bigg|_y dx + \frac{\partial z}{\partial y}\bigg|_x dy \quad \Rightarrow \quad \frac{\partial z}{\partial x}\bigg|_y = \frac{1-z}{2z+x+y}.$$

Returning to $\star$ we derive

$$\boxed{\frac{\partial w}{\partial x}\bigg|_y = 1 + \frac{1-z}{2z+x+y}.}$$

2. *suppose w, y both depend on x, z. Calculate,*

$$\frac{\partial w}{\partial x}\bigg|_z = \frac{\partial}{\partial x}\bigg|_z (x+y+z) = \frac{\partial x}{\partial x}\bigg|_z + \frac{\partial y}{\partial x}\bigg|_z + \frac{\partial z}{\partial x}\bigg|_z = 1 + \frac{\partial y}{\partial x}\bigg|_z + 0$$

To complete this calculation we need to eliminate w as before, using $\star\star$,

$$(1 - z)dy = (1 - z)dx - (2z + x + y)dz \quad \Rightarrow \quad \frac{\partial y}{\partial x}\bigg|_z = 1.$$

Therefore,

$$\boxed{\frac{\partial w}{\partial x}\bigg|_z = 2.}$$

I hope you can begin to see how the game is played. Basically the example above generalizes the idea of implicit differentiation to several equations of many variables. This is actually a pretty important type of calculation for engineering. The study of thermodynamics is full of variables which are intermittently used as either dependent or independent variables. The so-called equation of state can be given in terms of about a dozen distinct sets of state variables.

Example 3.3.9. *The ideal gas law states that for a fixed number of particles n the pressure P, volume V and temperature T are related by $PV = nRT$ where R is a constant. Calculate,*

$$\frac{\partial P}{\partial V}\bigg|_T = \frac{\partial}{\partial V}\left[\frac{nRT}{V}\right]\bigg|_T = -\frac{nRT}{V^2},$$

$$\frac{\partial V}{\partial T}\bigg|_P = \frac{\partial}{\partial T}\left[\frac{nRT}{P}\right]\bigg|_T = \frac{nR}{P},$$

$$\frac{\partial T}{\partial P}\bigg|_V = \frac{\partial}{\partial P}\left[\frac{PV}{nR}\right]\bigg|_T = \frac{V}{nR}.$$

You might expect that $\frac{\partial P}{\partial V}\big|_T \frac{\partial V}{\partial T}\big|_P \frac{\partial T}{\partial P}\big|_V = 1$. Is it true?

$$\frac{\partial P}{\partial V}\bigg|_T \frac{\partial V}{\partial T}\bigg|_P \frac{\partial T}{\partial P}\bigg|_V = -\frac{nRT}{V^2}\cdot\frac{nR}{P}\cdot\frac{V}{nR} = \frac{-nRT}{PV} = -1.$$

This is an example where naive cancellation of partials fails.

Example 3.3.10. *Suppose $F(x,y) = 0$ then $dF = F_x dx + F_y dy = 0$ and it follows that $dx = -\frac{F_y}{F_x}dy$ or $dy = -\frac{F_x}{F_y}dx$. Hence, $\frac{\partial x}{\partial y} = -\frac{F_y}{F_x}$ and $\frac{\partial y}{\partial x} = -\frac{F_x}{F_y}$. Therefore,*

$$\frac{\partial x}{\partial y}\frac{\partial y}{\partial x} = \frac{F_y}{F_x}\cdot\frac{F_x}{F_y} = 1$$

for (x,y) such that $F_x \neq 0$ and $F_y \neq 0$. The condition $F_x \neq 0$ suggests we can solve for $y = y(x)$ whereas the condition $F_y \neq 0$ suggests we can solve for $x = x(y)$.

3.4 the constant rank theorem

The implicit function theorem required we work with independent constraints. However, one does not always have that luxury. There is a theorem which deals with the slightly more general case. The base idea is that if the Jacobian has rank k then it locally injects a k-dimensional image into the codomain. If we are using a map as a parametrization then the rank k condition suggests the mapping does parametrize a k-fold, at least locally. On the other hand, if we are using the map to define a space as a level set then $F : \mathbb{R}^n \to \mathbb{R}^p$ has $F^{-1}(C)$ as a $(n-k)$-fold. Previously, we would have insisted $k = p$. I leave proof of this claim to the student. Perhaps you will find the red comments give further insight as to the meaning of the Jacobian.

Remark 3.4.1.

I have put remarks about the rank of the derivative in red for the examples below.

Example 3.4.2. *Let $f(t) = (t, t^2, t^3)$ then $f'(t) = (1, 2t, 3t^2)$. In this case we have*

$$f'(t) = [df_t] = \begin{bmatrix} 1 \\ 2t \\ 3t^2 \end{bmatrix}$$

The Jacobian here is a single column vector. It has rank 1 provided the vector is nonzero. We see that $f'(t) \neq (0, 0, 0)$ for all $t \in \mathbb{R}$. This corresponds to the fact that this space curve has a well-defined tangent line for each point on the path.

Example 3.4.3. *Let $f(\vec{x}, \vec{y}) = \vec{x} \cdot \vec{y}$ be a mapping from $\mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$. I'll denote the coordinates in the domain by $(x_1, x_2, x_3, y_1, y_2, y_3)$ thus $f(\vec{x}, \vec{y}) = x_1y_1 + x_2y_2 + x_3y_3$. Calculate,*

$$[df_{(\vec{x},\vec{y})}] = \nabla f(\vec{x}, \vec{y})^T = [y_1, y_2, y_3, x_1, x_2, x_3]$$

The Jacobian here is a single row vector. It has rank 6 provided all entries of the input vectors are nonzero.

Example 3.4.4. *Let $f(\vec{x}, \vec{y}) = \vec{x} \cdot \vec{y}$ be a mapping from $\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$. I'll denote the coordinates in the domain by $(x_1, \dots, x_n, y_1, \dots, y_n)$ thus $f(\vec{x}, \vec{y}) = \sum_{i=1}^n x_iy_i$. Calculate,*

$$\frac{\partial}{x_j}\left[\sum_{i=1}^n x_iy_i\right] = \sum_{i=1}^n \frac{\partial x_i}{x_j} y_i = \sum_{i=1}^n \delta_{ij}y_i = y_j$$

Likewise,

$$\frac{\partial}{y_j}\left[\sum_{i=1}^n x_iy_i\right] = \sum_{i=1}^n x_i\frac{\partial y_i}{y_j} = \sum_{i=1}^n x_i\delta_{ij} = x_j$$

Therefore, noting that $\nabla f = (\partial_{x_1} f, \dots, \partial_{x_n} f, \partial_{y_1} f, \dots, \partial_{y_n} f)$,

$$[df_{(\vec{x},\vec{y})}]^T = (\nabla f)(\vec{x}, \vec{y}) = \vec{y} \times \vec{x} = (y_1, \dots, y_n, x_1, \dots, x_n)$$

The Jacobian here is a single row vector. It has rank 2n provided all entries of the input vectors are nonzero.

Example 3.4.5. *Suppose $F(x, y, z) = (xyz, y, z)$ we calculate,*

$$\frac{\partial F}{\partial x} = (yz, 0, 0) \qquad \frac{\partial F}{\partial y} = (xz, 1, 0) \qquad \frac{\partial F}{\partial z} = (xy, 0, 1)$$

Remember these are actually column vectors in my sneaky notation; $(v_1, \dots, v_n) = [v_1, \dots, v_n]^T$. This means the **derivative** *or* **Jacobian matrix** *of F at (x, y, z) is*

$$F'(x, y, z) = [dF_{(x,y,z)}] = \begin{bmatrix} yz & xz & xy \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Note, $\text{rank}(F'(x, y, z)) = 3$ for all $(x, y, z) \in \mathbb{R}^3$ such that $y, z \neq 0$. There are a variety of ways to see that claim, one way is to observe $\det[F'(x, y, z)] = yz$ and this determinant is nonzero so long as neither y nor z is zero. In linear algebra we learn that a square matrix is invertible iff it has nonzero determinant iff it has linearly indpendent column vectors.

Example 3.4.6. *Suppose* $F(x,y,z) = (x^2+z^2, yz)$ *we calculate,*

$$\tfrac{\partial F}{\partial x} = (2x, 0) \qquad \tfrac{\partial F}{\partial y} = (0, z) \qquad \tfrac{\partial F}{\partial z} = (2z, y)$$

The derivative is a 2×3 *matrix in this example,*

$$F'(x,y,z) = [dF_{(x,y,z)}] = \left[\begin{array}{ccc} 2x & 0 & 2z \\ 0 & z & y \end{array}\right]$$

The maximum rank for F' *is* 2 *at a particular point* (x,y,z) *because there are at most two linearly independent vectors in* $\mathbb{R}^2$. *You can consider the three square submatrices to analyze the rank for a given point. If any one of these is nonzero then the rank (dimension of the column space) is two.*

$$M_1 = \left[\begin{array}{cc} 2x & 0 \\ 0 & z \end{array}\right] \qquad M_2 = \left[\begin{array}{cc} 2x & 2z \\ 0 & y \end{array}\right] \qquad M_3 = \left[\begin{array}{cc} 0 & 2z \\ z & y \end{array}\right]$$

We'll need either $det(M_1) = 2xz \neq 0$ *or* $det(M_2) = 2xy \neq 0$ *or* $det(M_3) = -2z^2 \neq 0$. *I believe the only point where all three of these fail to be true simulataneously is when* $x = y = z = 0$. *This mapping has maximal rank at all points except the origin.*

Example 3.4.7. *Suppose* $F(x,y) = (x^2+y^2, xy, x+y)$ *we calculate,*

$$\tfrac{\partial F}{\partial x} = (2x, y, 1) \qquad \tfrac{\partial F}{\partial y} = (2y, x, 1)$$

The derivative is a 3×2 *matrix in this example,*

$$F'(x,y) = [dF_{(x,y)}] = \left[\begin{array}{cc} 2x & 2y \\ y & x \\ 1 & 1 \end{array}\right]$$

The maximum rank is again 2, this time because we only have two columns. The rank will be two if the columns are not linearly dependent. We can analyze the question of rank a number of ways but I find determinants of submatrices a comforting tool in these sort of questions. If the columns are linearly dependent then all three sub-square-matrices of F' *will be zero. Conversely, if even one of them is nonvanishing then it follows the columns must be linearly independent. The submatrices for this problem are:*

$$M_1 = \left[\begin{array}{cc} 2x & 2y \\ y & x \end{array}\right] \qquad M_2 = \left[\begin{array}{cc} 2x & 2y \\ 1 & 1 \end{array}\right] \qquad M_3 = \left[\begin{array}{cc} y & x \\ 1 & 1 \end{array}\right]$$

You can see $det(M_1) = 2(x^2 - y^2)$, $det(M_2) = 2(x - y)$ *and* $det(M_3) = y - x$. *Apparently we have* $rank(F'(x,y,z)) = 2$ *for all* $(x,y) \in \mathbb{R}^2$ *with* $y \neq x$. *In retrospect this is not surprising.*

Example 3.4.8. *Let* $F(x,y) = (x, y, \sqrt{R^2 - x^2 - y^2})$ *for a constant* R. *We calculate,*

$$\nabla\sqrt{R^2 - x^2 - y^2} = \left(\tfrac{-x}{\sqrt{R^2-x^2-y^2}}, \tfrac{-y}{\sqrt{R^2-x^2-y^2}} \right)$$

Also, $\nabla x = (1,0)$ *and* $\nabla y = (0,1)$ *thus*

$$F'(x,y) = \left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \\ \frac{-x}{\sqrt{R^2-x^2-y^2}} & \frac{-y}{\sqrt{R^2-x^2-y^2}} \end{array}\right]$$

This matrix clearly has rank 2 where is is well-defined. Note that we need $R^2 - x^2 - y^2 > 0$ for the derivative to exist. Moreover, we could define $G(y, z) = (\sqrt{R^2 - y^2 - z^2}, y, z)$ and calculate,

$$G'(y, z) = \begin{bmatrix} 1 & 0 \\ \frac{-y}{\sqrt{R^2-y^2-z^2}} & \frac{-z}{\sqrt{R^2-y^2-z^2}} \\ 0 & 1 \end{bmatrix}.$$

Observe that $G'(y, z)$ exists when $R^2 - y^2 - z^2 > 0$. Geometrically, F parametrizes the sphere above the equator at $z = 0$ whereas G parametrizes the right-half of the sphere with $x > 0$. These parametrizations overlap in the first octant where both x and z are positive. In particular, $dom(F') \cap dom(G') = \{(x, y) \in \mathbb{R}^2 \mid x, y > 0 \text{ and } x^2 + y^2 < R^2\}$

Example 3.4.9. *Let $F(x, y, z) = (x, y, z, \sqrt{R^2 - x^2 - y^2 - z^2})$ for a constant R. We calculate,*

$$\nabla\sqrt{R^2 - x^2 - y^2 - z^2} = \left(\frac{-x}{\sqrt{R^2-x^2-y^2-z^2}}, \frac{-y}{\sqrt{R^2-x^2-y^2-z^2}}, \frac{-z}{\sqrt{R^2-x^2-y^2-z^2}} \right)$$

Also, $\nabla x = (1, 0, 0)$, $\nabla y = (0, 1, 0)$ and $\nabla z = (0, 0, 1)$ thus

$$F'(x, y, z) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ \frac{-x}{\sqrt{R^2-x^2-y^2-z^2}} & \frac{-y}{\sqrt{R^2-x^2-y^2-z^2}} & \frac{-z}{\sqrt{R^2-x^2-y^2-z^2}} \end{bmatrix}$$

This matrix clearly has rank 3 where is is well-defined. Note that we need $R^2-x^2-y^2-z^2 > 0$ for the derivative to exist. This mapping gives us a parametrization of the 3-sphere $x^2 + y^2 + z^2 + w^2 = R^2$ for $w > 0$. (drawing this is a little trickier)

Example 3.4.10. *Let $f(x, y, z) = (x + y, y + z, x + z, xyz)$. You can calculate,*

$$[df_{(x,y,z)}] = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \\ yz & xz & xy \end{bmatrix}$$

This matrix clearly has rank 3 and is well-defined for all of $\mathbb{R}^3$.

Example 3.4.11. *Let $f(x, y, z) = xyz$. You can calculate,*

$$[df_{(x,y,z)}] = \begin{bmatrix} yz & xz & xy \end{bmatrix}$$

This matrix fails to have rank 3 if x, y or z are zero. In other words, $f'(x, y, z)$ has rank 3 in $\mathbb{R}^3$ provided we are at a point which is not on some coordinate plane. (the coordinate planes are $x = 0, y = 0$ and $z = 0$ for the yz, zx and xy coordinate planes respective)

Example 3.4.12. *Let $f(x, y, z) = (xyz, 1 - x - y)$. You can calculate,*

$$[df_{(x,y,z)}] = \begin{bmatrix} yz & xz & xy \\ -1 & -1 & 0 \end{bmatrix}$$

This matrix has rank 3 if either $xy \neq 0$ or $(x - y)z \neq 0$. In contrast to the preceding example, the derivative does have rank 3 on certain points of the coordinate planes. For example, $f'(1, 1, 0)$ and $f'(0, 1, 1)$ both give $rank(f') = 3$.

Example 3.4.13. *Let $X(u,v) = (x,y,z)$ where x,y,z denote functions of u,v and I prefer to omit the explicit depedendence to reduce clutter in the equations to follow.*

$$\frac{\partial X}{\partial u} = X_u = (x_u, y_u, z_u) \quad \text{and} \quad \frac{\partial X}{\partial v} = X_v = (x_v, y_v, z_v)$$

Then the Jacobian is the 3×2 matrix

$$\left[dX_{(u,v)}\right] = \begin{bmatrix} x_u & x_v \\ y_u & y_v \\ z_u & z_v \end{bmatrix}$$

The matrix $\left[dX_{(u,v)}\right]$ has rank 2 if at least one of the determinants below is nonzero,

$$det \begin{bmatrix} x_u & x_v \\ y_u & y_v \end{bmatrix} \quad det \begin{bmatrix} x_u & x_v \\ z_u & z_v \end{bmatrix} \quad det \begin{bmatrix} y_u & y_v \\ z_u & z_v \end{bmatrix}$$

Chapter 4

two views of manifolds in $\mathbb{R}^n$

In this chapter we describe spaces inside $\mathbb{R}^n$ which are k-dimensional [1]. Technically, to make this precise we would need to study manifolds with boundary. Careful discussion of manifolds with boundary in euclidean space can be found in Munkres *Analysis on Manifolds.* In the interest of focusing on examples, I'll be a bit fuzzy about the defintion of a k-dimensional subspace S of euclidean space. This much we can say: there are two ways to envision the geometry of S:

(1.) **Parametrically**: provide a **patch** R such that $R : U \subseteq \mathbb{R}^k \to S \subseteq \mathbb{R}^n$. Here U is called the **parameter space** and R^{-1} is called a **coordinate chart**. The cannonical example:

$$R(x_1, \dots x_k) = (x_1, \dots x_k, 0, \dots, 0).$$

(2.) **Implicitly**: provide a **level function** $G : \mathbb{R}^k \times \mathbb{R}^p \to \mathbb{R}^p$ such that $S = G^{-1}\{c\} = S$. This viewpoint casts S as points in $x \in \mathbb{R}^k \times \mathbb{R}^p$ for which $G(x) = k$. The cannonical example:

$$G(x_1, \dots, x_{k+p}) = (x_{k+1}, \dots, x_{k+p}) = (0, \dots, 0).$$

The cannonical examples of (1.) and (2.) are both the $x_1 \dots x_k$-coordinate plane embedded in $\mathbb{R}^n$. Just to take it down a notch. If $n = 3$ then we could look at the xy-plane in either view as follows:

$$(1.)\ R(x, y) = (x, y, 0) \qquad (2.)\ G(x, y, z) = z = 0.$$

Which viewpoint should we adopt? What is the dimension of a given space S? How should we find tangent space to S? How should we find the normal space to S? These are the questions we set-out to answer in this chapter.

Orthogonal complements help us to understand how all of this fits together. This is possible since we deal with embedded manifolds for which the euclidean dot-product of $\mathbb{R}^n$ is available to sort out the geometry. Finally, we use this geometry and a few simple lemmas to justify the method of Lagrange multipliers. Lagrange's technique paired with the theory of multivariate Taylor polynomials form the basis for analyzing extrema for multivariate functions. In this chapter we deal with the question of extrema on the edges of a set. The second half of the story is found in the next chapter where we deal with the interior points via the theory of quadratic forms applied to the second-order approximation to a function of several variables.

[1] I'll try to stick with this notation for this chapter, $n \geq k$ and $n = p + k$

4.1 definition of level set

A level set is the solution set of some equation or system of equations. We confine our interest to level sets of $\mathbb{R}^n$. For example, the set of all (x, y) that satisfy

$$G(x, y) = c$$

is called a **level curve** in $\mathbb{R}^2$. Often we can use k to label the curve. You should also recall **level surfaces** in $\mathbb{R}^3$ are defined by an equation of the form

$$G(x, y, z) = c.$$

The set of all $(x_1, x_2, x_3, x_4) \in \mathbb{R}^4$ which solve $G(x_1, x_2, x_3, x_4) = c$ is a **level volume** in $\mathbb{R}^4$. We can obtain lower dimensional objects by simultaneously imposing several equations at once. For example, suppose $G_1(x, y, z) = z = 1$ and $G_2(x, y, z) = x^2+y^2+z^2 = 5$, points (x, y, z) which solve both of these equations are on the intersection of the plane $z = 1$ and the sphere $x^2 + y^2 + z^2 = 5$. Let $G = (G_1, G_2)$, note that $G(x, y, z) = (1, 5)$ describes a circle in $\mathbb{R}^3$. More generally:

Definition 4.1.1.

> Suppose $G : dom(G) \subseteq \mathbb{R}^k \times \mathbb{R}^p \to \mathbb{R}^p$. Let c be a vector of constants in $\mathbb{R}^p$ and suppose $S = \{x \in \mathbb{R}^k \times \mathbb{R}^p \mid G(x) = c\}$ is non-empty and G is continuously differentiable on an open set containing S. We say S is an k-**dimensional level** set iff $G'(x)$ has p linearly independent rows at each $x \in S$.

The condition of linear independence of the rows is give to eliminate possible redundancy in the system of equations. In the case that $p = 1$ the criteria reduces to $G'(x) \neq 0$ over the level set of dimension $n - 1$. Intuitively we think of each equation in $G(x) = c$ as removing one of the dimensions of the ambient space $\mathbb{R}^n = \mathbb{R}^k \times \mathbb{R}^p$. It is worthwhile to cite a useful result from linear algebra at this point:

Proposition 4.1.2.

> Let $A \in \mathbb{R}^{m \times n}$. The number of linearly independent columns in A is the same as the number of linearly independent rows in A. This invariant of A is called the **rank** of A.

Given the wisdom of linear algebra we see that we should require a k-dimensional level set $S = G^{-1}(c)$ to have a level function $G : \mathbb{R}^n \to \mathbb{R}^p$ whose derivative is of rank $n - k = p$ over all of S. We can either analyze linear independence of columns or rows.

Example 4.1.3. *Consider $G(x, y, z) = x^2 + y^2 - z^2$ and suppose $S = G^{-1}\{0\}$. Calculate,*

$$G'(x, y, z) = [2x, 2y, -2z]$$

Notice that $(0, 0, 0) \in S$ and $G'(0, 0, 0) = [0, 0, 0]$ hence G' is not rank one at the origin. At all other points in S we have $G'(x, y, z) \neq 0$ which means this is almost a $3 - 1 = 2$-dimensional level set. However, almost is not good enough in math. Under our definition the cone S is not a 2-dimensional level set since it fails to meet the full-rank criteria at the point of the cone.

A p-dimensional level set is an example of a p-dimensional manifold. The example above with the origin included is a manifold paired with a singular point, such spaces are known as **orbifolds**. The study of orbifolds has attracted considerable effort in recent years as the singularities of such orbifolds can be used to do physics in string theory. I digress. Let us examine another level set:

Example 4.1.4. *Let $G(x, y, z) = (x, y)$ and define $S = G^{-1}(a, b)$ for some fixed pair of constants $a, b \in \mathbb{R}$. We calculate that $G'(x, y, z) = I_2 \in \mathbb{R}^{2\times 2}$. We clearly have rank two at all points in S hence S is a $3 - 2 = 1$-dimensional level set. Perhaps you realize S is the vertical line which passes through $(a, b, 0)$ in the xy-plane.*

4.2 tangents and normals to a level set

There are many ways to define a tangent space for some subset of $\mathbb{R}^n$. One natural definition is that the tangent space to $p \in S$ is simply the set of all tangent vectors to curves on S which pass through the point p. In this section we study the geometry of curves on a level-set. We'll see how the tangent space is naturally a vector space in the particular context of level-sets in $\mathbb{R}^n$.

Throughout this section we assume that S is a k-dimensional level set defined by $G : \mathbb{R}^k \times \mathbb{R}^p \to \mathbb{R}^p$ where $G^{-1}(c) = S$. This means that we can apply the implicit function theorem to S and for any given point $p = (p_x, p_y) \in S$ where $p_x \in \mathbb{R}^k$ and $p_y \in \mathbb{R}^p$. There exists a local continuously differentiable solution $h : U \subseteq \mathbb{R}^k \to \mathbb{R}^p$ such that $h(p_x) = p_y$ and for all $x \in U$ we have $G(x, h(x)) = c$. We can view $G(x, y) = c$ for x near p as the graph of $y = h(x)$ for $x \in U$. With the set-up above in mind, suppose that $\gamma : \mathbb{R} \to U \subseteq S$. If we write $\gamma = (\gamma_x, \gamma_y)$ then it follows $\gamma = (\gamma_x, h \circ \gamma_x)$ over the subset $U \times h(U)$ of S. More explicitly, for all $t \in \mathbb{R}$ such that $\gamma(t) \in U \times h(U)$ we have

$$\gamma(t) = (\gamma_x(t), h(\gamma_x(t))).$$

Therefore, if $\gamma(0) = p$ then $\gamma(0) = (p_x, h(p_x))$. Differentiate, use the chain-rule in the second factor to obtain:

$$\gamma'(t) = (\gamma_x'(t), h'(\gamma_x(t))\gamma_x'(t)).$$

We find that the tangent vector to $p \in S$ of γ has a rather special form which was forced on us by the implicit function theorem:

$$\gamma'(0) = (\gamma_x'(0), h'(p_x)\gamma_x'(0)).$$

Or to cut through the notation a bit, if $\gamma'(0) = v = (v_x, v_y)$ then $v = (v_x, h'(p_x)v_x)$. The second component of the vector is not free of the first, it essentially redundant. This makes us suspect that the tangent space to S at p is k-dimensional.

Theorem 4.2.1.

Let $G : \mathbb{R}^k \times \mathbb{R}^p \to \mathbb{R}^p$ be a level-mappping which defines a k-dimensional level set S by $G^{-1}(c) = S$. Suppose $\gamma_1, \gamma_2 : \mathbb{R} \to S$ are differentiable curves with $\gamma_1'(0) = v_1$ and $\gamma_2'(0) = v_2$ then there exists a differentiable curve $\gamma : \mathbb{R} \to S$ such that $\gamma'(0) = v_1 + v_2$ and $\gamma(0) = p$. Moreover, there exists a differentiable curve $\beta : \mathbb{R} \to S$ such that $\beta'(0) = cv_1$ and $\beta(0) = p$.

Proof: It is convenient to define a map which gives a **local parametrization** of S at p. Since we have a description of S locally as a graph $y = h(x)$ (near p) it is simple to construct the parameterization. Define $\Phi : U \subseteq \mathbb{R}^k \to S$ by $\Phi(x) = (x, h(x))$. Clearly $\Phi(U) = U \times h(U)$ and there is an inverse mapping $\Phi^{-1}(x, y) = x$ is well-defined since $y = h(x)$ for each $(x, y) \in U \times h(U)$. Let $w \in \mathbb{R}^k$ and observe that

$$\psi(t) = \Phi(\Phi^{-1}(p) + tw) = \Phi(p_x + tw) = (p_x + tw, h(p_x + tw))$$

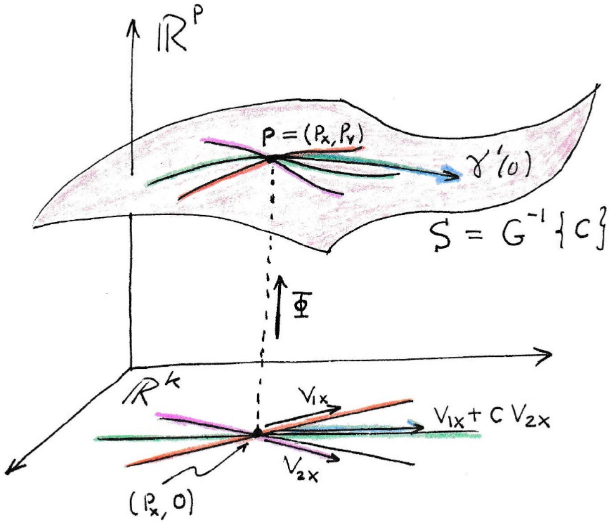
is a curve from $\mathbb{R}$ to $U \subseteq S$ such that $\psi(0) = (p_x, h(p_x)) = (p_x, p_y) = p$ and using the chain rule on the final form of $\psi(t)$:

$$\psi'(0) = (w, h'(p_x)w).$$

The construction above shows that any vector of the form $(v_x, h'(p_x)v_x)$ is the tangent vector of a particular differentiable curve in the level set (differentiability of ψ follows from the differentiability of h and the other maps which we used to construct ψ). In particular we can apply this to the case $w = v_{1x} + v_{2x}$ and we find $\gamma(t) = \Phi(\Phi^{-1}(p) + t(v_{1x} + v_{2x}))$ has $\gamma'(0) = v_1 + v_2$ and $\gamma(0) = p$.

Likewise, apply the construction to the case $w = cv_{1x}$ to write $\beta(t) = \Phi(\Phi^{-1}(p) + t(cv_{1x}))$ with $\beta'(0) = cv_1$ and $\beta(0) = p$. $\square$

The idea of the proof is encapsulated in the picture below. This idea of mapping lines in a flat domain to obtain standard curves in a curved domain is an idea which plays over and over as you study manifold theory. The particular redundancy of the x and y sub-vectors is special to the discussion level-sets, however anytime we have a local parametrization we'll be able to construct curves with tangents of our choosing by essentially the same construction. In fact, there are infinitely many curves which produce a particular tangent vector in the tangent space of a manifold.



Theorem 4.2.1 shows that the definition given below is logical. In particular, it is not at all obvious that the sum of two tangent vectors ought to again be a tangent vector. However, that is just what the Theorem 4.2.1 told us for level-sets[2].

Definition 4.2.2.

Suppose S is a k-dimensional level-set defined by $S = G^{-1}\{c\}$ for $G : \mathbb{R}^k \times \mathbb{R}^p \to \mathbb{R}^p$. We define the **tangent space at** $p \in S$ to be the set of pairs:

$$T_pS = \{(p, v) \mid \text{there exists differentiable } \gamma : \mathbb{R} \to S \text{ and } \gamma(0) = p \text{ where } v = \gamma'(0)\}$$

Moreover, we define (i.) **addition** and (ii.) **scalar multiplication** of vectors by the rules

$$(i.) \quad (p, v_1) + (p, v_2) = (p, v_1 + v_2) \qquad (ii.) \quad c(p, v_1) = (p, cv_1)$$

for all $(p, v_1), (p, v_2) \in T_pS$ and $c \in \mathbb{R}$.

When I picture T_pS in my mind I think of vectors pointing out from the base-point p. To make an explicit connection between the pairs of the above definition and the classical geometric form of the tangent space we simply take the image of T_pS under the mapping $\Psi(x, y) = x + y$ thus $\Psi(T_pS) = \{p + v \mid (p, v) \in T_pS\}$. I often picture T_pS as $\psi(T_pS)$[3]

[2] technically, there is another logical gap which I currently ignore. I wonder if you can find it.

[3] In truth, as you continue to study manifold theory you'll find at least three seemingly distinct objects which are all called "tangent vectors"; equivalence classes of curves, derivations, contravariant tensors.

We could set out to calculate tangent spaces in view of the definition above, but we are actually interested in more than just the tangent space for a level-set. In particular. we want a concrete description of all the vectors which are not in the tangent space.

Definition 4.2.3.

Suppose S is a k-dimensional level-set defined by $S = G^{-1}\{c\}$ for $G : \mathbb{R}^k \times \mathbb{R}^p \to \mathbb{R}^p$ and T_pS is the tangent space at p. Note that $T_pS \leq V_p$ where $V_p = \{p\} \times \mathbb{R}^k \times \mathbb{R}^p$ is given the natural vector space structure which we already exhibited on the subspace T_pS. We define the **inner product** on V_p as follows: for all $(p, v), (p, w) \in V_p$,

$$(p, v) \cdot (p, w) = v \cdot w.$$

The length of a vector (p, v) is naturally defined by $||(p, v)|| = ||v||$. Moreover, we say two vectors $(p, v), (p, w) \in V_p$ are **orthogonal** iff $v \cdot w = 0$. Given a set of vectors $R \subseteq V_p$ we define the **orthogonal complement** by

$$R^{\perp} = \{(p, v) \in V_p \mid (p, v) \cdot (p, r) \ \text{ for all } (p, r) \in R\}.$$

Suppose $W_1, W_2 \subseteq V_p$ then we say W_1 is **orthogonal** to W_2 iff $w_1 \cdot w_2 = 0$ for all $w_1 \in W_1$ and $w_2 \in W_2$. We denote orthogonality by writing $W_1 \perp W_2$. If every $v \in V_p$ can be written as $v = w_1 + w_2$ for a pair of $w_1 \in W_1$ and $w_2 \in W_2$ where $W_1 \perp W_2$ then we say that V_p is the **direct sum** of W_1 and W_2 which is denoted by $V_p = W_1 \oplus W_2$.

There is much more to say about orthogonality, however, our focus is not in that vein. We just need the langauge to properly define the normal space. The calculation below is probably the most important calculation to understand for a level-set. Suppose we have a curve $\gamma : \mathbb{R} \to S$ where $S = G^{-1}(c)$ is a k-dimensional level-set in $\mathbb{R}^k \times \mathbb{R}^p$. Observe that for all $t \in \mathbb{R}$,

$$G(\gamma(t)) = c \quad \Rightarrow \quad G'(\gamma(t))\gamma'(t) = 0.$$

In particular, suppose for $t = 0$ we have $\gamma(0) = p$ and $v = \gamma'(0)$ which makes $(p, v) \in T_pS$ with

$$G'(p)v = 0.$$

Recall $G : \mathbb{R}^k \times \mathbb{R}^p \to \mathbb{R}^p$ has an $p \times n$ derivative matrix where the j-th row is the gradient vector of the j-th component function. The equation $G'(p)v = 0$ gives us p-independent equations as we examine it componentwise. In particular, it reveals that (p, v) is orthogonal to $\nabla G_j(p)$ for $j = 1, 2, \dots, p$. We have derived the following theorem[4]:

Theorem 4.2.4.

Let $G : \mathbb{R}^k \times \mathbb{R}^p \to \mathbb{R}^p$ be a level-mappping which defines a k-dimensional level set S by $G^{-1}(c) = S$. The gradient vectors $\nabla G_j(p)$ are perpendicular to the tangent space at p; for each $j \in \mathbb{N}_p$

$$(p, \nabla(G_j(p))) \in (T_pS)^{\perp}.$$

It's time to do some counting. Observe that the mapping $\phi : \mathbb{R}^k \to T_pS$ defined by $\phi(v) = (p, v)$ is an isomorphism of vector spaces hence $dim(T_pS) = k$. But, by the same isomorphism we can see that $V_p = \phi(\mathbb{R}^k \times \mathbb{R}^p)$ hence $dim(V_p) = p + k$. In linear algebra we learn that if we have a

[4] notice if $G = (G_1, \dots, G_p)$ then $row_j(J_G) = \nabla G_j^T$. Since I have defined the gradient to be a column vector it follows we need a transpose to make ∇G_j into a row vector.

k-dimensional subspace W of an n-dimensional vector space V then the orthogonal complement $W^{\perp}$ is a subspace of V with **codimension** k. The term **codimension** is used to indicate a loss of dimension from the ambient space, in particular $dim(W^{\perp}) = n - k$. We should note that the direct sum of W and $W^{\perp}$ covers the whole space; $W \oplus W^{\perp} = V$. In the case of the tangent space, the codimension of $T_pS \leq V_p$ is found to be $p + k - k = p$. Thus $dim(T_pS)^{\perp} = p$. Any basis for this space must consist of p linearly independent vectors which are all orthogonal to the tangent space. Naturally, the subset of vectors $\{(p, (\nabla G_j(p)))_{j=1}^{p}$ forms just such a basis since it is given to be linearly independent by the $rank(G'(p)) = p$ condition. It follows that:

$$\boxed{(T_pS)^{\perp} \approx Row(G'(p)) = Col(G'(p)^T)}$$

where equality can be obtained by the slightly tedious equation

$$\boxed{(T_pS)^{\perp} = \phi(Col(G'(p)^T)) = \text{span}\{(p, (\nabla G_j(p)))_{j=1}^{p}}.$$

That equation simply does the following:

1. transpose $G'(p)$ to swap rows to columns
2. construct column space by taking span of columns in $G'(p)^T$
3. adjoin p to make pairs of vectors which live in V_p

many wiser authors wouldn't bother. The comments above are primarily about notation. Certainly hiding these details would make this section prettier, however, would it make it better? In linear algebra we learn that $(Row(A))^{\perp} = Null(A)$ and $Col(A)^{\perp} = Null(A^T)$. To see why these are true, consider $x \in Null(A)$ iff $Ax = 0$ iff $row_i(A) \bullet x = 0$ for each $i = 1, \dots, m$. Likewise, $y \in Null(A^T)$ iff $A^T y = 0$ iff $row_j(A^T)y = 0$ for each $j = 1, \dots, n$. But, $row_j(A^T) = col_j(A)^T$ hence $col_j(A)^T y = col_j(A) \bullet y = 0$ for each $j = 1, \dots, n$. Another useful identity for the "perp" is that $(A^{\perp})^{\perp} = A$. Consequently,

$$(T_pS)^{\perp} \approx Row(G'(p)) \quad \Rightarrow \quad T_pS \approx Row(G'(p))^{\perp} = Null(G'(p))$$

Let me once more replace $\approx$ by a more tedious, but explicit, procedure:

$$\boxed{T_pS = \phi(Null(G'(p)))}$$

Theorem 4.2.5.

Let $G : \mathbb{R}^k \times \mathbb{R}^p \to \mathbb{R}^p$ be a level-mappping which defines a k-dimensional level set S by $G^{-1}(c) = S$. The **tangent space** T_pS and the **normal space** N_pS at $p \in S$ are given by

$$T_pS = \{p\} \times Null(G'(p)) \qquad \& \qquad N_pS = \{p\} \times Col(G'(p)^T).$$

Moreover, $V_p = T_pS \oplus N_pS$. Every vector can be uniquely written as the sum of a tangent vector and a normal vector.

The fact that there are only tangents and normals is the key to the method of Lagrange multipliers. It forces two seemingly distinct objects to be in the same direction as one another.

Example 4.2.6. *Let $g : \mathbb{R}^4 \to \mathbb{R}$ be defined by $g(x, y, z, t) = t+x^2+y^2-2z^2$ note that $g(x, y, z, t) = 0$ gives a three dimensional subset of $\mathbb{R}^4$, let's call it M. Notice $\nabla g =< 2x, 2y, -4z, 1 >$ is nonzero everywhere. Let's focus on the point $(2, 2, 1, 0)$ note that $g(2, 2, 1, 0) = 0$ thus the point is on M. The tangent plane at $(2, 2, 1, 0)$ is formed from the union of all tangent vectors to $g = 0$ at the point $(2, 2, 1, 0)$. To find the equation of the tangent plane we suppose $\gamma : \mathbb{R} \to M$ is a curve with $\gamma' \neq 0$ and $\gamma(0) = (2, 2, 1, 0)$. By assumption $g(\gamma(s)) = 0$ since $\gamma(s) \in M$ for all $s \in \mathbb{R}$. Define $\gamma'(0) =< a, b, c, d >$, we find a condition from the chain-rule applied to $g \circ \gamma = 0$ at $s = 0$,*

$$\begin{aligned} \frac{d}{ds}\big(\ g\circ\gamma(s)\ \big) = \big(\nabla g\big)(\gamma(s))\cdot\gamma'(s) = 0 \qquad &\Rightarrow \qquad \nabla g(2,2,1,0)\cdot < a,b,c,d >= 0 \\ &\Rightarrow \qquad < 4,4,-4,1 > \ \cdot < a,b,c,d >= 0 \\ &\Rightarrow \qquad 4a + 4b - 4c + d = 0 \end{aligned}$$

Thus the equation of the tangent plane is $4(x - 2) + 4(y - 2) - 4(z - 1) + t = 0$. In invite the reader to find a vector in the tangent plane and check it is orthogonal to $\nabla g(2, 2, 1, 0)$. However, this should not be surprising, the condition the chain rule just gave us is just the statement that $< a, b, c, d >\in Null(\nabla g(2, 2, 1, 0)^T)$ and that is precisely the set of vector orthogonal to $\nabla g(2, 2, 1, 0)$.

Example 4.2.7. *Let $G : \mathbb{R}^4 \to \mathbb{R}^2$ be defined by $G(x, y, z, t) = (z + x^2 + y^2 - 2, z + y^2 + t^2 - 2)$. In this case $G(x, y, z, t) = (0, 0)$ gives a two-dimensional manifold in $\mathbb{R}^4$ let's call it M. Notice that $G_1 = 0$ gives $z + x^2 + y^2 = 2$ and $G_2 = 0$ gives $z + y^2 + t^2 = 2$ thus $G = 0$ gives the intersection of both of these three dimensional manifolds in $\mathbb{R}^4$ (no I can't "see" it either). Note,*

$$\nabla G_1 =< 2x, 2y, 1, 0 > \qquad \nabla G_2 =< 0, 2y, 1, 2t >$$

It turns out that the inverse mapping theorem says $G = 0$ describes a manifold of dimension 2 if the gradient vectors above form a linearly independent set of vectors. For the example considered here the gradient vectors are linearly dependent at the origin since $\nabla G_1(0) = \nabla G_2(0) = (0, 0, 1, 0)$. In fact, these gradient vectors are colinear along along the plane $x = t = 0$ since $\nabla G_1(0, y, z, 0) = \nabla G_2(0, y, z, 0) =< 0, 2y, 1, 0 >$. We again seek to contrast the tangent plane and its normal at some particular point. Choose $(1, 1, 0, 1)$ which is in M since $G(1, 1, 0, 1) = (0 + 1 + 1 - 2, 0 + 1 + 1 - 2) = (0, 0)$. Suppose that $\gamma : \mathbb{R} \to M$ is a path in M which has $\gamma(0) = (1, 1, 0, 1)$ whereas $\gamma'(0) =< a, b, c, d >$. Note that $\nabla G_1(1, 1, 0, 1) =< 2, 2, 1, 0 >$ and $\nabla G_2(1, 1, 0, 1) =< 0, 2, 1, 1 >$. Applying the chain rule to both G_1 and G_2 yields:

$$\begin{aligned} (G_1\circ\gamma)'(0) = \nabla G_1(\gamma(0))\cdot < a,b,c,d >= 0 \qquad &\Rightarrow \qquad < 2,2,1,0 > \ \cdot < a,b,c,d >= 0 \\ (G_2\circ\gamma)'(0) = \nabla G_2(\gamma(0))\cdot < a,b,c,d >= 0 \qquad &\Rightarrow \qquad < 0,2,1,1 > \ \cdot < a,b,c,d >= 0 \end{aligned}$$

This is two equations and four unknowns, we can solve it and write the vector in terms of two free variables correspondant to the fact the tangent space is two-dimensional. Perhaps it's easier to use matrix techiques to organize the calculation:

$$\begin{bmatrix} 2 & 2 & 1 & 0 \\ 0 & 2 & 1 & 1 \end{bmatrix}\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

We calculate[5], *$rref\begin{bmatrix} 2 & 2 & 1 & 0 \\ 0 & 2 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & -1/2 \\ 0 & 1 & 1/2 & 1/2 \end{bmatrix}$. It's natural to chose c, d as free variables then we can read that $a = d/2$ and $b = -c/2 - d/2$ hence*

$$< a, b, c, d >=< d/2, -c/2 - d/2, c, d >= \tfrac{c}{2} < 0, -1, 2, 0 > +\tfrac{d}{2} < 1, -1, 0, 2 >$$

[5] yes, I could short-cut this whole discussion and use the Theorem discussed earlier in this section to motivate the calculation of $Null(G'(p))$ for $p = (1, 1, 0, 1)$, my apologies these notes are a work in progress

We can see a basis for the tangent space. In fact, I can give parametric equations for the tangent space as follows:

$$X(u,v) = (1,1,0,1) + u < 0,-1,2,0 > +v < 1,-1,0,2 >$$

Not surprisingly the basis vectors of the tangent space are perpendicular to the gradient vectors $\nabla G_1(1,1,0,1) =< 2,2,1,0 >$ *and* $\nabla G_2(1,1,0,1) =< 0,2,1,1 >$ *which span the* **normal plane** N_p *to the tangent plane* T_p *at* $p = (1,1,0,1)$*. We find that* T_p *is orthogonal to* N_p*. In summary* $T_p^{\perp} = N_p$ *and* $T_p \oplus N_p = \mathbb{R}^4$*. This is just a fancy way of saying that the normal and the tangent plane only intersect at zero and they together span the entire ambient space.*

4.3 tangent and normal space from patches

I use the term **parametrization** in courses more basic than this, however, perhaps the term **patch** would be better. It's certainly easier to say and in our current context has the same meaning. I suppose the term **parametrization** is used in a bit less technical sense, so it fits calculus III better. In any event, we should make a definition of patched k-dimensional surface for the sake of concrete discussion in this section.

Definition 4.3.1.

Suppose $R : dom(R) \subseteq \mathbb{R}^k \to S \subseteq \mathbb{R}^n$. We say S is an **k-dimensional patch** iff $R'(t)$ has rank k for each $t \in dom(R)$. We also call S a k-dimensional parametrized subspace of $\mathbb{R}^n$.

The condition $R'(t)$ is just a slick way to say that the k-tangent vectors to S obtained by partial differentiation with respect to $t_1, \dots, t_k$ are linearly independent at $t = (t_1, \dots, t_k)$. I spent considerable effort justifying the formulae for the level-set case. I believe what follows should be intuitively clear given our previous efforts. Or, if that leaves you unsatisfied then read on to the examples. It's really not that complicated. This theorem is dual to Theorem 4.2.5.

Theorem 4.3.2.

Suppose $R : dom(R) \subseteq \mathbb{R}^k \to S \subseteq \mathbb{R}^n$ defines a **k-dimensional patch** of S. The **tangent space** T_pS and the **normal space** at $p = R(t) \in S$ are given by

$$T_pS = \{p\} \times Col(R'(t)) \qquad \& \qquad N_pS = \{p\} \times Col(R'(t))^{\perp} = \{p\} \times Null(R'(t)^T).$$

Moreover, $V_p = T_pS \oplus N_pS$. Every vector can be uniquely written as the sum of a tangent vector and a normal vector.

Once again, the vector space structure of T_pS and N_pS is given by the addition of vectors based at p. Let us begin with a reasonably simple example.

Example 4.3.3. *Let* $R : \mathbb{R}^2 \to \mathbb{R}^3$ *with* $R(x,y) = (x,y,xy)$ *define* $S \subset \mathbb{R}^3$*. We calculate,*

$$R'(x,y) = [\partial_x R | \partial_y R] = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ y & x \end{bmatrix}$$

If $p = (a,b,ab) \in S$ *then* $T_pS = \{(a,b,ab)\} \times span\{(1,0,b),(0,1,a)\}$*. The normal space is found from* $Null(R'(a,b)^T)$*. A short calculation shows that*

$$Null \begin{bmatrix} 1 & 0 & b \\ 0 & 1 & a \end{bmatrix} = span\{(-b,-a,1)\}$$

As a quick check, note $(1,0,b) \bullet (-b,-a,1) = 0$ *and* $(0,1,a) \bullet (-b,-a,1) = 0$*. We conclude, for* $p = (a,b,ab)$ *the normal space is simply:*

$$N_pS = \{(a,b,ab)\} \times span\{(-b,-a,1)\}.$$

In the previous example, we could rightly call T_pS the tangent plane at p and N_pS the normal line through p. Moreover, we could have used three-dimensional vector analysis to find the normal line from the cross-product. However, that will not be possible in what follows:

Example 4.3.4. *Let* $R : \mathbb{R}^2 \to \mathbb{R}^4$ *with* $R(s,t) = (s^2, t^2, t, s)$ *define* $S \subset \mathbb{R}^4$*. We calculate,*

$$R'(s,t) = [\partial_s R | \partial_t R] = \begin{bmatrix} 2s & 0 \\ 0 & 2t \\ 0 & 1 \\ 1 & 0 \end{bmatrix}$$

If $p = (1,9,3,1) \in S$ *then* $T_pS = \{(1,9,3,1)\} \times span\{(2,0,0,1),(0,6,3,0)\}$*. The normal space is found from* $Null(R'(1,3)^T)$*. A short calculation shows that*

$$Null \begin{bmatrix} 2 & 0 & 0 & 1 \\ 0 & 6 & 3 & 0 \end{bmatrix} = span\{(-1,0,0,2),(0,-3,6,0)\}$$

We conclude, for $p = (1,9,3,1)$ *the normal space is simply:*

$$N_pS = \{(1,9,3,1)\} \times span\{(-1,0,0,2),(0,-3,6,0)\}.$$

4.4 summary of tangent and normal spaces

Let me briefly draw together what we did thus far in this chapter: the notation below given in I is also used in $II.$ and $III.$ We studied a k-dimensional submanifold of $\mathbb{R}^n$ from two viewpoints. Let us contrast the set-ups of space, tangent space and normal space for (a.) the parametric or explicit view point and (b.) the solution set or implicit viewpoint:

(I.) a set S has dimension k if

(a) $\{\partial_1 R(t), \dots, \partial_k R(t)\}$ is pointwise linearly independent at each $t \in U$ where $R : U \to S$ is a patch.

(b) $\text{rank}(F'(x)) = p$ for all $x \in \tilde{S}$ where $\tilde{S}$ is open and contains $S = F^{-1}\{c\}$ for continuously differentiable $F : \mathbb{R}^k \times \mathbb{R}^p \to \mathbb{R}^p$

(II.) the tangent space at x_o for the k-dimensional set S is found from:

(a) attaching the span of the vectors $\{\partial_1 R(t_o), \dots, \partial_k R(t_o)\}$ to $x_o = R(t_o) \in S$.

(b) attaching the orthogonal complement of $\{\nabla F_1(x_o), \dots, \nabla F_p(x_o)\}$ to $x_o \in S$. To calculate the basis for the orthogonal complement we find basis for $Null(F'(x_o))$.

(III.) the normal space to a k-dimensional set S (embedded in $\mathbb{R}^n$) is found from:

(a) attaching $\{\partial_1 R(t_o), \dots, \partial_k R(t_o)\}^\perp$ to $x_o = R(t_o)$. To calculate the basis for normal space we can find a basis for $Null(R'(t)^T)$ and attach it to x_o.

(b) attaching $\{\nabla F_1(x_o), \dots, \nabla F_p(x_o)\}$ to $x_o \in S$ to $x_o \in S$.

I should mention, while it is possible to switch viewpoints for a given example which is sufficiently simple, it is generally a difficult problem, indeed an *art* to exchange the implicit for the explicit presentation of a submanifold or vice-versa.

4.5 method of Lagrange mulitpliers

Let us begin with a statement of the problem we wish to solve.

Problem: given an objective function $f : \mathbb{R}^n \to \mathbb{R}$ and continuously differentiable constraint function $G : \mathbb{R}^n \to \mathbb{R}^p$, find extreme values for the objective function f relative to the constraint $G(x) = c$.

Note that $G(x) = c$ is a vector notation for p-scalar equations. If we suppose $rank(G'(x)) = p$ then the constraint surface $G(x) = c$ will form an $(n-p)$-dimensional level set. Let us make that supposition throughout the remainder of this section.

In order to solve a problem it is sometimes helpful to find necessary conditions by assuming an answer exists. Let us do that here. Suppose x_o maps to the local extrema of $f(x_o)$ on $S = G^{-1}\{c\}$. This means there exists an open ball around x_o for which $f(x_o)$ is either an upper or lower bound of all the values of f over the ball intersected with S. One clear implication of this data is that if we take any continuously differentiable curve on S which passes through x_o, say $\gamma : \mathbb{R} \to \mathbb{R}^n$ with $\gamma(0) = x_o$ and $G(\gamma(t)) = c$ for all t, then the composite $f \circ \gamma$ is a function on $\mathbb{R}$ which takes an extreme value at $t = 0$. Fermat's theorem from calculus I applies and as $f \circ \gamma$ is differentiable near $t = 0$ we find $(f \circ \gamma)'(0) = 0$ is a necessary condition. But, this means we have two necessary conditions on γ:

1. $G(\gamma(t)) = c$
2. $(f \circ \gamma)'(0) = 0$

Let us expand a bit on both of these conditions:

1. $G'(x_o)\gamma'(0) = 0$
2. $f'(x_o)\gamma'(0) = 0$

The first of these conditions places $\gamma'(0) \in T_{x_o}S$ but then the second condition says that $f'(x_o) = (\nabla f)(x_o)^T$ is orthogonal to $\gamma'(0)$ hence $(\nabla f)(x_o)^T \in N_{x_o}$. Now, recall from the last section that the gradient vectors of the component functions to G span the normal space, this means any vector in N_{x_o} can be written as a linear combination of the gradient vectors. In particular, this means there exist constants $\lambda_1, \lambda_2, \dots, \lambda_p$ such that

$$(\nabla f)(x_o)^T = \lambda_1(\nabla G_1)(x_o)^T + \lambda_2(\nabla G_2)(x_o)^T + \cdots + \lambda_p(\nabla G_p)(x_o)^T$$

We may summarize the method of Lagrange multipliers as follows:

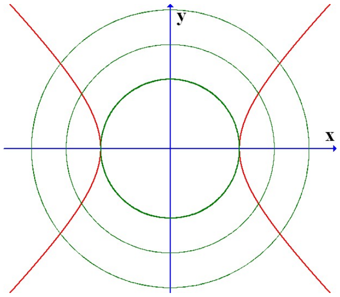
1. **choose n-variables which aptly describe your problem.**
2. **identify your objective function and write all constraints as level surfaces.**
3. **solve $\nabla f = \lambda_1 \nabla G_1 + \lambda_2 \nabla G_2 + \cdots + \lambda_p \nabla G_p$ subject to the constraint $G(x) = c$.**
4. **test the validity of your proposed extremal points.**

The obvious gap in the method is the supposition that an extrema exists for the restriction $f|_S$. Well examine a few examples before I reveal a sufficient condition. We'll also see how absence of that sufficient condition does allow the method to fail.

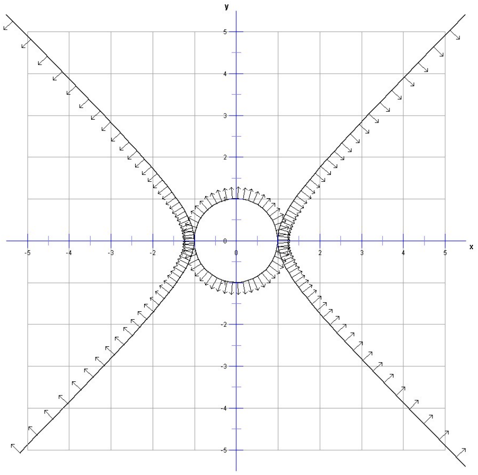
Example 4.5.1. *Suppose we wish to find maximum and minimum distance to the origin for points on the curve $x^2 - y^2 = 1$. In this case we can use the distance-squared function as our objective $f(x, y) = x^2 + y^2$ and the single constraint function is $g(x, y) = x^2 - y^2$. Observe that $\nabla f =< 2x, 2y >$ whereas $\nabla g =< 2x, -2y >$. We seek solutions of $\nabla f = \lambda \nabla g$ which gives us $< 2x, 2y >= \lambda < 2x, -2y >$. Hence $2x = 2\lambda x$ and $2y = -2\lambda y$. We must solve these equations subject to the condition $x^2 - y^2 = 1$. Observe that $x = 0$ is not a solution since $0 - y^2 = 1$ has no real solution. On the other hand, $y = 0$ does fit the constraint and $x^2 - 0 = 1$ has solutions $x = \pm 1$. Consider then*

$$2x = 2\lambda x \quad and \quad 2y = -2\lambda y \qquad \Rightarrow \qquad x(1 - \lambda) = 0 \quad and \quad y(1 + \lambda) = 0$$

Since $x \neq 0$ on the constraint curve it follows that $1 - \lambda = 0$ hence $\lambda = 1$ and we learn that $y(1 + 1) = 0$ hence $y = 0$. Consequently, $(1, 0$ and $(-1, 0)$ are the two point where we expect to find extreme-values of f. In this case, the method of Lagrange multipliers served it's purpose, as you can see in the graph. Below the green curves are level curves of the objective function whereas the particular red curve is the given constraint curve.



The picture below is a screen-shot of the Java applet created by David Lippman and Konrad Polthier to explore 2D and 3D graphs. Especially nice is the feature of adding vector fields to given objects, many other plotters require much more effort for similar visualization. See more at the website: http://dlippman.imathas.com/g1/GrapherLaunch.html.



Note how the gradient vectors to the objective function and constraint function line-up nicely at those points.

In the previous example, we actually got lucky. There are examples of this sort where we could get false maxima due to the nature of the constraint function.

Example 4.5.2. *Suppose we wish to find the points on the unit circle $g(x,y) = x^2 + y^2 = 1$ which give extreme values for the objective function $f(x,y) = x^2 - y^2$. Apply the method of Lagrange multipliers and seek solutions to $\nabla f = \lambda \nabla g$:*

$$< 2x, -2y >= \lambda < 2x, 2y >$$

We must solve $2x = 2x\lambda$ which is better cast as $(1-\lambda)x = 0$ and $-2y = 2\lambda y$ which is nicely written as $(1+\lambda)y = 0$. On the basis of these equations alone we have several options:

1. *if $\lambda = 1$ then $(1+1)y = 0$ hence $y = 0$*
2. *if $\lambda = -1$ then $(1-(1))x = 0$ hence $x = 0$*

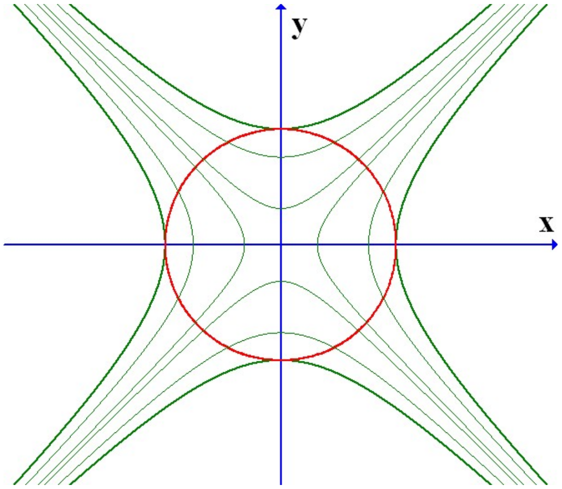
But, we also must fit the constraint $x^2 + y^2 = 1$ hence we find four solutions:

1. *if $\lambda = 1$ then $y = 0$ thus $x^2 = 1 \quad \Rightarrow \quad x = \pm 1 \quad \Rightarrow \quad (\pm 1, 0)$*
2. *if $\lambda = -1$ then $x = 0$ thus $y^2 = 1 \quad \Rightarrow \quad y = \pm 1 \quad \Rightarrow \quad (0, \pm 1)$*

We test the objective function at these points to ascertain which type of extrema we've located:

$$f(0,\pm 1) = 0^2 - (\pm 1)^2 = -1 \qquad \& \qquad f(\pm 1, 0) = (\pm 1)^2 - 0^2 = 1$$

When constrained to the unit circle we find the objective function attains a maximum value of 1 at the points $(1,0)$ and $(-1,0)$ and a minimum value of -1 at $(0,1)$ and $(0,-1)$. Let's illustrate the answers as well as a few non-answers to get perspective. Below the green curves are level curves of the objective function whereas the particular red curve is the given constraint curve.



The success of the last example was no accident. The fact that the constraint curve was a circle which is a closed and bounded subset of $\mathbb{R}^2$ means that is is a **compact** subset of $\mathbb{R}^2$. A well-known theorem of analysis states that any real-valued continuous function on a compact domain attains both maximum and minimum values. The objective function is continuous and the domain is compact hence the theorem applies and the method of Lagrange multipliers succeeds. In contrast, the constraint curve of the preceding example was a hyperbola which is not compact. We have no assurance of the existence of any extrema. Indeed, we only found minima but no maxima in Example 4.5.1.

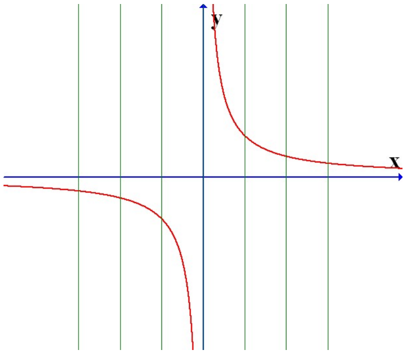
The generality of the method of Lagrange multipliers is naturally limited to smooth constraint curves and smooth objective functions. We must insist the gradient vectors exist at all points of inquiry. Otherwise, the method breaks down. If we had a constraint curve which has sharp corners then the method of Lagrange breaks down at those corners. In addition, if there are points of discontinuity in the constraint then the method need not apply. This is not terribly surprising, even in calculus I the main attack to analyze extrema of function on $\mathbb{R}$ assumed continuity, differentiability and sometimes twice differentiability. Points of discontinuity require special attention in whatever context you meet them.

At this point it is doubtless the case that some of you are, to misquote an ex-student of mine, "not-impressed". Perhaps the following examples better illustrate the dangers of non-compact constraint curves.

Example 4.5.3. *Suppose we wish to find extrema of $f(x,y) = x$ when constrained to $xy = 1$. Identify $g(x,y) = xy = 1$ and apply the method of Lagrange multipliers and seek solutions to $\nabla f = \lambda \nabla g$:*

$$< 1, 0 >= \lambda < y, x > \quad \Rightarrow \quad 1 = \lambda y \quad \textit{and} \quad 0 = \lambda x$$

If $\lambda = 0$ then $1 = \lambda y$ is impossible to solve hence $\lambda \neq 0$ and we find $x = 0$. But, if $x = 0$ then $xy = 1$ is not solvable. Therefore, we find no solutions. Well, I suppose we have succeeded here in a way. We just learned there is no extreme value of x on the hyperbola $xy = 1$. Below the green curves are level curves of the objective function whereas the particular red curve is the given constraint curve.



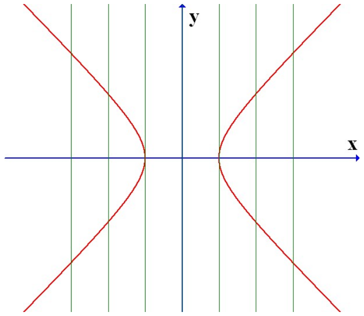
Example 4.5.4. *Suppose we wish to find extrema of $f(x,y) = x$ when constrained to $x^2 - y^2 = 1$. Identify $g(x,y) = x^2 - y^2 = 1$ and apply the method of Lagrange multipliers and seek solutions to $\nabla f = \lambda \nabla g$:*

$$< 1, 0 >= \lambda < 2x, -2y > \quad \Rightarrow \quad 1 = 2\lambda x \quad \textit{and} \quad 0 = -2\lambda y$$

If $\lambda = 0$ then $1 = 2\lambda x$ is impossible to solve hence $\lambda \neq 0$ and we find $y = 0$. If $y = 0$ and $x^2 - y^2 = 1$ then we must solve $x^2 = 1$ whence $x = \pm 1$. We are tempted to conclude that:

1. *the objective function $f(x,y) = x$ attains a maximum on $x^2 - y^2 = 1$ at $(1,0)$ since $f(1,0) = 1$*
2. *the objective function $f(x,y) = x$ attains a minimum on $x^2 - y^2 = 1$ at $(-1,0)$ since $f(1,0) = -1$*

But, both conclusions are false. Note $\sqrt{2}^2 - 1^2 = 1$ hence $(\pm\sqrt{2}, 1)$ are points on the constraint curve and $f(\sqrt{2}, 1) = \sqrt{2}$ and $f(-\sqrt{2}, 1) = -\sqrt{2}$. The error of the method of Lagrange multipliers in this context is the supposition that there exists extrema to find, in this case there are no such points. It is possible for the gradient vectors to line-up at points where there are no extrema. Below the green curves are level curves of the objective function whereas the particular red curve is the given constraint curve.



Incidentally, if you want additional discussion of Lagrange multipliers for two-dimensional problems one very nice source I certainly profitted from was the YouTube video by Edward Frenkel of Berkley. See his website http://math.berkeley.edu/ frenkel/ for links.

Chapter 5

critical point analysis for several variables

In the typical calculus sequence you learn the first and second derivative tests in calculus I. Then in calculus II you learn about power series and Taylor's Theorem. Finally, in calculus III, in many popular texts, you learn an essentially ad-hoc procedure for judging the nature of critical points as minimum, maximum or saddle. These topics are easily seen as disconnected events. In this chapter, we connect them. We learn that the geometry of quadratic forms is ellegantly revealed by eigenvectors and more than that this geometry is precisely what elucidates the proper classifications of critical points of multivariate functions with real values.

5.1 multivariate power series

We set aside the issue of convergence for now. We will suppose the series discussed in this section exist on and converge on some domain, but we do not seek to treat that topic here. Our focus is computational. How should we calculate the Taylor series for $f(x,y)$ at (a,b)? Or, what about $f(x)$ at $x_o \in \mathbb{R}^n$?.

5.1.1 taylor's polynomial for one-variable

If $f : U \subseteq \mathbb{R} \to \mathbb{R}$ is analytic at $x_o \in U$ then we can write

$$f(x) = f(x_o) + f'(x_o)(x - x_o) + \frac{1}{2}f''(x_o)(x - x_o)^2 + \cdots = \sum_{n=0}^{\infty} \frac{f^{(n)}(x_o)}{n!}(x - x_o)^n$$

We could write this in terms of the operator $D = \frac{d}{dt}$ and the evaluation of $t = x_o$

$$f(x) = \left[\sum_{n=0}^{\infty} \frac{1}{n!}(x - t)^n D^n f(t) \right]_{t=x_o} =$$

I remind the reader that a function is called **entire** if it is analytic on all of $\mathbb{R}$, for example $e^x, \cos(x)$ and $\sin(x)$ are all entire. In particular, you should know that:

$$e^x = 1 + x + \frac{1}{2}x^2 + \cdots = \sum_{n=0}^{\infty} \frac{1}{n!}x^n$$

$$\cos(x) = 1 - \frac{1}{2}x^2 + \frac{1}{4!}x^4 \cdots = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!}x^{2n}$$

$$\sin(x) = x - \frac{1}{3!}x^3 + \frac{1}{5!}x^5 \cdots = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!}x^{2n+1}$$

Since $e^x = \cosh(x) + \sinh(x)$ it also follows that

$$\cosh(x) = 1 + \frac{1}{2}x^2 + \frac{1}{4!}x^4 \cdots = \sum_{n=0}^{\infty} \frac{1}{(2n)!}x^{2n}$$

$$\sinh(x) = x + \frac{1}{3!}x^3 + \frac{1}{5!}x^5 \cdots = \sum_{n=0}^{\infty} \frac{1}{(2n+1)!}x^{2n+1}$$

The geometric series is often useful, for $a, r \in \mathbb{R}$ with $|r| < 1$ it is known

$$a + ar + ar^2 + \cdots = \sum_{n=0}^{\infty} ar^n = \frac{a}{1-r}$$

This generates a whole host of examples, for instance:

$$\frac{1}{1+x^2} = 1 - x^2 + x^4 - x^6 + \cdots$$

$$\frac{1}{1-x^3} = 1 + x^3 + x^6 + x^9 + \cdots$$

$$\frac{x^3}{1-2x} = x^3(1 + 2x + (2x)^2 + \cdots) = x^3 + 2x^4 + 4x^5 + \cdots$$

Moreover, the term-by-term integration and differentiation theorems yield additional results in conjuction with the geometric series:

$$\tan^{-1}(x) = \int \frac{dx}{1+x^2} = \int \sum_{n=0}^{\infty}(-1)^n x^{2n} dx = \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1}x^{2n+1} = x - \frac{1}{3}x^3 + \frac{1}{5}x^5 + \cdots$$

$$\ln(1-x) = \int \frac{d}{dx}\ln(1-x)dx = \int \frac{-1}{1-x}dx = -\int \sum_{n=0}^{\infty} x^n dx = \sum_{n=0}^{\infty} \frac{-1}{n+1}x^{n+1}$$

Of course, these are just the basic building blocks. We also can twist things and make the student use algebra,

$$e^{x+2} = e^x e^2 = e^2(1 + x + \frac{1}{2}x^2 + \cdots)$$

or trigonmetric identities,

$$\sin(x) = \sin(x - 2 + 2) = \sin(x-2)\cos(2) + \cos(x-2)\sin(2)$$

$$\Rightarrow \quad \sin(x) = \cos(2)\sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!}(x-2)^{2n+1} + \sin(2)\sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!}(x-2)^{2n}.$$

Feel free to peruse my most recent calculus II materials to see a host of similarly sneaky calculations.

5.1.2 taylor's multinomial for two-variables

Suppose we wish to find the taylor polynomial centered at $(0,0)$ for $f(x,y) = e^x \sin(y)$. It is a simple as this:

$$f(x,y) = \Big(1 + x + \frac{1}{2}x^2 + \cdots\Big)\Big(y - \frac{1}{6}y^3 + \cdots\Big) = y + xy + \frac{1}{2}x^2y - \frac{1}{6}y^3 + \cdots$$

the resulting expression is called a multinomial since it is a polynomial in multiple variables. If all functions $f(x,y)$ could be written as $f(x,y) = F(x)G(y)$ then multiplication of series known from calculus II would often suffice. However, many functions do not possess this very special form. For example, how should we expand $f(x,y) = \cos(xy)$ about $(0,0)$?. We need to derive the two-dimensional Taylor's theorem.

We already know Taylor's theorem for functions on $\mathbb{R}$,

$$g(x) = g(a) + g'(a)(x-a) + \frac{1}{2}g''(a)(x-a)^2 + \cdots + \frac{1}{k!}g^{(k)}(a)(x-a)^k + R_k$$

and... If the remainder term vanishes as $k \to \infty$ then the function g is represented by the Taylor series given above and we write:

$$g(x) = \sum_{k=0}^{\infty} \frac{1}{k!} g^{(k)}(a)(x-a)^k.$$

Consider the function of two variables $f : U \subseteq \mathbb{R}^2 \to \mathbb{R}$ which is smooth with smooth partial derivatives of all orders. Furthermore, let $(a,b) \in U$ and construct a line through (a,b) with direction vector (h_1, h_2) as usual:

$$\phi(t) = (a,b) + t(h_1,h_2) = (a + th_1, b + th_2)$$

for $t \in \mathbb{R}$. Note $\phi(0) = (a,b)$ and $\phi'(t) = (h_1,h_2) = \phi'(0)$. Construct $g = f \circ \phi : \mathbb{R} \to \mathbb{R}$ and choose $dom(g)$ such that $\phi(t) \in U$ for $t \in dom(g)$. This function g is a real-valued function of a real variable and we will be able to apply Taylor's theorem from calculus II on g. However, to differentiate g we'll need tools from calculus III to sort out the derivatives. In particular, as we differentiate g, note we use the chain rule for functions of several variables:

$$\begin{aligned} g'(t) = (f\circ\phi)'(t) &= f'(\phi(t))\phi'(t) \\ &= \nabla f(\phi(t)) \cdot (h_1,h_2) \\ &= h_1 f_x(a+th_1, b+th_2) + h_2 f_y(a+th_1, b+th_2) \end{aligned}$$

Note $g'(0) = h_1 f_x(a,b) + h_2 f_y(a,b)$. Differentiate again (I omit $(\phi(t))$ dependence in the last steps),

$$\begin{aligned} g''(t) &= h_1 f_x'(a+th_1, b+th_2) + h_2 f_y'(a+th_1, b+th_2) \\ &= h_1 \nabla f_x(\phi(t)) \cdot (h_1,h_2) + h_2 \nabla f_y(\phi(t)) \cdot (h_1,h_2) \\ &= h_1^2 f_{xx} + h_1h_2 f_{yx} + h_2h_1 f_{xy} + h_2^2 f_{yy} \\ &= h_1^2 f_{xx} + 2h_1h_2 f_{xy} + h_2^2 f_{yy} \end{aligned}$$

Thus, making explicit the point dependence, $g''(0) = h_1^2 f_{xx}(a,b) + 2h_1h_2 f_{xy}(a,b) + h_2^2 f_{yy}(a,b)$. We may construct the Taylor series for g up to quadratic terms:

$$\begin{aligned} g(0+t) &= g(0) + tg'(0) + \frac{1}{2}g''(0) + \cdots \\ &= f(a,b) + t[h_1 f_x(a,b) + h_2 f_y(a,b)] + \frac{t^2}{2}\left[h_1^2 f_{xx}(a,b) + 2h_1h_2 f_{xy}(a,b) + h_2^2 f_{yy}(a,b)\right] + \cdots \end{aligned}$$

Note that $g(t) = f(a + th_1, b + th_2)$ hence $g(1) = f(a + h_1, b + h_2)$ and consequently,

$$f(a + h_1, b + h_2) = f(a, b) + h_1 f_x(a, b) + h_2 f_y(a, b) + \\ + \frac{1}{2}\left[h_1^2 f_{xx}(a, b) + 2h_1h_2 f_{xy}(a, b) + h_2^2 f_{yy}(a, b)\right] + \cdots$$

Omitting point dependence on the 2^{nd} derivatives,

$$\boxed{f(a + h_1, b + h_2) = f(a, b) + h_1 f_x(a, b) + h_2 f_y(a, b) + \tfrac{1}{2}\left[h_1^2 f_{xx} + 2h_1h_2 f_{xy} + h_2^2 f_{yy}\right] + \cdots}$$

Sometimes we'd rather have an expansion about (x, y). To obtain that formula simply substitute $x - a = h_1$ and $y - b = h_2$. Note that the point (a, b) is fixed in this discussion so the derivatives are not modified in this substitution,

$$f(x, y) = f(a, b) + (x - a)f_x(a, b) + (y - b)f_y(a, b) + \\ + \frac{1}{2}\left[(x - a)^2 f_{xx}(a, b) + 2(x - a)(y - b)f_{xy}(a, b) + (y - b)^2 f_{yy}(a, b)\right] + \cdots$$

At this point we ought to recognize the first three terms give the tangent plane to $z = f(z, y)$ at $(a, b, f(a, b))$. The higher order terms are nonlinear corrections to the linearization, these quadratic terms form a *quadratic form.* If we computed third, fourth or higher order terms we will find that, using $a = a_1$ and $b = a_2$ as well as $x = x_1$ and $y = x_2$,

$$\boxed{f(x, y) = \sum_{n=0}^{\infty} \sum_{i_1=0}^{2} \sum_{i_2=0}^{2} \cdots \sum_{i_n=0}^{2} \frac{1}{n!} \frac{\partial^{(n)} f(a_1, a_2)}{\partial x_{i_1} \partial x_{i_2} \cdots \partial x_{i_n}} (x_{i_1} - a_{i_1})(x_{i_2} - a_{i_2}) \cdots (x_{i_n} - a_{i_n})}$$

Example 5.1.1. *Expand* $f(x, y) = \cos(xy)$ *about* $(0, 0)$. *We calculate derivatives,*

$$f_x = -y \sin(xy) \qquad f_y = -x \sin(xy)$$

$$f_{xx} = -y^2 \cos(xy) \qquad f_{xy} = -\sin(xy) - xy \cos(xy) \qquad f_{yy} = -x^2 \cos(xy)$$

$$f_{xxx} = y^3 \sin(xy) \qquad f_{xxy} = -y \cos(xy) - y \cos(xy) + xy^2 \sin(xy)$$

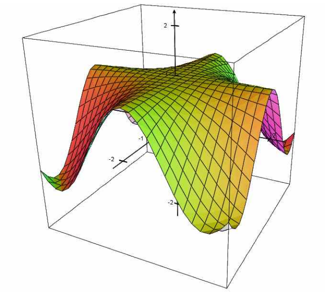
$$f_{xyy} = -x \cos(xy) - x \cos(xy) + x^2 y \sin(xy) \qquad f_{yyy} = x^3 \sin(xy)$$

Next, evaluate at $x = 0$ *and* $y = 0$ *to find* $f(x, y) = 1 + \cdots$ *to third order in* x, y *about* $(0, 0)$. *We can understand why these derivatives are all zero by approaching the expansion a different route: simply expand cosine directly in the variable* (xy),

$$f(x, y) = 1 - \frac{1}{2}(xy)^2 + \frac{1}{4!}(xy)^4 + \cdots = 1 - \frac{1}{2}x^2y^2 + \frac{1}{4!}x^4y^4 + \cdots .$$

Apparently the given function only has nontrivial derivatives at $(0, 0)$ *at orders* $0, 4, 8, ...$ *We can deduce that* $f_{xxxxy}(0, 0) = 0$ *without furthter calculation.*

This is actually a very interesting function, I think it defies our analysis in the later portion of this chapter. The second order part of the expansion reveals nothing about the nature of the critical point $(0, 0)$. Of course, any student of trigonometry should recognize that $f(0, 0) = 1$ is likely a local maximum, it's certainly not a local minimum. The graph reveals that $f(0, 0)$ is a local maxium for f restricted to certain rays from the origin whereas it is constant on several special directions (the coordinate axes).



And, if you were wondering, yes, we could also derive this from subsitution of $u = xy$ into the standard expansion for $\cos(u) = 1 - \frac{1}{2}u^2 + \frac{1}{4!}u^4 + \cdots$. Often such subsitutions are the quickest way to generate interesting examples.

5.1.3 taylor's multinomial for many-variables

Suppose $f : dom(f) \subseteq \mathbb{R}^n \to \mathbb{R}$ is a function of n-variables and we seek to derive the Taylor series centered at $a = (a_1, a_2, \dots, a_n)$. Once more consider the composition of f with a line in $dom(f)$. In particular, let $\phi : \mathbb{R} \to \mathbb{R}^n$ be defined by $\phi(t) = a + th$ where $h = (h_1, h_2, \dots, h_n)$ gives the direction of the line and clearly $\phi'(t) = h$. Let $g : dom(g) \subseteq \mathbb{R} \to \mathbb{R}$ be defined by $g(t) = f(\phi(t))$ for all $t \in \mathbb{R}$ such that $\phi(t) \in dom(f)$. Differentiate, use the multivariate chain rule, recall here that $\nabla = e_1\frac{\partial}{\partial x_1} + e_2\frac{\partial}{\partial x_2} + \cdots + e_n\frac{\partial}{\partial x_n} = \sum_{i=1}^n e_i\partial_i$,

$$g'(t) = \nabla f(\phi(t)) \cdot \phi'(t) = \nabla f(\phi(t)) \cdot h = \sum_{i=1}^n h_i(\partial_i f)(\phi(t))$$

If we omit the explicit dependence on $\phi(t)$ then we find the simple formula $g'(t) = \sum_{i=1}^n h_i\partial_i f$. Differentiate a second time,

$$g''(t) = \frac{d}{dt}\left[\sum_{i=1}^n h_i\partial_i f(\phi(t))\right] = \sum_{i=1}^n h_i\frac{d}{dt}\Big[(\partial_i f)(\phi(t))\Big] = \sum_{i=1}^n h_i\big(\nabla\partial_i f\big)(\phi(t)) \cdot \phi'(t)$$

Omitting the $\phi(t)$ dependence and once more using $\phi'(t) = h$ we find

$$g''(t) = \sum_{i=1}^n h_i\nabla\partial_i f \cdot h$$

Recall that $\nabla = \sum_{j=1}^n e_j\partial_j$ and expand the expression above,

$$g''(t) = \sum_{i=1}^n h_i\left(\sum_{j=1}^n e_j\partial_j\partial_i f\right) \cdot h = \sum_{i=1}^n\sum_{j=1}^n h_ih_j\partial_j\partial_i f$$

where we should remember $\partial_j\partial_i f$ depends on $\phi(t)$. It should be clear that if we continue and take k-derivatives then we will obtain:

$$g^{(k)}(t) = \sum_{i_1=1}^n\sum_{i_2=1}^n\cdots\sum_{i_k=1}^n h_{i_1}h_{i_2}\cdots h_{i_k}\partial_{i_1}\partial_{i_2}\cdots\partial_{i_k}f$$

More explicitly,

$$g^{(k)}(t) = \sum_{i_1=1}^n\sum_{i_2=1}^n\cdots\sum_{i_k=1}^n h_{i_1}h_{i_2}\cdots h_{i_k}(\partial_{i_1}\partial_{i_2}\cdots\partial_{i_k}f)(\phi(t))$$

Hence, by Taylor's theorem, provided we are sufficiently close to $t = 0$ as to bound the remainder[1]

$$g(t) = \sum_{k=0}^{\infty} \frac{1}{k!}\Big(\sum_{i_1=1}^{n}\sum_{i_2=1}^{n}\cdots\sum_{i_k=1}^{n} h_{i_1}h_{i_2}\cdots h_{i_k}(\partial_{i_1}\partial_{i_2}\cdots\partial_{i_k}f)(\phi(t))\Big)t^k$$

Recall that $g(t) = f(\phi(t)) = f(a + th)$. Put[2] $t = 1$ and bring in the $\frac{1}{k!}$ to derive

$$\boxed{f(a+h) = \sum_{k=0}^{\infty}\sum_{i_1=1}^{n}\sum_{i_2=1}^{n}\cdots\sum_{i_k=1}^{n} \frac{1}{k!}\big(\partial_{i_1}\partial_{i_2}\cdots\partial_{i_k}f\big)(a)\ h_{i_1}h_{i_2}\cdots h_{i_k}.}$$

Naturally, we sometimes prefer to write the series expansion about a as an expresssion in $x = a+h$. With this substitution we have $h = x - a$ and $h_{i_j} = (x-a)_{i_j} = x_{i_j} - a_{i_j}$ thus

$$\boxed{f(x) = \sum_{k=0}^{\infty}\sum_{i_1=1}^{n}\sum_{i_2=1}^{n}\cdots\sum_{i_k=1}^{n} \frac{1}{k!}\big(\partial_{i_1}\partial_{i_2}\cdots\partial_{i_k}f\big)(a)\ (x_{i_1}-a_{i_1})(x_{i_2}-a_{i_2})\cdots(x_{i_k}-a_{i_k}).}$$

Example 5.1.2. *Suppose $f : \mathbb{R}^3 \to \mathbb{R}$ let's unravel the Taylor series centered at $(0,0,0)$ from the general formula boxed above. Utilize the notation $x = x_1, y = x_2$ and $z = x_3$ in this example.*

$$f(x) = \sum_{k=0}^{\infty}\sum_{i_1=1}^{3}\sum_{i_2=1}^{3}\cdots\sum_{i_k=1}^{3} \frac{1}{k!}\big(\partial_{i_1}\partial_{i_2}\cdots\partial_{i_k}f\big)(0)\ x_{i_1}x_{i_2}\cdots x_{i_k}.$$

The terms to order 2 are as follows:

$$\begin{aligned} f(x) &= f(0) + f_x(0)x + f_y(0)y + f_z(0)z \\ &+\tfrac{1}{2}\Big(f_{xx}(0)x^2 + f_{yy}(0)y^2 + f_{zz}(0)z^2 + \\ &\quad +f_{xy}(0)xy + f_{xz}(0)xz + f_{yz}(0)yz + f_{yx}(0)yx + f_{zx}(0)zx + f_{zy}(0)zy \Big) + \cdots \end{aligned}$$

Partial derivatives commute for smooth functions hence,

$$\begin{aligned} f(x) &= f(0) + f_x(0)x + f_y(0)y + f_z(0)z \\ &+\tfrac{1}{2}\Big(f_{xx}(0)x^2 + f_{yy}(0)y^2 + f_{zz}(0)z^2 + 2f_{xy}(0)xy + 2f_{xz}(0)xz + 2f_{yz}(0)yz \Big) \\ &+\tfrac{1}{3!}\Big(f_{xxx}(0)x^3 + f_{yyy}(0)y^3 + f_{zzz}(0)z^3 + 3f_{xxy}(0)x^2y + 3f_{xxz}(0)x^2z \\ &\quad +3f_{yyz}(0)y^2z + 3f_{xyy}(0)xy^2 + 3f_{xzz}(0)xz^2 + 3f_{yzz}(0)yz^2 + 6f_{xyz}(0)xyz \Big) + \cdots \end{aligned}$$

[1] there exist smooth examples for which no neighborhood is small enough, the bump function in one-variable has higher-dimensional analogues, we focus our attention to functions for which it is possible for the series below to converge

[2] if $t = 1$ is not in the domain of g then we should rescale the vector h so that $t = 1$ places $\phi(1)$ in $dom(f)$, if f is smooth on some neighborhood of a then this is possible

Example 5.1.3. *Suppose $f(x,y,z) = e^{xyz}$. Find a quadratic approximation to f near $(0,1,2)$. Observe:*

$$f_x = yze^{xyz} \qquad f_y = xze^{xyz} \qquad f_z = xye^{xyz}$$

$$f_{xx} = (yz)^2e^{xyz} \qquad f_{yy} = (xz)^2e^{xyz} \qquad f_{zz} = (xy)^2e^{xyz}$$

$$f_{xy} = ze^{xyz} + xyz^2e^{xyz} \qquad f_{yz} = xe^{xyz} + x^2yze^{xyz} \qquad f_{xz} = ye^{xyz} + xy^2ze^{xyz}$$

Evaluating at $x = 0, y = 1$ and $z = 2$,

$$f_x(0,1,2) = 2 \qquad f_y(0,1,2) = 0 \qquad f_z(0,1,2) = 0$$

$$f_{xx}(0,1,2) = 4 \qquad f_{yy}(0,1,2) = 0 \qquad f_{zz}(0,1,2) = 0$$

$$f_{xy}(0,1,2) = 2 \qquad f_{yz}(0,1,2) = 0 \qquad f_{xz}(0,1,2) = 1$$

Hence, as $f(0,1,2) = e^0 = 1$ we find

$$f(x,y,z) = 1 + 2x + 2x^2 + 2x(y-1) + 2x(z-2) + \cdots$$

Another way to calculate this expansion is to make use of the adding zero trick,

$$f(x,y,z) = e^{x(y-1+1)(z-2+2)} = 1 + x(y-1+1)(z-2+2) + \frac{1}{2}\big[x(y-1+1)(z-2+2)\big]^2 + \cdots$$

Keeping only terms with two or less of x, $(y-1)$ and $(z-2)$ variables,

$$f(x,y,z) = 1 + 2x + x(y-1)(2) + x(1)(z-2) + \frac{1}{2}x^2(1)^2(2)^2 + \cdots$$

Which simplifies once more to $f(x,y,z) = 1 + 2x + 2x(y-1) + x(z-2) + 2x^2 + \cdots$.

5.2 a brief introduction to the theory of quadratic forms

Definition 5.2.1.

Generally, a **quadratic form** Q is a function $Q : \mathbb{R}^n \to \mathbb{R}$ whose formula can be written $Q(\vec{x}) = \vec{x}^T A\vec{x}$ for all $\vec{x} \in \mathbb{R}^n$ where $A \in \mathbb{R}^{\ n\times n}$ such that $A^T = A$. In particular, if $\vec{x} = (x,y)$ and $A = \left[\begin{array}{cc} a & b \\ b & c \end{array}\right]$ then

$$Q(\vec{x}) = \vec{x}^T A\vec{x} = ax^2 + bxy + byx + cy^2 = ax^2 + 2bxy + y^2.$$

The $n = 3$ case is similar,denote $A = [A_{ij}]$ and $\vec{x} = (x,y,z)$ so that

$$Q(\vec{x}) = \vec{x}^T A\vec{x} = A_{11}x^2 + 2A_{12}xy + 2A_{13}xz + A_{22}y^2 + 2A_{23}yz + A_{33}z^2.$$

Generally, if $[A_{ij}] \in \mathbb{R}^{\ n\times n}$ and $\vec{x} = [x_i]^T$ then the associated quadratic form is

$$Q(\vec{x}) = \vec{x}^T A\vec{x} = \sum_{i,j} A_{ij}x_ix_j = \sum_{i=1}^{n} A_{ii}x_i^2 + \sum_{i<j} 2A_{ij}x_ix_j.$$

In case you wondering, yes you could write a given quadratic form with a different matrix which is not symmetric, but we will find it convenient to insist that our matrix is symmetric since that choice is always possible for a given quadratic form.

It is at times useful to use the dot-product to express a given quadratic form:

$$\vec{x}^T A\vec{x} = \vec{x}\cdot(A\vec{x}) = (A\vec{x})\cdot\vec{x} = \vec{x}^T A^T\vec{x}$$

Some texts actually use the middle equality above to define a symmetric matrix.

Example 5.2.2.

$$2x^2+2xy+2y^2 = \begin{bmatrix} x & y \end{bmatrix}\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$$

Example 5.2.3.

$$2x^2+2xy+3xz-2y^2-z^2 = \begin{bmatrix} x & y & z \end{bmatrix}\begin{bmatrix} 2 & 1 & 3/2 \\ 1 & -2 & 0 \\ 3/2 & 0 & -1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

Proposition 5.2.4.

The values of a quadratic form on $\mathbb{R}^n - \{0\}$ is completely determined by it's values on the $(n-1)$-sphere $S_{n-1} = \{\vec{x}\in\mathbb{R}^n \mid ||\vec{x}|| = 1\}$. In particular, $Q(\vec{x}) = ||\vec{x}||^2 Q(\hat{x})$ where $\hat{x} = \frac{1}{||\vec{x}||}\vec{x}$.

Proof: Let $Q(\vec{x}) = \vec{x}^T A\vec{x}$. Notice that we can write any nonzero vector as the product of its magnitude $||x||$ and its direction $\hat{x} = \frac{1}{||\vec{x}||}\vec{x}$,

$$Q(\vec{x}) = Q(||\vec{x}||\hat{x}) = (||\vec{x}||\hat{x})^T A||\vec{x}||\hat{x} = ||\vec{x}||^2\hat{x}^T A\hat{x} = ||x||^2 Q(\hat{x}).$$

Therefore $Q(\vec{x})$ is simply proportional to $Q(\hat{x})$ with proportionality constant $||\vec{x}||^2$. □

The proposition above is very interesting. It says that if we know how Q works on unit-vectors then we can extrapolate its action on the remainder of $\mathbb{R}^n$. If $f : S \to \mathbb{R}$ then we could say $f(S) > 0$ iff $f(s) > 0$ for all $s \in S$. Likewise, $f(S) < 0$ iff $f(s) < 0$ for all $s \in S$. The proposition below follows from the proposition above since $||\vec{x}||^2$ ranges over all nonzero positive real numbers in the equations above.

Proposition 5.2.5.

If Q is a quadratic form on $\mathbb{R}^n$ and we denote $\mathbb{R}^n_* = \mathbb{R}^n - \{0\}$

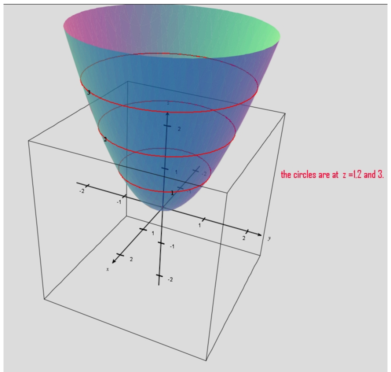
1.(negative definite) $Q(\mathbb{R}^n_*) < 0$ iff $Q(S_{n-1}) < 0$

2.(positive definite) $Q(\mathbb{R}^n_*) > 0$ iff $Q(S_{n-1}) > 0$

3.(non-definite) $Q(\mathbb{R}^n_*) = \mathbb{R} - \{0\}$ iff $Q(S_{n-1})$ has both positive and negative values.

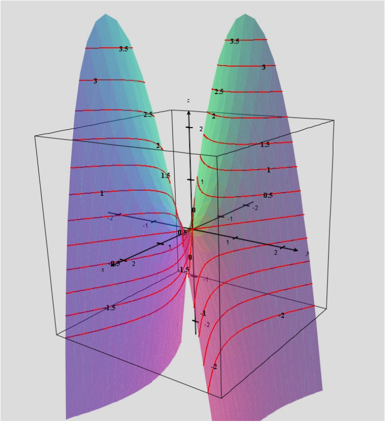
Before I get too carried away with the theory let's look at a couple examples.

Example 5.2.6. *Consider the quadric form $Q(x,y) = x^2 + y^2$. You can check for yourself that $z = Q(x,y)$ is a cone and Q has positive outputs for all inputs except $(0,0)$. Notice that $Q(v) = ||v||^2$ so it is clear that $Q(S_1) = 1$. We find agreement with the preceding proposition. Next, think about the application of $Q(x,y)$ to level curves; $x^2 + y^2 = k$ is simply a circle of radius $\sqrt{k}$ or just the origin. Here's a graph of $z = Q(x,y)$:*



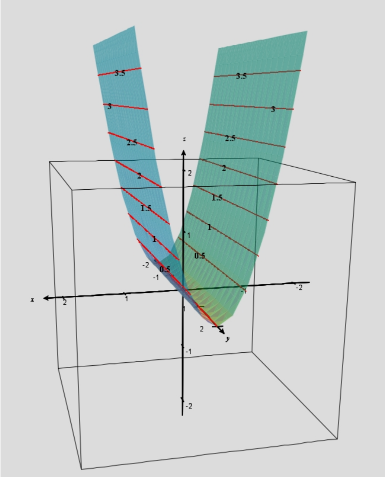
Notice that $Q(0,0) = 0$ is the absolute minimum for Q. Finally, let's take a moment to write $Q(x,y) = [x,y]\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$ in this case the matrix is diagonal and we note that the e-values are $\lambda_1 = \lambda_2 = 1$.

Example 5.2.7. *Consider the quadric form $Q(x,y) = x^2 - 2y^2$. You can check for yourself that $z = Q(x,y)$ is a hyperboloid and Q has non-definite outputs since sometimes the x^2 term dominates whereas other points have $-2y^2$ as the dominent term. Notice that $Q(1,0) = 1$ whereas $Q(0,1) = -2$ hence we find $Q(S_1)$ contains both positive and negative values and consequently we find agreement with the preceding proposition. Next, think about the application of $Q(x,y)$ to level curves; $x^2 - 2y^2 = k$ yields either hyperbolas which open vertically $(k > 0)$ or horizontally $(k < 0)$ or a pair of lines $y = \pm\frac{x}{2}$ in the $k = 0$ case. Here's a graph of $z = Q(x,y)$:*



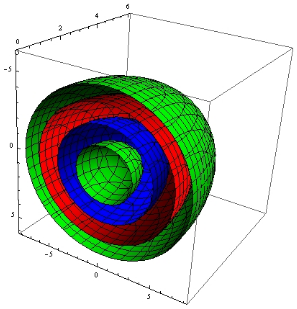
The origin is a **saddle point**. *Finally, let's take a moment to write $Q(x,y) = [x,y]\begin{bmatrix} 1 & 0 \\ 0 & -2 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$ in this case the matrix is diagonal and we note that the e-values are $\lambda_1 = 1$ and $\lambda_2 = -2$.*

Example 5.2.8. *Consider the quadric form $Q(x, y) = 3x^2$. You can check for yourself that $z = Q(x, y)$ is parabola-shaped trough along the y-axis. In this case Q has positive outputs for all inputs except $(0, y)$, we would call this form* **positive semi-definite**. *A short calculation reveals that $Q(S_1) = [0, 3]$ thus we again find agreement with the preceding proposition (case 3). Next, think about the application of $Q(x, y)$ to level curves; $3x^2 = k$ is a pair of vertical lines: $x = \pm\sqrt{k/3}$ or just the y-axis. Here's a graph of $z = Q(x, y)$:*



Finally, let's take a moment to write $Q(x, y) = [x, y]\begin{bmatrix} 3 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$ in this case the matrix is diagonal and we note that the e-values are $\lambda_1 = 3$ and $\lambda_2 = 0$.

Example 5.2.9. *Consider the quadric form $Q(x, y, z) = x^2 + 2y^2 + 3z^2$. Think about the application of $Q(x, y, z)$ to level surfaces; $x^2 + 2y^2 + 3z^2 = k$ is an ellipsoid. I can't graph a function of three variables, however, we can look at level surfaces of the function. I use Mathematica to plot several below:*



Finally, let's take a moment to write $Q(x, y, z) = [x, y, z]\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix}$ in this case the matrix is diagonal and we note that the e-values are $\lambda_1 = 1$ and $\lambda_2 = 2$ and $\lambda_3 = 3$.

5.2.1 diagonalizing forms via eigenvectors

The examples given thus far are the simplest cases. We don't really need linear algebra to understand them. In contrast, e-vectors and e-values will prove a useful tool to unravel the later examples[3]

[3] this is the one place in this course where we need eigenvalues and eigenvector calculations, I include these to illustrate the structure of quadratic forms in general, however, as linear algebra is not a prerequisite you may find some things in this section mysterious. The homework and study guide will elaborate on what is required this semester

Definition 5.2.10.

Let $A \in \mathbb{R}^{n\times n}$. If $v \in \mathbb{R}^{n\times 1}$ is **nonzero** and $Av = \lambda v$ for some $\lambda \in \mathbb{C}$ then we say v is an **eigenvector** with **eigenvalue** λ of the matrix A.

Proposition 5.2.11.

Let $A \in \mathbb{R}^{n\times n}$ then λ is an eigenvalue of A iff $det(A-\lambda I) = 0$. We say $P(\lambda) = det(A-\lambda I)$ the **characteristic polynomial** and $det(A-\lambda I) = 0$ is the **characteristic equation**.

Proof: Suppose λ is an eigenvalue of A then there exists a nonzero vector v such that $Av = \lambda v$ which is equivalent to $Av - \lambda v = 0$ which is precisely $(A-\lambda I)v = 0$. Notice that $(A-\lambda I)0 = 0$ thus the matrix $(A-\lambda I)$ is singular as the equation $(A-\lambda I)x = 0$ has more than one solution. Consequently $det(A-\lambda I) = 0$.

Conversely, suppose $det(A-\lambda I) = 0$. It follows that $(A-\lambda I)$ is singular. Clearly the system $(A-\lambda I)x = 0$ is consistent as $x = 0$ is a solution hence we know there are infinitely many solutions. In particular there exists at least one vector $v \neq 0$ such that $(A-\lambda I)v = 0$ which means the vector v satisfies $Av = \lambda v$. Thus v is an eigenvector with eigenvalue λ for A. □

Remark 5.2.12.

I found a pretty derivation of the eigenvector condition from the method of Lagrange multipliers. I shared in the Lecture 10 part 1. It's likely I cover that argument again in Lecture this year, my apologies it has not made it to these notes at this time.

Example 5.2.13. *Let* $A = \begin{bmatrix} 3 & 1 \\ 3 & 1 \end{bmatrix}$ *find the e-values and e-vectors of* A.

$$det(A-\lambda I) = det\begin{bmatrix} 3-\lambda & 1 \\ 3 & 1-\lambda \end{bmatrix} = (3-\lambda)(1-\lambda) - 3 = \lambda^2 - 4\lambda = \lambda(\lambda-4) = 0$$

We find $\lambda_1 = 0$ *and* $\lambda_2 = 4$. *Now find the e-vector with e-value* $\lambda_1 = 0$, *let* $u_1 = [u,v]^T$ *denote the e-vector we wish to find. Calculate,*

$$(A-0I)u_1 = \begin{bmatrix} 3 & 1 \\ 3 & 1 \end{bmatrix}\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} 3u+v \\ 3u+v \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Obviously the equations above are redundant and we have infinitely many solutions of the form $3u + v = 0$ *which means* $v = -3u$ *so we can write,* $u_1 = \begin{bmatrix} u \\ -3u \end{bmatrix} = u\begin{bmatrix} 1 \\ -3 \end{bmatrix}$. *In applications we often make a choice to select a particular e-vector. Most modern graphing calculators can calculate e-vectors. It is customary for the e-vectors to be chosen to have length one. That is a useful choice for certain applications as we will later discuss. If you use a calculator it would likely give* $u_1 = \frac{1}{\sqrt{10}}\begin{bmatrix} 1 \\ -3 \end{bmatrix}$ *although the* $\sqrt{10}$ *would likely be approximated unless your calculator is smart.*

Continuing we wish to find eigenvectors $u_2 = [u,v]^T$ *such that* $(A-4I)u_2 = 0$. *Notice that* u, v *are disposable variables in this context, I do not mean to connect the formulas from the* $\lambda = 0$ *case with the case considered now.*

$$(A-4I)u_1 = \begin{bmatrix} -1 & 1 \\ 3 & -3 \end{bmatrix}\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} -u+v \\ 3u-3v \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Again the equations are redundant and we have infinitely many solutions of the form $v = u$. Hence, $u_2 = \begin{bmatrix} u \\ u \end{bmatrix} = u \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is an eigenvector for any $u \in \mathbb{R}$ such that $u \neq 0$.

Theorem 5.2.14.

A matrix $A \in \mathbb{R}^{\,n\times n}$ is symmetric iff there exists an orthonormal eigenbasis for A.

There is a geometric proof of this theorem in Edwards[4] (see Theorem 8.6 pgs 146-147) . I prove half of this theorem in my linear algebra notes by a non-geometric argument (full proof is in Appendix C of Insel,Spence and Friedberg). It might be very interesting to understand the connection between the geometric verse algebraic arguments. We'll content ourselves with an example here:

Example 5.2.15. *Let $A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 2 & 1 \end{bmatrix}$. Observe that $det(A - \lambda I) = -\lambda(\lambda+1)(\lambda-3)$ thus $\lambda_1 = 0, \lambda_2 = -1, \lambda_3 = 3$. We can calculate orthonormal e-vectors of $v_1 = [1,0,0]^T$, $v_2 = \frac{1}{\sqrt{2}}[0,1,-1]^T$ and $v_3 = \frac{1}{\sqrt{2}}[0,1,1]^T$. I invite the reader to check the validity of the following equation:*

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

Its really neat that to find the inverse of a matrix of orthonormal e-vectors we need only take the transpose; note

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Proposition 5.2.16.

If Q is a quadratic form on $\mathbb{R}^n$ with matrix A and e-values $\lambda_1, \lambda_2, \dots, \lambda_n$ with orthonormal e-vectors $v_1, v_2, \dots, v_n$ then

$$Q(v_i) = {\lambda_i}^2$$

for $i = 1, 2, \dots, n$. Moreover, if $P = [v_1|v_2|\cdots|v_n]$ then

$$Q(\vec{x}) = (P^T\vec{x})^T P^T A P P^T \vec{x} = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2$$

where we defined $\vec{y} = P^T\vec{x}$.

Let me restate the proposition above in simple terms: we can transform a given quadratic form to a diagonal form by finding orthonormalized e-vectors and performing the appropriate coordinate transformation. Since P is formed from orthonormal e-vectors we know that P will be either a rotation or reflection. This proposition says we can remove "cross-terms" by transforming the quadratic forms with an appropriate rotation.

[4] think about it, there is a 1-1 correspondance between symmetric matrices and quadratic forms

Example 5.2.17. *Consider the quadric form $Q(x,y) = 2x^2+2xy+2y^2$. It's not immediately obvious (to me) what the level curves $Q(x,y) = k$ look like. We'll make use of the preceding proposition to understand those graphs. Notice $Q(x,y) = [x,y]\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$. Denote the matrix of the form by A and calculate the e-values/vectors:*

$$det(A-\lambda I) = det\begin{bmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{bmatrix} = (\lambda-2)^2-1 = \lambda^2-4\lambda+3 = (\lambda-1)(\lambda-3) = 0$$

Therefore, the e-values are $\lambda_1 = 1$ and $\lambda_2 = 3$.

$$(A-I)\vec{u}_1 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \Rightarrow \quad \vec{u}_1 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

I just solved $u+v=0$ to give $v=-u$ choose $u=1$ then normalize to get the vector above. Next,

$$(A-3I)\vec{u}_2 = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \Rightarrow \quad \vec{u}_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

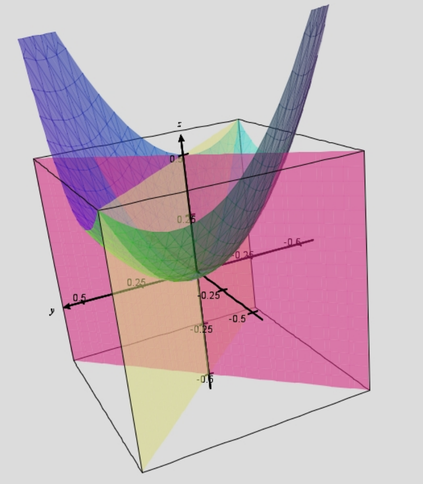
I just solved $u-v=0$ to give $v=u$ choose $u=1$ then normalize to get the vector above. Let $P = [\vec{u}_1|\vec{u}_2]$ and introduce new coordinates $\vec{y} = [\bar{x},\bar{y}]^T$ defined by $\vec{y} = P^T\vec{x}$. Note these can be inverted by multiplication by P to give $\vec{x} = P\vec{y}$. Observe that

$$P = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \quad \Rightarrow \quad \begin{array}{ll} x & = \frac{1}{2}(\bar{x}+\bar{y}) \\ y & = \frac{1}{2}(-\bar{x}+\bar{y}) \end{array} \quad or \quad \begin{array}{ll} \bar{x} & = \frac{1}{2}(x-y) \\ \bar{y} & = \frac{1}{2}(x+y) \end{array}$$

The proposition preceding this example shows that substitution of the formulas above into Q yield[5]:

$$\tilde{Q}(\bar{x},\bar{y}) = \bar{x}^2+3\bar{y}^2$$

It is clear that in the barred coordinate system the level curve $Q(x,y) = k$ is an ellipse. If we draw the barred coordinate system superposed over the xy-coordinate system then you'll see that the graph of $Q(x,y) = 2x^2+2xy+2y^2 = k$ is an ellipse rotated by 45 degrees. Or, if you like, we can plot $z = Q(x,y)$:



[5] technically $\tilde{Q}(\bar{x},\bar{y})$ is $Q(x(\bar{x},\bar{y}),y(\bar{x},\bar{y}))$

Example 5.2.18. *Consider the quadric form $Q(x,y) = x^2+2xy+y^2$. It's not immediately obvious (to me) what the level curves $Q(x,y)=k$ look like. We'll make use of the preceding proposition to understand those graphs. Notice $Q(x,y) = [x,y]\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$. Denote the matrix of the form by A and calculate the e-values/vectors:*

$$det(A-\lambda I) = det\begin{bmatrix} 1-\lambda & 1 \\ 1 & 1-\lambda \end{bmatrix} = (\lambda-1)^2 - 1 = \lambda^2 - 2\lambda = \lambda(\lambda-2) = 0$$

Therefore, the e-values are $\lambda_1 = 0$ and $\lambda_2 = 2$.

$$(A-0)\vec{u}_1 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \Rightarrow \quad \vec{u}_1 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

I just solved $u+v=0$ to give $v=-u$ choose $u=1$ then normalize to get the vector above. Next,

$$(A-2I)\vec{u}_2 = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \Rightarrow \quad \vec{u}_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

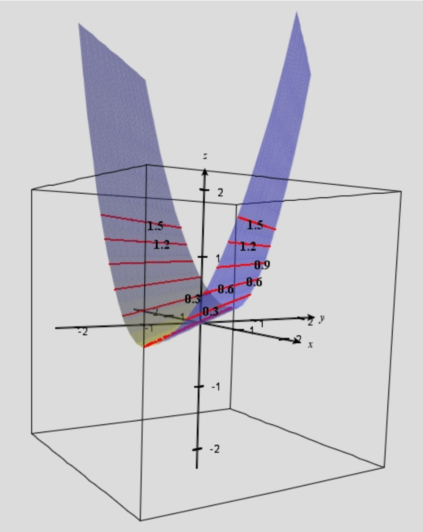
I just solved $u-v=0$ to give $v=u$ choose $u=1$ then normalize to get the vector above. Let $P = [\vec{u}_1|\vec{u}_2]$ and introduce new coordinates $\vec{y} = [\bar{x},\bar{y}]^T$ defined by $\vec{y} = P^T\vec{x}$. Note these can be inverted by multiplication by P to give $\vec{x} = P\vec{y}$. Observe that

$$P = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \quad \Rightarrow \quad \begin{array}{ll} x & = \frac{1}{2}(\bar{x}+\bar{y}) \\ y & = \frac{1}{2}(-\bar{x}+\bar{y}) \end{array} \quad or \quad \begin{array}{ll} \bar{x} & = \frac{1}{2}(x-y) \\ \bar{y} & = \frac{1}{2}(x+y) \end{array}$$

The proposition preceding this example shows that substitution of the formulas above into Q yield:

$$\tilde{Q}(\bar{x},\bar{y}) = 2\bar{y}^2$$

It is clear that in the barred coordinate system the level curve $Q(x,y)=k$ is a pair of paralell lines. If we draw the barred coordinate system superposed over the xy-coordinate system then you'll see that the graph of $Q(x,y) = x^2+2xy+y^2 = k$ is a line with slope -1. Indeed, with a little algebraic insight we could have anticipated this result since $Q(x,y) = (x+y)^2$ so $Q(x,y)=k$ implies $x+y=\sqrt{k}$ thus $y=\sqrt{k}-x$. Here's a plot which again verifies what we've already found:



Example 5.2.19. *Consider the quadric form $Q(x,y) = 4xy$. It's not immediately obvious (to me) what the level curves $Q(x,y) = k$ look like. We'll make use of the preceding proposition to understand those graphs. Notice $Q(x,y) = [x,y]\left[\begin{array}{cc} 0 & 2 \\ 0 & 2 \end{array}\right]\left[\begin{array}{c} x \\ y \end{array}\right]$. Denote the matrix of the form by A and calculate the e-values/vectors:*

$$det(A-\lambda I) = det\left[\begin{array}{cc} -\lambda & 2 \\ 2 & -\lambda \end{array}\right] = \lambda^2 - 4 = (\lambda+2)(\lambda-2) = 0$$

Therefore, the e-values are $\lambda_1 = -2$ and $\lambda_2 = 2$.

$$(A+2I)\vec{u}_1 = \left[\begin{array}{cc} 2 & 2 \\ 2 & 2 \end{array}\right]\left[\begin{array}{c} u \\ v \end{array}\right] = \left[\begin{array}{c} 0 \\ 0 \end{array}\right] \quad \Rightarrow \quad \vec{u}_1 = \frac{1}{\sqrt{2}}\left[\begin{array}{c} 1 \\ -1 \end{array}\right]$$

I just solved $u+v=0$ to give $v=-u$ choose $u=1$ then normalize to get the vector above. Next,

$$(A-2I)\vec{u}_2 = \left[\begin{array}{cc} -2 & 2 \\ 2 & -2 \end{array}\right]\left[\begin{array}{c} u \\ v \end{array}\right] = \left[\begin{array}{c} 0 \\ 0 \end{array}\right] \quad \Rightarrow \quad \vec{u}_2 = \frac{1}{\sqrt{2}}\left[\begin{array}{c} 1 \\ 1 \end{array}\right]$$

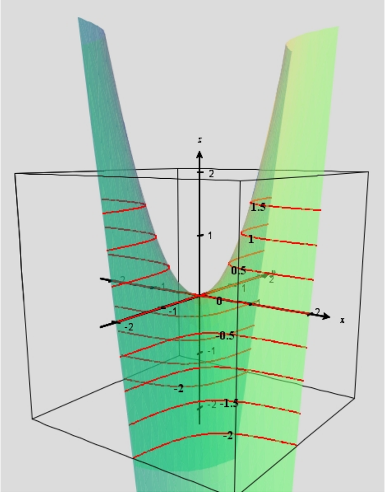
I just solved $u-v=0$ to give $v=u$ choose $u=1$ then normalize to get the vector above. Let $P = [\vec{u}_1|\vec{u}_2]$ and introduce new coordinates $\vec{y} = [\bar{x},\bar{y}]^T$ defined by $\vec{y} = P^T\vec{x}$. Note these can be inverted by multiplication by P to give $\vec{x} = P\vec{y}$. Observe that

$$P = \frac{1}{2}\left[\begin{array}{cc} 1 & 1 \\ -1 & 1 \end{array}\right] \quad \Rightarrow \quad \begin{array}{rl} x & = \frac{1}{2}(\bar{x}+\bar{y}) \\ y & = \frac{1}{2}(-\bar{x}+\bar{y}) \end{array} \quad or \quad \begin{array}{rl} \bar{x} & = \frac{1}{2}(x-y) \\ \bar{y} & = \frac{1}{2}(x+y) \end{array}$$

The proposition preceding this example shows that substitution of the formulas above into Q yield:

$$\tilde{Q}(\bar{x},\bar{y}) = -2\bar{x}^2 + 2\bar{y}^2$$

It is clear that in the barred coordinate system the level curve $Q(x,y) = k$ is a hyperbola. If we draw the barred coordinate system superposed over the xy-coordinate system then you'll see that the graph of $Q(x,y) = 4xy = k$ is a hyperbola rotated by 45 degrees. The graph $z = 4xy$ is thus a hyperbolic paraboloid:



The fascinating thing about the mathematics here is that if you don't want to graph $z = Q(x,y)$, but you do want to know the general shape then you can determine which type of quadraic surface you're dealing with by simply calculating the eigenvalues of the form.

Remark 5.2.20.

I made the preceding triple of examples all involved the same rotation. This is purely for my lecturing convenience. In practice the rotation could be by all sorts of angles. In addition, you might notice that a different ordering of the e-values would result in a redefinition of the barred coordinates. [6]

We ought to do at least one 3-dimensional example.

Example 5.2.21. *Consider the quadric form Q defined below:*

$$Q(x,y,z) = [x,y,z]\begin{bmatrix} 6 & -2 & 0 \\ -2 & 6 & 0 \\ 0 & 0 & 5 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

Denote the matrix of the form by A and calculate the e-values/vectors:

$$\begin{aligned} det(A-\lambda I) &= det\begin{bmatrix} 6-\lambda & -2 & 0 \\ -2 & 6-\lambda & 0 \\ 0 & 0 & 5-\lambda \end{bmatrix} \\ &= [(\lambda-6)^2-4](5-\lambda) \\ &= (5-\lambda)[\lambda^2-12\lambda+32](5-\lambda) \\ &= (\lambda-4)(\lambda-8)(5-\lambda) \end{aligned}$$

Therefore, the e-values are $\lambda_1 = 4, \ \ \lambda_2 = 8$ and $\lambda_3 = 5$. After some calculation we find the following orthonormal e-vectors for A:

$$\vec{u}_1 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \qquad \vec{u}_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \qquad \vec{u}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

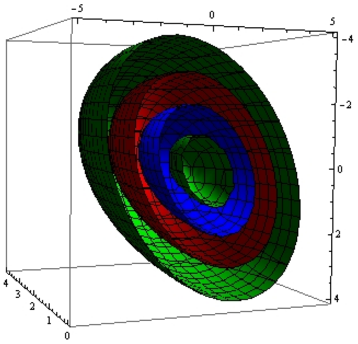
Let $P = [\vec{u}_1|\vec{u}_2|\vec{u}_3]$ and introduce new coordinates $\vec{y} = [\bar{x},\bar{y},\bar{z}]^T$ defined by $\vec{y} = P^T\vec{x}$. Note these can be inverted by multiplication by P to give $\vec{x} = P\vec{y}$. Observe that

$$P = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & \sqrt{2} \end{bmatrix} \ \Rightarrow \ \begin{array}{lcl} x &=& \frac{1}{2}(\bar{x}+\bar{y}) \\ y &=& \frac{1}{2}(-\bar{x}+\bar{y}) \\ z &=& \bar{z} \end{array} \quad or \quad \begin{array}{lcl} \bar{x} &=& \frac{1}{2}(x-y) \\ \bar{y} &=& \frac{1}{2}(x+y) \\ \bar{z} &=& z \end{array}$$

The proposition preceding this example shows that substitution of the formulas above into Q yield:

$$\tilde{Q}(\bar{x},\bar{y},\bar{z}) = 4\bar{x}^2 + 8\bar{y}^2 + 5\bar{z}^2$$

It is clear that in the barred coordinate system the level surface $Q(x,y,z) = k$ is an ellipsoid. If we draw the barred coordinate system superposed over the xyz-coordinate system then you'll see that the graph of $Q(x,y,z) = k$ is an ellipsoid rotated by 45 degrees around the $z-axis$. Plotted below are a few representative ellipsoids:



In summary, the behaviour of a quadratic form $Q(x) = x^TAx$ is governed by it's set of eigenvalues[7] $\{\lambda_1, \lambda_2, \dots, \lambda_k\}$. Moreover, the form can be written as $Q(y) = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_k y_k^2$ by choosing the coordinate system which is built from the orthonormal eigenbasis of $col(A)$. In this coordinate system the shape of the level-sets of Q becomes manifest from the signs of the e-values.)

Remark 5.2.22.

If you would like to read more about conic sections or quadric surfaces and their connection to e-values/vectors I reccommend sections 9.6 and 9.7 of Anton's linear algebra text. I have yet to add examples on how to include translations in the analysis. It's not much more trouble but I decided it would just be an unecessary complication this semester. Also, section 7.1,7.2 and 7.3 in Lay's linear algebra text show a bit more about how to use this math to solve concrete applied problems. You might also take a look in Gilbert Strang's linear algebra text, his discussion of tests for positive-definite matrices is much more complete than I will give here.

5.3 second derivative test in many-variables

There is a connection between the shape of level curves $Q(x_1, x_2, \dots, x_n) = k$ and the graph $x_{n+1} = f(x_1, x_2, \dots, x_n)$ of f. I'll discuss $n = 2$ but these comments equally well apply to $w = f(x, y, z)$ or higher dimensional examples. Consider a critical point (a, b) for $f(x, y)$ then the Taylor expansion about (a, b) has the form

$$f(a + h, b + k) = f(a, b) + Q(h, k)$$

where $Q(h, k) = \frac{1}{2}h^2 f_{xx}(a, b) + hkf_{xy}(a, b) + \frac{1}{2}h^2 f_{yy}(a, b) = [h, k][Q](h, k)$. Since $[Q]^T = [Q]$ we can find orthonormal e-vectors $\vec{u}_1, \vec{u}_2$ for $[Q]$ with e-values λ_1 and λ_2 respective. Using $U = [\vec{u}_1|\vec{u}_2]$ we can introduce rotated coordinates $(\bar{h}, \bar{k}) = U(h, k)$. These will give

$$Q(\bar{h}, \bar{k}) = \lambda_1 \bar{h}^2 + \lambda_2 \bar{k}^2$$

Clearly if $\lambda_1 > 0$ and $\lambda_2 > 0$ then $f(a, b)$ yields the local minimum whereas if $\lambda_1 < 0$ and $\lambda_2 < 0$ then $f(a, b)$ yields the local maximum. Edwards discusses these matters on pgs. 148-153. In short, supposing $f \approx f(p) + Q$, if all the e-values of Q are positive then f has a local minimum of $f(p)$ at p whereas if all the e-values of Q are negative then f reaches a local maximum of $f(p)$ at p. Otherwise Q has both positive and negative e-values and we say Q is non-definite and the function has a saddle point. If all the e-values of Q are positive then Q is said to be **positive-definite** whereas if all the e-values of Q are negative then Q is said to be **negative-definite**. Edwards gives a few nice tests for ascertaining if a matrix is positive definite without explicit computation of e-values. Finally, if one of the e-values is zero then the graph will be like a trough.

[7] this set is called the spectrum of the matrix

Example 5.3.1. Suppose $f(x,y) = exp(-x^2 - y^2 + 2y - 1)$ expand f about the point $(0,1)$:

$$f(x,y) = exp(-x^2)exp(-y^2 + 2y - 1) = exp(-x^2)exp(-(y-1)^2)$$

expanding,

$$f(x,y) = (1 - x^2 + \cdots)(1 - (y-1)^2 + \cdots) = 1 - x^2 - (y-1)^2 + \cdots$$

Recenter about the point $(0,1)$ by setting $x = h$ and $y = 1 + k$ so

$$f(h, 1+k) = 1 - h^2 - k^2 + \cdots$$

If (h,k) is near $(0,0)$ then the dominant terms are simply those we've written above hence the graph is like that of a quadraic surface with a pair of negative e-values. It follows that $f(0,1)$ is a local maximum. In fact, it happens to be a global maximum for this function.

Example 5.3.2. Suppose $f(x,y) = 4-(x-1)^2+(y-2)^2+Aexp(-(x-1)^2-(y-2)^2)+2B(x-1)(y-2)$ for some constants A, B. Analyze what values for A, B will make $(1,2)$ a local maximum, minimum or neither. Expanding about $(1,2)$ we set $x = 1 + h$ and $y = 2 + k$ in order to see clearly the local behaviour of f at $(1,2)$,

$$\begin{aligned} f(1+h, 2+k) &= 4 - h^2 - k^2 + Aexp(-h^2 - k^2) + 2Bhk \\ &= 4 - h^2 - k^2 + A(1 - h^2 - k^2) + 2Bhk \cdots \\ &= 4 + A - (A+1)h^2 + 2Bhk - (A+1)k^2 + \cdots \end{aligned}$$

There is no nonzero linear term in the expansion at $(1,2)$ which indicates that $f(1,2) = 4 + A$ may be a local extremum. In this case the quadratic terms are nontrivial which means the graph of this function is well-approximated by a quadraic surface near $(1,2)$. The quadratic form $Q(h,k) = -(A+1)h^2 + 2Bhk - (A+1)k^2$ has matrix

$$[Q] = \begin{bmatrix} -(A+1) & B \\ B & -(A+1)^2 \end{bmatrix}.$$

The characteristic equation for Q is

$$det([Q] - \lambda I) = det \begin{bmatrix} -(A+1) - \lambda & B \\ B & -(A+1)^2 - \lambda \end{bmatrix} = (\lambda + A + 1)^2 - B^2 = 0$$

We find solutions $\lambda_1 = -A - 1 + B$ and $\lambda_2 = -A - 1 - B$. The possibilities break down as follows:

1. if $\lambda_1, \lambda_2 > 0$ then $f(1,2)$ is local minimum.
2. if $\lambda_1, \lambda_2 < 0$ then $f(1,2)$ is local maximum.
3. if just one of λ_1, λ_2 is zero then f is constant along one direction and min/max along another so technically it is a local extremum.
4. if $\lambda_1 \lambda_2 < 0$ then $f(1,2)$ is not a local etremum, however it is a saddle point.

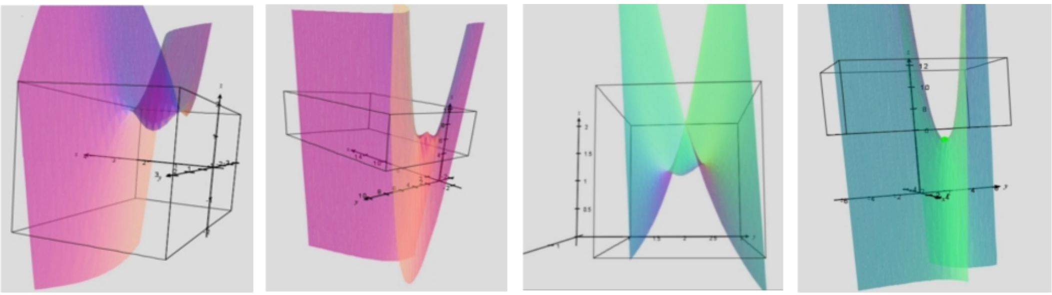
In particular, the following choices for A, B will match the choices above

1. Let $A = -3$ and $B = 1$ so $\lambda_1 = 3$ and $\lambda_2 = 1$;
2. Let $A = 3$ and $B = 1$ so $\lambda_1 = -3$ and $\lambda_2 = -5$

3. *Let $A = -3$ and $B = -2$ so $\lambda_1 = 0$ and $\lambda_2 = 4$*

4. *Let $A = 1$ and $B = 3$ so $\lambda_1 = 1$ and $\lambda_2 = -5$*

Here are the graphs of the cases above, note the analysis for case 3 is more subtle for Taylor approximations as opposed to simple quadraic surfaces. In this example, case 3 was also a local minimum. In contrast, in Example 5.2.18 the graph was like a trough. The behaviour of f away from the critical point includes higher order terms whose influence turns the trough into a local minimum.



Example 5.3.3. *Suppose $f(x,y) = \sin(x)\cos(y)$ to find the Taylor series centered at $(0,0)$ we can simply multiply the one-dimensional result $\sin(x) = x - \frac{1}{3!}x^3 + \frac{1}{5!}x^5 + \cdots$ and $\cos(y) = 1 - \frac{1}{2!}y^2 + \frac{1}{4!}y^4 + \cdots$ as follows:*

$$\begin{aligned} f(x,y) &= (x - \tfrac{1}{3!}x^3 + \tfrac{1}{5!}x^5 + \cdots)(1 - \tfrac{1}{2!}y^2 + \tfrac{1}{4!}y^4 + \cdots) \\ &= x - \tfrac{1}{2}xy^2 + \tfrac{1}{24}xy^4 - \tfrac{1}{6}x^3 - \tfrac{1}{12}x^3y^2 + \cdots \\ &= x + \cdots \end{aligned}$$

The origin $(0,0)$ is a critical point since $f_x(0,0) = 0$ and $f_y(0,0) = 0$, however, this particular critical point escapes the analysis via the quadratic form term since $Q = 0$ in the Taylor series for this function at $(0,0)$. This is analogous to the inconclusive case of the 2nd derivative test in calculus III.

Example 5.3.4. *Suppose $f(x,y,z) = xyz$. Calculate the multivariate Taylor expansion about the point $(1,2,3)$. I'll actually calculate this one via differentiation, I have used tricks and/or calculus II results to shortcut any differentiation in the previous examples. Calculate first derivatives*

$$f_x = yz \qquad f_y = xz \qquad f_z = xy,$$

and second derivatives,

$$\begin{array}{ccc} f_{xx} = 0 & f_{xy} = z & f_{xz} = y \\ f_{yx} = z & f_{yy} = 0 & f_{yz} = x \\ f_{zx} = y & f_{zy} = x & f_{zz} = 0, \end{array}$$

and the nonzero third derivatives,

$$f_{xyz} = f_{yzx} = f_{zxy} = f_{zyx} = f_{yxz} = f_{xzy} = 1.$$

It follows,

$$\begin{aligned} &f(a+h, b+k, c+l) = \\ &\quad = f(a,b,c) \ + \ f_x(a,b,c)h \ + \ f_y(a,b,c)k \ + \ f_z(a,b,c)l \ + \\ &\qquad \tfrac{1}{2}(\ f_{xx}hh + f_{xy}hk + f_{xz}hl + f_{yx}kh + f_{yy}kk + f_{yz}kl + f_{zx}lh + f_{zy}lk + f_{zz}ll\ ) + \cdots \end{aligned}$$

Of course certain terms can be combined since $f_{xy} = f_{yx}$ etc... for smooth functions (we assume smooth in this section, moreover the given function here is clearly smooth). In total,

$$f(1+h, 2+k, 3+l) = 6 + 6h + 3k + 2l + \frac{1}{2}\big(3hk + 2hl + 3kh + kl + 2lh + lk\big) + \frac{1}{3!}(6)hkl$$

Of course, we could also obtain this from simple algebra:

$$f(1+h, 2+k, 3+l) = (1+h)(2+k)(3+l) = 6 + 6h + 3k + l + 3hk + 2hl + kl + hkl.$$

Chapter 6

introduction to variational calculus

6.1 history

The problem of variational calculus is almost as old as modern calculus. Variational calculus seeks to answer questions such as:

Remark 6.1.1.

1. what is the shortest path between two points on a surface ?
2. what is the path of least time for a mass sliding without friction down some path between two given points ?
3. what is the path which minimizes the energy for some physical system ?
4. given two points on the x-axis and a particular area what curve has the longest perimeter and bounds that area between those points and the x-axis?

You'll notice these all involve a variable which is not a real variable or even a vector-valued-variable. Instead, the answers to the questions posed above will be **paths** or **curves** depending on how you wish to frame the problem. In variational calculus the variable is a function and we wish to find extreme values for a **functional**. In short, a functional is an abstract function of functions. A functional takes as an input a function and gives as an output a number. The space from which these functions are taken varies from problem to problem. Often we put additional **contraints** or **conditions** on the **space of admissable solutions**. To read about the full generality of the problem you should look in a text such as Hans Sagan's. Our treatment is introductory in this chapter, my aim is to show you why it is plausible and then to show you how we use variational calculus.

We will see that the problem of finding an extreme value for a functional is equivalent to solving the Euler-Lagrange equations or Euler equations for the functional. Euler predates Lagrange in his discovery of the equations bearing their names. Eulers's initial attack of the problem was to chop the hypothetical solution curve up into a polygonal path. The unknowns in that approach were the coordinates of the vertices in the polygonal path. Then through some ingenious calculations he arrived at the Euler-Lagrange equations. Apparently there were logical flaws in Euler's original treatment. Lagrange later derived the same equations using the viewpoint that the variable was a function and the **variation** was one of shifting by an arbitrary function. The treatment of variational calculus in Edwards is neither Euler nor Lagrange's approach, it is a refined version which takes in the contributions of generations of mathematicians working on the subject and then merges it with careful functional analysis. I'm no expert of the full history, I just give you a rough sketch of what I've gathered from reading a few variational calculus texts.

Physics played a large role in the development of variational calculus. Lagrange was a physicist as well as a mathematician. At the present time, every physicist takes course(s) in *Lagrangian Mechanics*. Moreover, the use of variational calculus is fundamental since Hamilton's principle says that all physics can be derived from the principle of least action. In short this means that nature is lazy. The solutions realized in the physical world are those which minimize the action. The action

$$S[y] = \int L(y, y', t)\, dt$$

is constructed from the Lagrangian $L = T - U$ where T is the kinetic energy and U is the potential energy. In the case of classical mechanics the Euler Lagrange equations are precisely Newton's equations. The Hamiltonian $H = T + U$ is similar to the Lagrangian except that the fundamental variables are taken to be momentum and position in contrast to velocity and position in Lagrangian mechanics.

Hamiltonians and Lagrangians are used to set-up new physical theories. Euler-Lagrange equations are said to give the so-called *classical limit* of modern field theories. The concept of a force is not so useful to quantum theories, instead the concept of energy plays the central role. Moreover, the problem of quantizing and then renormalizing field theory brings in very sophisiticated mathematics. In fact, the math of modern physics is not understood. In this chapter I'll just show you a few famous classical mechanics problems which are beatifully solved by Lagrange's approach. We'll also see how expressing the Lagrangian in non-Cartesian coordinates can give us an easy way to derive forces that arise from geometric contraints.

I am following the typical physics approach to variational calculus. Edwards' last chapter is more natural mathematically but I think the math is a bit much for your first exposure to the subject. The treatment given here is close to that of Arfken and Weber's Mathematical Physics text, however I suspect you can find these calculations in dozens of classical mechanics texts. More or less our approach is that of Lagrange.

6.2 the variational problem

Our goal in what follows here is to maximize or minimize a particular function of functions. Suppose $\mathcal{F}_o$ is a set of functions with some particular property. For now, we may could assume that all the functions in $\mathcal{F}_o$ have graphs that include (x_1, y_1) and (x_2, y_2). Consider a functional $J : \mathcal{F}_o \to \mathcal{F}_o$ which is defined by an integral of some function f which we call the **Lagrangian**,

$$J[y] = \int_{x_1}^{x_2} f(y, y', x)\, dx.$$

We suppose that f is given but y is a variable. Consider that if we are given a function $y^* \in \mathcal{F}_o$ and another function η such that $\eta(x_1) = \eta(x_2) = 0$ then we can reach a whole family of functions indexed by a real variable α as follows (relabel $y^*(x)$ by $y(x, 0)$ so it matches the rest of the family of functions):

$$y(x, \alpha) = y(x, 0) + \alpha\eta(x)$$

Note that $x \mapsto y(x, \alpha)$ gives a function in $\mathcal{F}_o$. We define the **variation** of y to be

$$\boxed{\delta y = \alpha \eta(x)}$$

This means $y(x, \alpha) = y(x, 0) + \delta y$. We may write J as a function of α given the variation we just described:

$$J(\alpha) = \int_{x_1}^{x_2} f(y(x, \alpha), y(x, \alpha)', x)\, dx.$$

It is intuitively obvious that if the function $y^*(x) = y(x, 0)$ is an extremum of the functional then we ought to expect

$$\left[\frac{\partial J(\alpha)}{\partial \alpha}\right]_{\alpha=0} = 0$$

Notice that we can calculate the derivative above using multivariate calculus. Remember that $y(x, \alpha) = y(x, 0) + \alpha\eta(x)$ hence $y(x, \alpha)' = y(x, 0)' + \alpha\eta(x)'$ thus $\frac{\partial y}{\partial \alpha} = \eta$ and $\frac{\partial y'}{\partial \alpha} = \eta' = \frac{d\eta}{dx}$. Consider that:

$$\begin{aligned}
\frac{\partial J(\alpha)}{\partial \alpha} &= \frac{\partial}{\partial \alpha}\left[\int_{x_1}^{x_2} f(y(x, \alpha), y(x, \alpha)', x)\, dx \right] \\
&= \int_{x_1}^{x_2} \left(\frac{\partial f}{\partial y}\frac{\partial y}{\partial \alpha} + \frac{\partial f}{\partial y'}\frac{\partial y'}{\partial \alpha} + \frac{\partial f}{\partial x}\frac{\partial x}{\partial \alpha} \right) dx \\
&= \int_{x_1}^{x_2} \left(\frac{\partial f}{\partial y}\eta + \frac{\partial f}{\partial y'}\frac{d\eta}{dx} \right) dx \qquad (6.1)
\end{aligned}$$

Observe that

$$\frac{d}{dx}\left[\frac{\partial f}{\partial y'}\eta\right] = \frac{d}{dx}\left[\frac{\partial f}{\partial y'}\right]\eta + \frac{\partial f}{\partial y'}\frac{d\eta}{dx}$$

Hence continuing Equation 6.1 in view of the product rule above,

$$\begin{aligned}
\frac{\partial J(\alpha)}{\partial \alpha} &= \int_{x_1}^{x_2} \left(\frac{\partial f}{\partial y}\eta + \frac{d}{dx}\left[\frac{\partial f}{\partial y'}\eta\right] - \frac{d}{dx}\left[\frac{\partial f}{\partial y'}\right]\eta \right) dx \\
&= \frac{\partial f}{\partial y'}\eta\Big|_{x_1}^{x_2} + \int_{x_1}^{x_2} \left(\frac{\partial f}{\partial y}\eta - \frac{d}{dx}\left[\frac{\partial f}{\partial y'}\right]\eta \right) dx \qquad (6.2) \\
&= \int_{x_1}^{x_2} \left(\frac{\partial f}{\partial y} - \frac{d}{dx}\left[\frac{\partial f}{\partial y'}\right] \right)\eta\, dx
\end{aligned}$$

Note we used the conditions $\eta(x_1) = \eta(x_2)$ to see that $\frac{\partial f}{\partial y'}\eta\big|_{x_1}^{x_2} = \frac{\partial f}{\partial y'}\eta(x_2) - \frac{\partial f}{\partial y'}\eta(x_1) = 0$. Our goal is to find the extreme values for the functional J. Let me take a few sentences to again restate our set-up. Generally, we take a function y then J maps to a new function $J[y]$. The family of functions indexed by α gives a whole ensemble of functions in $\mathcal{F}_o$ which are near y^* according to the formula,

$$y(x, \alpha) = y^*(x) + \alpha\eta(x)$$

Let's call this set of functions W_η. If we took another function like η, say ζ such that $\zeta(x_1) = \zeta(x_2) = 0$ then we could look at another family of functions:

$$y(x, \alpha) = y^*(x) + \alpha\zeta(x)$$

and we could denote the set of all such functions generated from ζ to be W_ζ. The total variation of y based at y^* should include all possible families of functions in $\mathcal{F}_o$. You could think of W_η and W_ζ be two different subspaces in $\mathcal{F}_o$. If $\eta \neq \zeta$ then these subspaces of $\mathcal{F}_o$ are likely disjoint except for the proposed extremal solution y^*. It is perhaps a bit unsettling to realize there are infinitely many such subspaces because there are infinitely many choices for the function η or ζ. In any event, each possible variation of y^* must satisfy the condition $\left[\frac{\partial J(\alpha)}{\partial \alpha}\right]_{\alpha=0} = 0$ since we **assume** that y^* is an extreme value of the functional J. It follows that the Equation 6.2 holds for all possible η. Therefore, we ought to expect that any extreme value of the functional $J[y] = \int_{x_1}^{x_2} f(y, y', x)\, dx$ must solve the **Euler Lagrange Equations:**

$$\boxed{\frac{\partial f}{\partial y} - \frac{d}{dx}\left[\frac{\partial f}{\partial y'}\right] = 0 \ \ \text{Euler-Lagrange Equations for} \ \ J[y] = \int_{x_1}^{x_2} f(y, y', x)\, dx}$$

6.3 variational derivative

The role that η played in the discussion in the preceding section is somewhat similar to the role that the "h" plays in the definition $f'(a) = \lim_{h \to 0} \frac{f(a+h)-f(a)}{h}$. You might hope we could replace arguments in η with a more direct approach. Physicists have a heuristic way of making such arguments in terms of the variation δ. They would cast the arguments in the last page by just "taking the variation of J". Let me give you their formal argument,

$$\begin{aligned}
\delta J &= \delta\left[\int_{x_1}^{x_2} f(y, y', x)\, dx\right] \\
&= \left[\int_{x_1}^{x_2} \delta f(y, y', x)\, dx\right] \\
&= \int_{x_1}^{x_2} \left(\frac{\partial f}{\partial y}\delta y + \frac{\partial f}{\partial y'}\delta\big(\tfrac{dy}{dx}\big) + \frac{\partial f}{\partial x}\delta x\right) dx \\
&= \int_{x_1}^{x_2} \left(\frac{\partial f}{\partial y}\delta y + \frac{\partial f}{\partial y'}\frac{d}{dx}(\delta y)\right) dx \qquad (6.3) \\
&= \frac{\partial f}{\partial y'}\delta y\Big|_{x_1}^{x_2} + \int_{x_1}^{x_2}\left(\frac{\partial f}{\partial y} - \frac{d}{dx}\left[\frac{\partial f}{\partial y'}\right]\right)\delta y\, dx
\end{aligned}$$

Therefore, since $\delta y = 0$ at the endpoints of integration, the Euler-Lagrange equations follow from $\delta J = 0$. Now, if you're like me, the argument above is less than satisfying since we never actually defined what it means to "take δ" of something. Also, why could I commute the variational δ and $\frac{d}{dx}$)? That said, the formal method is not without use since it allows the focus to be on the Euler Lagrange equations rather than the technical details of the variation.

Remark 6.3.1.

The more adept reader at this point should realize the hypocrisy of me calling the above calculation formal since even my presentation here was formal. I also used an analogy, I assumed that the theory of extreme values for multivariate calculus extends to function space. But, $\mathcal{F}_o$ is not $\mathbb{R}^n$, it's much bigger. Edwards builds the correct formalism for a rigourous calculation of the variational derivative. To be careful we'd need to develop the norm on function space and prove a number of results about infinite dimensional linear algebra. Take a look at the last chapter in Edwards' text if you're interested. I don't believe I'll have time to go over that material this semester.

6.4 Euler-Lagrange examples

I present a few standard examples in this section. We make use of the calculation in the last section. Also, we will use a result from your homework which states an equivalent form of the Euler-Lagrange equation is

$$\frac{\partial f}{\partial x} - \frac{d}{dx}\left[f - y'\frac{\partial f}{\partial y'}\right] = 0.$$

This form of the Euler Lagrange equation yields better differential equations for certain examples.

6.4.1 shortest distance between two points in plane

If s denotes the arclength in the xy-plane then the pythagorean theorem gives $ds^2 = dx^2 + dy^2$ infinitesimally. Thus, $ds = \sqrt{1 + \frac{dy}{dx}^2}\, dx$ and we may add up all the little distances ds to find the total length between two given points (x_1, y_1) and (x_2, y_2):

$$J[y] = \int_{x_1}^{x_2} \sqrt{1 + (y')^2}\, dx$$

Identify that we have $f(y, y', x) = \sqrt{1 + (y')^2}$. Calculate then,

$$\frac{\partial f}{\partial y} = 0 \qquad \text{and} \qquad \frac{\partial f}{\partial y'} = \frac{y'}{\sqrt{1 + (y')^2}}.$$

Euler Lagrange equations yield,

$$\frac{d}{dx}\left[\frac{\partial f}{\partial y'}\right] = \frac{\partial f}{\partial y} \qquad \Rightarrow \qquad \frac{d}{dx}\left[\frac{y'}{\sqrt{1 + (y')^2}}\right] = 0 \qquad \Rightarrow \qquad \frac{y'}{\sqrt{1 + (y')^2}} = k$$

where $k \in \mathbb{R}$ is constant with respect to x. Moreover, square both sides to reveal

$$\frac{(y')^2}{1 + (y')^2} = k^2 \qquad \Rightarrow \qquad (y')^2 = \frac{k^2}{1 - k^2} \qquad \Rightarrow \qquad \frac{dy}{dx} = \pm\sqrt{\frac{k^2}{1 - k^2}} = m$$

where I have defined m is defined in the obvious way. We find solutions $y = mx + b$. Finally, we can find m, b to fit the given pair of points (x_1, y_1) and (x_2, y_2) as follows:

$$y_1 = mx_1 + b \qquad \text{and} \qquad y_2 = mx_2 + b \qquad \Rightarrow \qquad y = y_1 + \frac{y_2 - y_1}{x_2 - x_1}(x - x_1)$$

provided $x_1 \neq x_2$. If $x_1 \neq x_2$ and $y_1 \neq y_2$ then we could perform the same calculation as above with the roles of x and y interchanged,

$$J[x] = \int_{y_1}^{y_2} \sqrt{1 + (x')^2}\, dy$$

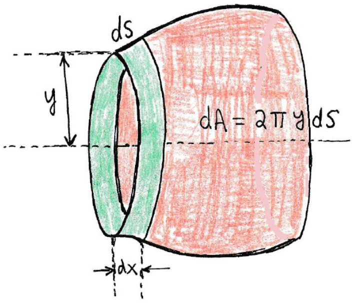
where $x' = dx/dy$ and the Euler Lagrange equations would yield the solution

$$x = x_1 + \frac{x_2 - x_1}{y_2 - y_1}(y - y_1).$$

Finally, if both coordinates are equal then $(x_1, y_1) = (x_2, y_2)$ and the shortest path between these points is the trivial path, the armchair solution. Silly comments aside, we have shown that a straight line provides the curve with the shortest arclength between any two points in the plane.

6.4.2 surface of revolution with minimal area

Suppose we wish to revolve some curve which connects (x_1, y_1) and (x_2, y_2) around the x-axis. A surface constructed in this manner is called a **surface of revolution**. In calculus we learn how to calculate the surface area of such a shape. One can imagine deconstructing the surface into a sequence of ribbons. Each ribbon at position x will have a "radius" of y and a width of dx however, because the shape is tilted the area of the ribbon works out to $dA = 2\pi y ds$ where ds is the arclength. I made a ribbon green in the picture below. You can imagine many ribbons approximating the surface, although, I made no attempt to draw those here:



If we choose x as the parameter this yields $dA = 2\pi y\sqrt{1+(y')^2}\,dx$. To find the surface of minimal surface area we ought to consider the functional:

$$A[y] = \int_{x_1}^{x_2} 2\pi y\sqrt{1+(y')^2}\,dx$$

Identify that $f(y, y', x) = 2\pi y\sqrt{1+(y')^2}$ hence $f_y = 2\pi\sqrt{1+(y')^2}$ and $f_{y'} = 2\pi y y'/\sqrt{1+(y')^2}$. The usual Euler-Lagrange equations are not easy to solve for this problem, it's easier to work with the equations you derived in homework,

$$\frac{\partial f}{\partial x} - \frac{d}{dx}\left[f - y'\frac{\partial f}{\partial y'}\right] = 0.$$

Hence,

$$\frac{d}{dx}\left[2\pi y\sqrt{1+(y')^2} - \frac{2\pi y(y')^2}{\sqrt{1+(y')^2}}\right] = 0$$

Dividing by 2π and making a common denominator,

$$\frac{d}{dx}\left[\frac{y}{\sqrt{1+(y')^2}}\right] = 0 \qquad \Rightarrow \qquad \frac{y}{\sqrt{1+(y')^2}} = k$$

where k is a constant with respect to x. Squaring the equation above yields

$$\frac{y^2}{1+(\frac{dy}{dx})^2} = k^2 \qquad \Rightarrow \qquad y^2 - k^2 = k^2(\tfrac{dy}{dx})^2$$

Solve for dx, integrate, assuming the given points are in the first quadrant,

$$x = \int dx = \int \frac{k dy}{\sqrt{y^2-k^2}} = k\cosh^{-1}(\tfrac{y}{k}) + c$$

Hence,

$$\boxed{y = k\cosh\left(\frac{x-c}{k}\right)}$$

generates the surface of revolution of least area between two points. These shapes are called **Catenoids** they can be observed in the formation of soap bubble between rings. There is a vast literature on this subject and there are many cases to consider, I simply exhibit a simple solution. For a given pair of points it is not immediately obvious if there exists a solution to the Euler-Lagrange equations which fits the data. (see page 622 of Arfken).

6.4.3 Braichistochrone

Suppose a particle slides freely along some curve from (x_1, y_1) to $(x_2, y_2) = (0, 0)$ under the influence of gravity where we take y to be the vertical direction. **What is the curve of quickest descent?** Notice that if $x_1 = 0$ then the answer is easy to see, however, if $x_1 \neq 0$ then the question is not trivial. To solve this problem we must first offer a functional which accounts for the time of descent. Note that the speed $v = ds/dt$ so we'd clearly like to minimize $J = \int_{(0,0)}^{(x_1,y_1)} \frac{ds}{v}$. Since the object is assumed to fall freely we may assume that energy is conserved in the motion hence

$$\frac{1}{2}mv^2 = mg(y - y_1) \qquad \Rightarrow \qquad v = \sqrt{2g(y_1 - y)}$$

As we've discussed in previous examples, $ds = \sqrt{1 + (y')^2}dt$ so we find

$$J[y] = \int_0^{x_1} \underbrace{\sqrt{\frac{1 + (y')^2}{2g(y_1 - y)}}}_{f(y,y',x)} \, dx$$

Notice that the modified Euler-Lagrange equations $\frac{\partial f}{\partial x} - \frac{d}{dx}\left[f - y'\frac{\partial f}{\partial y'}\right] = 0$ are convenient since $f_x = 0$. We calculate that

$$\frac{\partial f}{\partial y'} = \frac{1}{2\sqrt{\frac{1+(y')^2}{2g(y_1-y)}}} \frac{2y'}{2g(y_1 - y)} = \frac{y'}{\sqrt{2g(y_1 - y)(1 + (y')^2)}}$$

Hence there should exist some constant $1/(k\sqrt{2g})$ such that

$$\sqrt{\frac{1 + (y')^2}{2g(y_1 - y)}} - \frac{(y')^2}{\sqrt{2g(y_1 - y)(1 + (y')^2)}} = \frac{1}{k\sqrt{2g}}$$

It follows that,

$$\frac{1}{\sqrt{(y_1 - y)(1 + (y')^2)}} = \frac{1}{k} \qquad \Rightarrow \qquad \left(y_1 - y\right)\left(1 + \left(\frac{dy}{dx}\right)^2\right) = k^2$$

We need to solve for dy/dx,

$$(y_1 - y)\left(\frac{dy}{dx}\right)^2 = k^2 - y_1 + y \qquad \Rightarrow \qquad \left(\frac{dy}{dx}\right)^2 = \frac{y + k^2 - y_1}{y_1 - y}$$

Or, relabeling constants $a = y_1$ and $b = k^2 - y_1$ and we must solve

$$\frac{dy}{dx} = \pm\sqrt{\frac{b + y}{a - y}} \qquad \Rightarrow \qquad x = \pm\int\sqrt{\frac{a - y}{b + y}}\, dy$$

The integral is not trivial. It turns out that the solution is a cycloid (Arfken p. 624):

$$\boxed{x = \frac{a+b}{2}\Big(\theta + \sin(\theta)\Big) - d \qquad y = \frac{a+b}{2}\Big(1 - \cos(\theta)\Big) - b}$$

This is the curve that is traced out by a point on a wheel as it travels. If you take this solution and calculate $J[y_{cycloid}]$ you can show the time of descent is simply

$$T = \frac{\pi}{2}\sqrt{\frac{y_1}{2g}}$$

if the mass begins to descend from (x_2, y_2). But, this point has no connection with (x_1, y_1) except that they both reside on the same cycloid. It follows that the period of a pendulum that follows a cycloidal path is indpendent of the starting point on the path. This is not true for a circular pendulum in general, we need the small angle approximation to derive simple harmonic motion.

It turns out that it is possible to make a pendulum follow a cycloidal path if you let the string be guided by a frame which is also cycloidal. The neat thing is that even as it loses energy it still follows a cycloidal path and hence has the same period. The "Brachistochrone" problem was posed by Johann Bernoulli in 1696 and it actually predates the variational calculus of Lagrange by some 50 or so years. This problem and ones like it are what eventually prompted Lagrange and Euler to systematically develop the subject. Apparently Galileo also studied this problem however lacked the mathematics to crack it.

See this Geogebra demonstration to compare and contrast lines, verses parabolas, verses the cycloid. A google search will show you dozens of these.

6.5 Euler-Lagrange equations for several dependent variables

We still consider problems with just one independent parameter underlying everything. For problems of classical mechanics this is almost always time t. In anticipation of that application we choose to use the usual physics notation in the section. We suppose that our functional depends on functions $y_1, y_2, \dots, y_n$ of time t along with their time derivatives $\dot{y}_1, \dot{y}_2, \dots, \dot{y}_n$. We again suppose the functional of interest is an integral of a **Lagrangian** function f from time t_1 to time t_2,

$$J[(y_i)] = \int_{t_1}^{t_2} f(y_i, \dot{y}_i, t)\, dt$$

here we use (y_i) as shorthand for $(y_1, y_2, \dots, y_n)$ and $(\dot{y}_i)$ as shorthand for $(\dot{y}_1, \dot{y}_2, \dots, \dot{y}_n)$. We suppose that n-conditions are given for each of the endpoints in this problem; $y_i(t_1) = y_{i1}$ and $y_i(t_2) = y_{i2}$. Moreover, we define $\mathcal{F}_o$ to be the set of paths from $\mathbb{R}$ to $\mathbb{R}^n$ subject to the conditions just stated. We now set out to find necessary conditions on a proposed solution to the extreme value problem for the functional J above. As before let's assume that an extremal solution $y* \in \mathcal{F}_o$ exists. Moreover, imagine varying the solution by some variational function $\eta = (\eta_i)$ which has $\eta(t_1) = (0, 0, \dots, 0)$ and $\eta(t_2) = (0, 0, \dots, 0)$. Consequently the family of paths defined below are all in $\mathcal{F}_o$,

$$y(t, \alpha) = y^*(t) + \alpha\eta(t)$$

Thus $y(t, 0) = y^*$. In terms of component functions we have that

$$y_i(t, \alpha) = y_i^*(t) + \alpha\eta_i(t).$$

You can identify that $\delta y_i = y_i(t,\alpha) - y_i^*(t) = \alpha\eta_i(t)$. Since y^* is an extreme solution we should expect that $\left(\frac{\partial J}{\partial \alpha}\right)_{\alpha=0} = 0$. Differentiate the functional with respect to α and make use of the chain rule for f which is a function of some $2n+1$ variables,

$$\begin{aligned}
\frac{\partial J(\alpha)}{\partial \alpha} &= \frac{\partial}{\partial \alpha}\left[\int_{t_1}^{t_2} f(y_i(t,\alpha), \dot{y}_i(t,\alpha), t)\, dt\right] \\
&= \int_{t_1}^{t_2} \sum_{j=1}^{n}\left(\frac{\partial f}{\partial y_j}\frac{\partial y_j}{\partial \alpha} + \frac{\partial f}{\partial \dot{y}_j}\frac{\partial \dot{y}_j}{\partial \alpha}\right) dt \\
&= \int_{t_1}^{t_2} \sum_{j=1}^{n}\left(\frac{\partial f}{\partial y_j}\eta_j + \frac{\partial f}{\partial \dot{y}_j}\frac{d\eta_j}{dt}\right) dt \qquad (6.4) \\
&= \sum_{j=1}^{n} \frac{\partial f}{\partial \dot{y}_j}\eta\Big|_{t_1}^{t_2} + \int_{t_1}^{t_2}\sum_{j=1}^{n}\left(\frac{\partial f}{\partial y_j} - \frac{d}{dt}\frac{\partial f}{\partial \dot{y}_j}\right)\eta_j\, dt
\end{aligned}$$

Since $\eta(t_1) = \eta(t_2) = 0$ the first term vanishes. Moreover, since we may repeat this calculation for all possible variations about the optimal solution y^* it follows that we obtain a set of Euler-Lagrange equations for each component function of the solution:

$$\boxed{\frac{\partial f}{\partial y_j} - \frac{d}{dt}\left[\frac{\partial f}{\partial \dot{y}_j}\right] = 0 \quad j = 1, 2, \dots n \quad \text{Euler-Lagrange Eqns. for} \quad J[(y_i)] = \int_{t_1}^{t_2} f(y_i, \dot{y}_i, t)\, dt}$$

Often we simply use $y_1 = x$, $y_2 = y$ and $y_3 = z$ which denote the position of particle or perhaps just the component functions of a path which gives the geodesic on some surface. In either case we should have 3 sets of Euler-Lagrange equations, one for each coordinate. We will also use non-Cartesian coordinates to describe certain Lagrangians. We develop many useful results for set-up of Lagrangians in non-Cartesian coordinates in the next section.

6.5.1 free particle Lagrangian

For a particle of mass m the kinetic energy K is given in terms of the time derivatives of the coordinate functions x, y, z as follows:

$$K = \tfrac{m}{2}\left(\dot{x}^2 + \dot{y}^2 + \dot{z}^2\right)$$

Construct a functional by integrating the kinetic energy over time t,

$$S = \int_{t_1}^{t_2} \tfrac{m}{2}\left(\dot{x}^2 + \dot{y}^2 + \dot{z}^2\right)\, dt$$

The Euler-Lagrange equations for this functional are

$$\frac{\partial K}{\partial x} = \frac{d}{dt}\left[\frac{\partial K}{\partial \dot{x}}\right] \qquad \frac{\partial K}{\partial y} = \frac{d}{dt}\left[\frac{\partial K}{\partial \dot{y}}\right] \qquad \frac{\partial K}{\partial z} = \frac{d}{dt}\left[\frac{\partial K}{\partial \dot{z}}\right]$$

Since $\frac{\partial K}{\partial \dot{x}} = m\dot{x}$, $\frac{\partial K}{\partial \dot{y}} = m\dot{y}$ and $\frac{\partial K}{\partial \dot{z}} = m\dot{z}$ it follows that

$$\boxed{0 = m\ddot{x} \qquad 0 = m\ddot{y} \qquad 0 = m\ddot{z}.}$$

You should recognize these as Newton's equation for a particle with no force applied. The solution is $(x(t), y(t), z(t)) = (x_o + tv_x, y_o + tv_y, z_o + tv_z)$ which is uniform rectilinear motion at constant velocity (v_x, v_y, v_z). The solution to Newton's equation minimizes the integral of the Kinetic energy. Generally the quantity S is called the **action** and Hamilton's Principle states that the laws of physics all arise from minimizing the action of the physical phenomena. We'll return to this discussion in a later section.

6.5.2 geodesics in $\mathbb{R}^3$

A **geodesic** is the path of minimal length between a pair of points on some manifold. Note we already proved that geodesics in the plane are just lines. In general, for $\mathbb{R}^3$, the square of the infinitesimal arclength element is $ds^2 = dx^2 + dy^2 + dz^2$. The arclength integral from $p = 0$ to $q = (q_x, q_y, q_z)$ in $\mathbb{R}^3$ is most naturally given from the parametric viewpoint:

$$S = \int_0^1 \sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}\, dt$$

We assume $(x(0), y(0), z(0)) = (0, 0, 0)$ and $(x(1), y(1), z(1)) = q$ and it should be clear that the integral above calculates the arclength. The Euler-Lagrange equations for x, y, z are

$$\frac{d}{dt}\left[\frac{\dot{x}}{\sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}}\right] = 0, \qquad \frac{d}{dt}\left[\frac{\dot{y}}{\sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}}\right] = 0, \qquad \frac{d}{dt}\left[\frac{\dot{z}}{\sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}}\right] = 0.$$

It follows that there exist constants, say a, b and c, such that

$$a = \frac{\dot{x}}{\sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}}, \qquad b = \frac{\dot{y}}{\sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}}, \qquad c = \frac{\dot{z}}{\sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}}.$$

These equations are said to be **coupled** since each involves derivatives of the others. We usually need a way to uncouple the equations if we are to be successful in solving the system. We can calculate, and equate each with the constant 1:

$$1 = \frac{\dot{x}}{a\sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}} = \frac{\dot{y}}{b\sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}} = \frac{\dot{z}}{c\sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}}.$$

But, multiplying by the denominator reveals an interesting identity

$$\sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2} = \frac{\dot{x}}{a} = \frac{\dot{y}}{b} = \frac{\dot{z}}{c}$$

The solution has the form, $x(t) = tq_x$, $y(t) = tq_y$ and $z(t) = tq_z$. Therefore,

$$(x(t), y(t), z(t)) = t(q_x, q_y, q_z) = tq.$$

for $0 \leq t \leq 1$. These are the parametric equations for the line segment from the origin to q.

6.6 the Euclidean metric

The square root in the functional of the last subsection certainly complicated the calculation. It is intuitively clear that if we add up squared line elements ds^2 to give a minimum then that ought to correspond to the minimum for the sum of the positive square roots ds of those elements. Let's check if my conjecture works for $\mathbb{R}^3$:

$$S = \int_0^1 (\ \underbrace{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}_{f(x,y,z,\dot{x},\dot{y},\dot{z})}\ )\, dt$$

This gives us the Euler Lagrange equations below:

$$\ddot{x} = 0, \qquad \ddot{y} = 0, \qquad \ddot{z} = 0$$

The solution of these equations is clearly a line. In this formalism the equations were uncoupled from the outset.

Definition 6.6.1.

The Euclidean metric is $ds^2 = dx^2 + dy^2 + dz^2$. Generally, for orthogonal curvelinear coordinates u, v, w we calculate $ds^2 = \frac{1}{||\nabla u||^2}du^2 + \frac{1}{||\nabla v||^2}dv^2 + \frac{1}{||\nabla w||^2}dw^2$. We use this as a guide for constructing functionals which calculate arclength or speed

The beauty of the metric is that it allows us to calculate in other coordinates, consider

$$x = r\cos(\theta) \qquad y = r\sin(\theta)$$

For which we have implicit inverse coordinate transformations $r^2 = x^2 + y^2$ and $\theta = \tan^{-1}(y/x)$. From these inverse formulas we calculate:

$$\nabla r \ = < x/r, y/r > \qquad \nabla\theta \ = < -y/r^2, x/r^2 >$$

Thus, $||\nabla r|| = 1$ whereas $||\nabla\theta|| = 1/r$. We find that the metric in polar coordinates takes the form:

$$\boxed{ds^2 = dr^2 + r^2 d\theta^2}$$

Physicists and engineers tend to like to think of these as arising from calculating the length of infinitesimal displacements in the r or θ directions. Generically, for u, v, w coordinates

$$dl_u = \frac{1}{||\nabla u||}du \qquad dl_v = \frac{1}{||\nabla v||}dv \qquad dl_w = \frac{1}{||\nabla w||}dw$$

and $ds^2 = dl_u^2 + dl_v^2 + dl_w^2$. So in that notation we just found $dl_r = dr$ and $dl_\theta = rd\theta$. Notice then that cylindircal coordinates have the metric,

$$\boxed{ds^2 = dr^2 + r^2 d\theta^2 + dz^2.}$$

For spherical coordinates $x = r\cos(\phi)\sin(\theta)$, $y = r\sin(\phi)\sin(\theta)$ and $z = r\cos(\theta)$ (here $0 \leq \phi \leq 2\pi$ and $0 \leq \theta \leq \pi$, physics notation). Calculation of the metric follows from the line elements,

$$dl_r = dr \qquad dl_\phi = r\sin(\theta)d\phi \qquad dl_\theta = rd\theta$$

Thus,

$$\boxed{ds^2 = dr^2 + r^2\sin^2(\theta)d\phi^2 + r^2 d\theta^2.}$$

We now have all the tools we need for examples in spherical or cylindrical coordinates. What about other cases? In general, given some p-manifold in $\mathbb{R}^n$ how does one find the metric on that manifold? If we are to follow the approach of this section we'll need to find coordinates on $\mathbb{R}^n$ such that the manifold S is described by setting all but p of the coordinates to a constant. For example, in $\mathbb{R}^4$ we have generalized cylindircal coordinates (r, ϕ, z, t) defined implicitly by the equations below

$$x = r\cos(\phi), \qquad y = r\sin(\phi), \qquad z = z, \qquad t = t$$

On the hyper-cylinder $r = R$ we have the metric $ds^2 = R^2 d\theta^2 + dz^2 + dw^2$. There are mathematicians/physicists whose careers are founded upon the discovery of a metric for some manifold. This is generally a difficult task.

6.7 geodesics

A **geodesic** is a path of smallest distance on some manifold. In general relativity, it turns out that the solutions to Eistein's field equations are geodesics in 4-dimensional curved spacetime. Particles that fall freely are following geodesics, for example projectiles or planets in the absense of other frictional/non-gravitational forces. We don't follow a geodesic in our daily life because the earth pushes back up with a normal force. Also, do be honest, the idea of length in general relativity is a bit more abstract that the geometric length studied in this section. The metric of general relativity is non-Euclidean. General relativity is based on semi-Riemannian geometry whereas this section is all Riemannian geometry. The metric in Riemannian geometry is positive definite. The metric in semi-Riemannian geometry can be written as a quadratic form with both positive and negative eigenvalues. In any event, if you want to know more I know some books you might like.

6.7.1 geodesic on cylinder

The equation of a cylinder of radius R is most easily framed in cylindrical coordinates (r, θ, z); the equation is merely $r = R$ hence the metric reads

$$ds^2 = R^2 d\theta^2 + dz^2$$

Therefore, we ought to minimize the following functional in order to locate the parametric equations of a geodesic on the cylinder: note $ds^2 = \big(R^2 \frac{d\theta^2}{dt^2} + \frac{dz^2}{dt^2}\big)dt^2$ thus:

$$S = \int \left(R^2 \dot{\theta}^2 + \dot{z}^2\right) dt$$

Euler-Lagrange equations for the dependent variables θ and z are simply:

$$\ddot{\theta} = 0 \qquad \ddot{z} = 0.$$

We can integrate twice to find solutions

$$\boxed{\theta(t) = \theta_o + At \qquad z(t) = z_o + Bt}$$

Therefore, the geodesic on a cylinder is simply the line connecting two points in the plane which is curved to assemble the cylinder. Simple cases that are easy to understand:

1. Geodesic from $(R\cos(\theta_o), R\sin(\theta_o), z_1)$ to $(R\cos(\theta_o), R\sin(\theta_o), z_2)$ is parametrized by $\theta(t) = \theta_o$ and $z(t) = z_1 + t(z_2 - z_1)$ for $0 \le t \le 1$. Technically, there is some ambiguity here since I never declared over what range the t is to range. Could pick other intervals, we could use z at the parameter is we wished then $\theta(z) = \theta_o$ and $z = z$ for $z_1 \le z \le z_2$

2. Geodesic from $(R\cos(\theta_1), R\sin(\theta_1), z_o)$ to $(R\cos(\theta_2), R\sin(\theta_2), z_o)$ is parametrized by $\theta(t) = \theta_1 + t(\theta_2 - \theta_1)$ and $z(t) = z_o$ for $0 \le t \le 1$.

3. Geodesic from $(R\cos(\theta_1), R\sin(\theta_1), z_1)$ to $(R\cos(\theta_2), R\sin(\theta_2), z_2)$ is parametrized by

   $$\theta(t) = \theta_1 + t(\theta_2 - \theta_1) \qquad z(t) = z_1 + t(z_2 - z_1)$$

   You can eliminate t and find the equation $z = \frac{z_2 - z_1}{\theta_2 - \theta_1}(\theta - \theta_1)$ which again just goes to show you this is a line in the curved coordinates.

6.7.2 geodesic on sphere

The equation of a sphere of radius R is most easily framed in spherical coordinates (r, ϕ, θ); the equation is merely $r = R$ hence the metric reads

$$ds^2 = R^2 \sin^2(\theta) d\phi^2 + R^2 d\theta^2.$$

Therefore, we ought to minimize the following functional in order to locate the parametric equations of a geodesic on the sphere: note $ds^2 = \big(R^2 \sin^2(\theta)\frac{d\phi^2}{dt^2} + R^2 \frac{d\theta^2}{dt^2}\big)dt^2$ thus:

$$S = \int \Big(\ \underbrace{R^2 \sin^2(\theta)\dot{\phi}^2 + R^2\dot{\theta}^2}_{f(\theta,\phi,\dot{\theta},\dot{\phi})}\ \Big)\ dt$$

Euler-Lagrange equations for the dependent variables ϕ and θ are simply: $f_\theta = \frac{d}{dt}(f_{\dot{\theta}})$ and $f_\phi = \frac{d}{dt}(f_{\dot{\phi}})$ which yield:

$$2R^2 \sin(\theta)\cos(\theta)\dot{\phi}^2 = \tfrac{d}{dt}(2R^2\dot{\theta}) \qquad 0 = \frac{d}{dt}\Big(2R^2 \sin^2(\theta)\dot{\phi}\Big).$$

We find a **constant of motion** $L = 2R^2 \sin^2(\theta)\dot{\phi}$ inserting this in the equation for the azmuthial angle θ yields:

$$2R^2 \sin(\theta)\cos(\theta)\dot{\phi}^2 = \tfrac{d}{dt}(2R^2\dot{\theta}) \qquad 0 = \frac{d}{dt}\Big(2R^2 \sin^2(\theta)\dot{\phi}\Big).$$

If you can solve these and demonstrate through some reasonable argument that the solutions are great circles then would be happy to add your argument to these notes. That said, most of the solutions I've found online use $t = \theta$ to reduce the problem to solving a differential equation for ϕ as a function of θ[1]

6.8 Lagrangian mechanics

6.8.1 basic equations of classical mechanics summarized

Classical mechanics is the study of massive particles at relatively low velocities. Let me refresh your memory about the basics equations of Newtonian mechanics. Our goal in this section will be to rephrase Newtonian mechanics in the variational langauge and then to solve problems with the Euler-Lagrange equations. Newton's equations tell us how a particle of mass m evolves through time according to the net-force impressed on m. In particular,

$$m\frac{d^2\vec{r}}{dt^2} = \vec{F}$$

If m is not constant then you may recall that it is better to use momentum $\vec{P} = m\vec{v} = m\frac{d\vec{r}}{dt}$ to set-up Newton's 2nd Law:

$$\frac{d\vec{P}}{dt} = \vec{F}$$

In terms of components we have a system of differential equations with indpendent variable time t. If we use position as the dependent variable then Newton's 2nd Law gives three second order ODEs,

$$m\ddot{x} = F_x \qquad m\ddot{y} = F_y \qquad m\ddot{z} = F_z$$

[1] beware, when looking at a calculation involving spherical coordinates there are several choices for how to define θ and ϕ.

where $\vec{r} = (x, y, z)$ and the dots denote time-derivatives. Moreover, $\vec{F} =< F_x, F_y, F_z >$ is the sum of the forces that act on m. In contrast, if you work with momentum then you would want to solve six first order ODEs,

$$\dot{P}_x = F_x \qquad \dot{P}_y = F_y \qquad \dot{P}_z = F_z$$

and $P_x = m\dot{x}$, $P_y = m\dot{y}$ and $P_z = m\dot{z}$. These equations are easiest to solve when the force is not a function of velocity or time. In particular, if the force $\vec{F}$ is conservative then there exists a potential energy function $U : \mathbb{R}^3 \to \mathbb{R}$ such that $\vec{F} = -\nabla U$. We can prove that in the case the force is conservative the total energy is conserved.

6.8.2 kinetic and potential energy, formulating the Lagrangian

Recall the kinetic energy is $T = \frac{1}{2}m||\vec{v}||^2$, in Cartesian coordinates this gives us the formula:

$$T = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2).$$

If $\vec{F}$ is a conservative force then it is independent of path so we may construct the potential energy function as follows:

$$U(\vec{r}) = -\int_{\mathcal{O}}^{\vec{r}} \vec{F} \cdot d\vec{r}$$

Here $\mathcal{O}$ is the origin for the potential and we can prove that the potential energy constructed in this manner has $\vec{F} = -\nabla U$. We can prove that the total (mechanical) energy $E = T + U$ for a conservative system is a constant; $dE/dt = 0$. Hopefully these comments are at least vaguely familiar from some physics course in your distant memory. If not relax, calculationally this chapter is self-contained, read onward.

We already calculated that if we use T as the Lagrangian then the Euler-Lagrange equations produce Newton's equations in the case that the force is zero (see 6.5.1). Suppose that we define the Lagrangian to be $L = T - U$ for a system governed by a conservative force with potential energy function U. We seek to prove the Euler-Lagrange equations are precisely Newton's equations for this conservative system[2] Generically we have a Lagrangian of the form

$$L(x, y, z, \dot{x}, \dot{y}, \dot{z}) = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) - U(x, y, z).$$

We wish to find extrema for the functional $S = \int L(t)\, dt$. This yields three sets of Euler-Lagrange equations, one for each dependent variable x, y or z

$$\frac{d}{dt}\left[\frac{\partial L}{\partial \dot{x}}\right] = \frac{\partial L}{\partial x} \qquad \frac{d}{dt}\left[\frac{\partial L}{\partial \dot{y}}\right] = \frac{\partial L}{\partial y} \qquad \frac{d}{dt}\left[\frac{\partial L}{\partial \dot{z}}\right] = \frac{\partial L}{\partial z}.$$

Note that $\frac{\partial L}{\partial \dot{x}} = m\dot{x}$, $\frac{\partial L}{\partial \dot{y}} = m\dot{y}$ and $\frac{\partial L}{\partial \dot{z}} = m\dot{z}$. Also note that $\frac{\partial L}{\partial x} = -\frac{\partial U}{\partial x} = F_x$, $\frac{\partial L}{\partial y} = -\frac{\partial U}{\partial y} = F_y$ and $\frac{\partial L}{\partial z} = -\frac{\partial U}{\partial z} = F_z$. It follows that

$$\boxed{m\ddot{x} = F_x \qquad m\ddot{y} = F_y \qquad m\ddot{z} = F_z.}$$

Of course this is precisely $m\vec{a} = \vec{F}$ for a net-force $\vec{F} =< F_x, F_y, F_z >$. We have shown that **Hamilton's principle** reproduces Newton's Second Law for conservative forces. Let me take a moment to state it.

[2] don't mistake this example as an admission that Lagrangian mechanics is limited to conservative systems. Quite the contrary, Lagrangian mechanics is actually more general than the orginal framework of Newton!

Definition 6.8.1. Hamilton's Principle:

If a physical system has generalized coordinates q_j with velocities $\dot{q}_j$ and Lagrangian $L = T - U$ then the solutions of physics will minimize the action S defined below:

$$\boxed{S = \int_{t_1}^{t_2} L(q_j, \dot{q}_j, t)\, dt}$$

Mathematically, this means the variation $\delta S = 0$ for physical trajectories.

This is a necessary condition for solutions of the equations of physics. Sufficient conditions are known, you can look in any good variational calculus text. You'll find analogues to the second derivative test for variational differentiation. As far as I can tell physicists don't care about this logical gap, probably because the solutions to the Euler-Lagrange equations are the ones for which they are looking.

6.8.3 easy physics examples

Now, you might just see this whole exercise as some needless multiplication of notation and formalism. After all, I just told you we just get Newton's equations back from the Euler-Lagrange equations. To my taste the impressive thing about Lagrangian mechanics is that you get to start the problem with energy. Moreover, the Lagrangian formalism handles non-Cartesian coordinates with ease. If you search your memory from classical mechanics you'll notice that you either do constant acceleration, circular motion or motion along a line. What if you had a particle constrained to move in some frictionless ellipsoidal bowl. Or what if you had a pendulum hanging off another pendulum? How would you even write Newtons' equations for such systems? In contrast, the problem is at least easy to set-up in the Lagrangian approach. Of course, solutions may be less easy to obtain.

Example 6.8.2. Projectile motion: *take z as the vertical direction and suppose a bullet is fired with initial velocity $v_o =< v_{ox}, v_{oy}, v_{oz} >$. The potential energy due to gravity is simply $U = mgz$ and kinetic energy is given by $T = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2)$. Thus,*

$$L = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) - mgz$$

Euler-Lagrange equations are simply:

$$\frac{d}{dt}\Big[m\dot{x}\Big] = 0 \qquad \frac{d}{dt}\Big[m\dot{y}\Big] = 0 \qquad \frac{d}{dt}\Big[m\dot{z}\Big] = \frac{\partial}{\partial z}(-mgz) = -mg.$$

Integrating twice and applying initial conditions gives us the (possibly familiar) equations

$$x(t) = x_o + v_{ox}t, \qquad y(t) = y_o + v_{oy}t, \qquad z(t) = z_o + v_{oz}t - \tfrac{1}{2}gt^2.$$

Example 6.8.3. Simple Pendulum: *let θ denote angle measured off the vertical for a simple pendulum of mass m and length l. Trigonmetry tells us that*

$$x = l\sin(\theta) \qquad y = l\cos(\theta) \qquad \Rightarrow \qquad \dot{x} = l\cos(\theta)\dot{\theta} \qquad y = -l\sin(\theta)\dot{\theta}$$

Thus $T = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) = \frac{1}{2}ml^2\dot{\theta}^2$. Also, the potential energy due to gravity is $U = -mgl\cos(\theta)$ which gives us

$$L = \frac{1}{2}ml^2\dot{\theta}^2 + mgl\cos(\theta)$$

Then, the Euler-Lagrange equation in θ is simply:

$$\frac{d}{dt}\left[\frac{\partial L}{\partial \dot{\theta}}\right] = \frac{\partial L}{\partial \theta} \qquad \Rightarrow \qquad \frac{d}{dt}(ml^2\dot{\theta}) = -mgl\sin(\theta) \qquad \Rightarrow \qquad \ddot{\theta} + \frac{g}{l}\sin(\theta) = 0.$$

In the small angle approximation, $\sin(\theta) = \theta$ then we have the solution $\theta(t) = \theta_o\cos(\omega t + \phi_o)$ for angular frequency $\omega = \sqrt{g/l}$

Remark 6.8.4.

I intend to show a pdf in lecture where I show how to solve the central force problem. That discussion brings us to an explicit solution of Kepler's Laws which derives the explicit elliptical form of the orbits.

Chapter 7

multilinear algebra

The principle aim of this chapter is to introduce how to calculate with $\otimes$ and $\wedge$. We take a very concrete approach where the tensor and wedge product are understood in terms of multilinear mappings for which they form a basis. That said, there is a univerisal algebraic approach to construct the tensor and wedge products, I encourage the reader to study Dummit and Foote's *Abstract Algebra* Part III on *Modules and Vector Spaces* where these constructions are explained in a much broader algebraic context.

Beyond the basic definitions, we also study how wedge products capture determinants and give a natural language to ask certain questions of linear dependence. We also study metrics with particular attention to four dimensional Minkowski space with signature $(-1, 1, 1, 1)$. Hodge duality is detailed for three dimensional Euclidean space and four dimensional Minkowski space. However, there is sufficient detail that one ought to be able to extrapolate to euclidean spaces of other dimension. Moreover, the Hodge duality is reduced to a few tables for computational convenience. I encourage the reader to see David Bleeker's text for a broader discussion of Hodge duality with a physical bent.

When I lecture this material I'll probably just give examples to drive home the computational aspects. Also, it should be noted this Chapter can be studied without delving deeply into Sections 7.4 and 7.6. However, Chapter 9 requires some understanding of both of those sections.

7.1 dual space

Definition 7.1.1.

Suppose V is a vector space over $\mathbb{R}$. We define the **dual space** to V to be the set of all linear functions from V to $\mathbb{R}$. In particular, we denote:

$$V^* = \{f : V \to \mathbb{R} \mid f(x+y) = f(x) + f(y) \text{ and } f(cx) = cf(x) \ \ \forall x, y \in V \text{ and } c \in \mathbb{R}\}$$

If $\alpha \in V^*$ then we say α is a **dual vector**.

I offer several abstract examples to begin, however the majority of this section concerns $\mathbb{R}^n$.

Example 7.1.2. *Suppose $\mathcal{F}$ denotes the set of continuous functions on $\mathbb{R}$. Define $\alpha(f) = \int_0^1 f(t)\, dt$. The mapping $\alpha : \mathcal{F} \to \mathbb{R}$ is linear by properties of definite integrals therefore we identify the definite integral defines a dual-vector to the vector space of continuous functions.*

Example 7.1.3. *Suppose $V = \mathcal{F}(W, \mathbb{R})$ denotes a set of functions from a vector space W to $\mathbb{R}$. Note that V is a vector space with respect to point-wise defined addition and scalar multiplication of functions. Let $w_o \in W$ and define $\alpha(f) = f(w_o)$. The mapping $\alpha : V \rightarrow \mathbb{R}$ is linear since $\alpha(cf + g) = (cf + g)(w_o) = cf(w_o) + g(w_o) = c\alpha(f) + \alpha(g)$ for all $f, g \in V$ and $c \in \mathbb{R}$. We find that the* **evaluation** *map defines a dual-vector $\alpha \in V^*$.*

Example 7.1.4. *The determinant is a mapping from $\mathbb{R}^{\ n\times n}$ to $\mathbb{R}$ but it does not define a dual-vector to the vector space of square matrices since $det(A + B) \neq det(A) + det(B)$.*

Example 7.1.5. *Suppose $\alpha(x) = x \cdot v$ for a particular vector $v \in \mathbb{R}^n$. We argue $\alpha \in V^*$ where we recall $V = \mathbb{R}^n$ is a vector space. Additivity follows from a property of the dot-product on $\mathbb{R}^n$,*

$$\alpha(x + y) = (x + y) \cdot v = x \cdot v + y \cdot v = \alpha(x) + \alpha(y)$$

for all $x, y \in \mathbb{R}^n$. Likewise, homogeneity follows from another property of the dot-product: observe

$$\alpha(cx) = (cx) \cdot v = c(x \cdot v) = c\alpha(x)$$

for all $x \in \mathbb{R}^n$ and $c \in \mathbb{R}$.

Example 7.1.6. *Let $\alpha(x, y) = 2x + 5y$ define a function $\alpha : \mathbb{R}^2 \rightarrow \mathbb{R}$. Note that*

$$\alpha(x, y) = (x, y) \cdot (2, 5)$$

hence by the preceding example we find $\alpha \in (\mathbb{R}^2)^*$.

The preceding example is no accident. It turns out there is a one-one correspondance between row vectors and dual vectors on $\mathbb{R}^n$. Let $v \in \mathbb{R}^n$ then we define $\alpha_v(x) = x \cdot v$. We proved in Example 7.1.5 that $\alpha_v \in (\mathbb{R}^n)^*$. Suppose $\alpha \in (\mathbb{R}^n)^*$ we see to find $v \in \mathbb{R}^n$ such that $\alpha = \alpha_v$. Recall that a linear function is uniquely defined by its values on a basis; the values of α on the standard basis will show us how to choose v. This is a standard technique. Consider: $v \in \mathbb{R}^n$ with[1] $v = \sum_{j=1}^n v^j e_j$

$$\alpha(x) = \underbrace{\alpha(\sum_{j=1}^n x^j e_j) = \sum_{j=1}^n}_{additivity} \underbrace{\alpha(x^j e_j) = \sum_{j=1}^n}_{homogeneity} x^j \alpha(e_j) = x \cdot v$$

where we define $v = (\alpha(e_1), \alpha(e_2), \dots, \alpha(e_n)) \in \mathbb{R}^n$. The vector which corresponds naturally[2] to α is simply the vector of of the values of α on the standard basis.

The dual space to $\mathbb{R}^n$ is a vector space and the correspondance $v \rightarrow \alpha_v$ gives an isomorphism of $\mathbb{R}^n$ and $(\mathbb{R}^n)^*$. The image of a basis under an isomorphism is once more a basis. Define $\Phi : \mathbb{R}^n \rightarrow (\mathbb{R})^*$ by $\Phi(v) = \alpha_v$ to give the correspondance an explicit label. The image of the standard basis under Φ is called the **standard dual basis** for $(\mathbb{R}^n)^*$. Consider $\Phi(e_j)$, let $x \in \mathbb{R}^n$ and calculate

$$\Phi(e_j)(x) = \alpha_{e_j}(x) = x \cdot e_j$$

In particular, notice that when $x = e_i$ then $\Phi(e_j)(e_i) = e_i \cdot e_j = \delta_{ij}$. Dual vectors are linear transformations therefore we can define the dual basis by its values on the standard basis.

[1] the super-index is not a power in this context, it is just a notation to emphasize v^j is the component of a vector.

[2] some authors will say $\mathbb{R}^{n\times 1}$ is dual to $\mathbb{R}^{1\times n}$ since $\alpha_v(x) = v^T x$ and v^T is a row vector, I will avoid that langauge in these notes.

Definition 7.1.7.

The **standard dual basis** of $(\mathbb{R}^n)^*$ is denoted $\{e^1, e^2, \dots, e^n\}$ where we define $e^j : \mathbb{R}^n \to \mathbb{R}$ to be the linear transformation such that $e^j(e_i) = \delta_{ij}$ for all $i, j \in \mathbb{N}_n$. Generally, given a vector space V with basis $\beta = \{f_1, f_2, \dots, f_m\}$ we say the basis $\beta^* = \{f^1, f^2, \dots, f^n\}$ is dual to β iff $f^j(f_i) = \delta_{ij}$ for all $i, j \in \mathbb{N}_n$.

The term *basis* indicates that $\{e^1, e^2, \dots, e^n\}$ is linearly independent[3] and *span*$\{e^1, e^2, \dots, e^n\} = (\mathbb{R}^n)^*$. The following calculation is often useful: if $x \in \mathbb{R}^n$ with $x = \sum_{j=1}^n x^j e_j$ then

$$e^i(x) = e^i\left(\sum_{j=1}^n x^j e_j\right) = \sum_{j=1}^n x^j e^i(e_j) = \sum_{j=1}^n x^j \delta_{ij} = x^i \quad \Rightarrow \quad \boxed{e^i(x) = x^i.}$$

The calculation above is a prototype for many that follow in this chapter. Next, suppose $\alpha \in (\mathbb{R}^n)^*$ and suppose $x \in \mathbb{R}^n$ with $x = \sum_{j=1}^n x^j e_j$. Calculate,

$$\alpha(x) = \alpha\left(\sum_{i=1}^n x^i e_i\right) = \sum_{i=1}^n \alpha(e_i) e^i(x) \quad \Rightarrow \quad \boxed{\alpha = \sum_{i=1}^n \alpha(e_i) e^i}$$

this shows every dual vector is in the span of the dual basis $\{e^j\}_{j=1}^n$.

7.2 multilinearity and the tensor product

A multilinear mapping is a function of a Cartesian product of vector spaces which is linear with respect to each ”slot”. The goal of this section is to explain what that means. It turns out the set of all multilinear mappings on a particular set of vector spaces forms a vector space and we'll show how the tensor product can be used to construct an explicit basis by tensoring a bases which are dual to the bases in the domain. We also examine the concepts of symmetric and antisymmetric multilinear mappings, these form interesting subspaces of the set of all multilinear mappings. Our approach in this section is to treat the case of bilinearity in depth then transition to the case of multilinearity. Naturally this whole discussion demands a familarity with the preceding section.

7.2.1 bilinear maps

Definition 7.2.1.

Suppose V_1, V_2 are vector spaces then $b : V_1 \times V_2 \to \mathbb{R}$ is a **binear mapping** on $V_1 \times V_2$ iff for all $x, y \in V_1$, $z, w \in V_2$ and $c \in \mathbb{R}$:

$$\begin{array}{lrcll}
(1.) & b(cx + y, z) & = & cb(x, z) + b(y, z) & \text{(linearity in the first slot)} \\
(2.) & b(x, cz + w) & = & cb(x, z) + b(x, w) & \text{(linearity in the second slot).}
\end{array}$$

bilinear maps on $V \times V$

When $V_1 = V_2 = V$ we simply say that $b : V \times V \to \mathbb{R}$ is a **bilinear mapping on** V. The **set of all bilinear maps of** V **is denoted** $T^2_0 V$. You can show that $T^2_0 V$ forms a vector space under the usual point-wise defined operations of function addition and scalar multiplication[4]. Hopefully you are familar with the example below.

[3] direct proof of LI is left to the reader

[4] sounds like homework

Example 7.2.2. *Define $b : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ by $b(x, y) = x \cdot y$ for all $x, y \in \mathbb{R}^n$. Linearity in each slot follows easily from properties of dot-products:*

$$b(cx + y, z) = (cx + y) \cdot z = cx \cdot z + y \cdot z = cb(x, z) + b(y, z)$$

$$b(x, cy + z) = x \cdot (cy + z) = cx \cdot y + x \cdot z = cb(x, y) + b(x, z).$$

We can use matrix multiplication to generate a large class of examples with ease.

Example 7.2.3. *Suppose $A \in \mathbb{R}^{\ n\times n}$ and define $b : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ by $b(x, y) = x^T Ay$ for all $x, y \in \mathbb{R}^n$. Observe that, by properties of matrix multiplication,*

$$b(cx + y, z) = (cx + y)^T Az = (cx^T + y^T)Az = cx^T Az + y^T Az = cb(x, z) + b(y, z)$$

$$b(x, cy + z) = x^T A(cy + z) = cx^T Ay + x^T Az = cb(x, y) + b(x, z)$$

for all $x, y, z \in \mathbb{R}^n$ and $c \in \mathbb{R}$. It follows that b is bilinear on $\mathbb{R}^n$.

Suppose $b : V \times V \to \mathbb{R}$ is bilinear and suppose $\beta = \{e_1, e_2, \dots, e_n\}$ is a basis for V whereas $\beta^* = \{e^1, e^2, \dots, e^n\}$ is a basis of V^* with $e^j(e_i) = \delta_{ij}$

$$\begin{aligned} b(x, y) &= b\bigg(\sum_{i=1}^n x^i e_i, \ \sum_{j=1}^n y^j e_j \bigg) \qquad (7.1) \\ &= \sum_{i,j=1}^n b(x^i e_i, y^j e_j) \\ &= \sum_{i,j=1}^n x^i y^j b(e_i, e_j) \\ &= \sum_{i,j=1}^n b(e_i, e_j) e^i(x)^j(y) \end{aligned}$$

Therefore, if we define $b_{ij} = b(e_i, e_j)$ then we may compute $b(x, y) = \sum_{i,j=1}^n b_{ij} x^i y^j$. The calculation above also indicates that b is a linear combination of certain basic bilinear mappings. In particular, b can be written a linear combination of a tensor product of dual vectors on V.

Definition 7.2.4.

> Suppose V is a vector space with dual space V^*. If $\alpha, \beta \in V^*$ then we define $\alpha \otimes \beta : V \times V \to \mathbb{R}$ by $(\alpha \otimes \beta)(x, y) = \alpha(x)\beta(y)$ for all $x, y \in V$.

Given the notation[5] preceding this definition, we note $(e^i \otimes e^j)(x, y) = e^i(x)e^j(y)$ hence for all $x, y \in V$ we find:

$$b(x, y) = \sum_{i,j=1}^n b(e_i, e_j)(e^i \otimes e^j)(x, y) \ \text{ therefore, } \boxed{b = \sum_{i,j=1}^n b(e_i, e_j) e^i \otimes e^j}$$

We find[6] that $T^2_0 V = span\{e^i \otimes e^j\}_{i,j=1}^n$. Moreover, it can be argued[7] that $\{e^i \otimes e^j\}_{i,j=1}^n$ is a linearly independent set, therefore $\{e^i \otimes e^j\}_{i,j=1}^n$ forms a basis for $T^2_0 V$. We can count there are n^2 vectors

[5] perhaps you would rather write $(e^i \otimes e^j)(x, y)$ as $e^i \otimes e^j(x, y)$, that is also fine.

[6] with the help of your homework where you will show $\{e^i \otimes e^j\}_{i,j=1}^n \subseteq T^2_0 V$

[7] yes, again, in your homework

in $\{e^i \otimes e^j\}_{i,j=1}^n$ hence $dim(\,T^2_0\, V\,) = n^2$.

If $V = \mathbb{R}^n$ and if $\{e^i\}_{i=1}^n$ denotes the standard dual basis, then there is a standard notation for the set of coefficients found in the summation for b. In particular, we denote $B = [b]$ where $B_{ij} = b(e_i, e_j)$ hence, following Equation 7.1,

$$b(x,y) = \sum_{i,j=1}^{n} x^i y^j b(e_i, e_j) = \sum_{i=1}^{n}\sum_{j=1}^{n} x^i B_{ij} y^j = x^T B y$$

Definition 7.2.5.

Suppose $b : V \times V \to \mathbb{R}$ is a bilinear mapping then we say:

1. b is **symmetric** iff $b(x,y) = b(y,x)$ for all $x, y \in V$
2. b is **antisymmetric** iff $b(x,y) = -b(y,x)$ for all $x, y \in V$

Any bilinear mapping on V can be written as the sum of a symmetric and antisymmetric bilinear mapping, this claim follows easily from the calculation below:

$$b(x,y) = \underbrace{\frac{1}{2}\Big(b(x,y) + b(y,x)\Big)}_{symmetric} + \underbrace{\frac{1}{2}\Big(b(x,y) - b(y,x)\Big)}_{antisymmetric}.$$

We say S_{ij} is **symmetric** in i, j iff $S_{ij} = S_{ji}$ for all i, j. Likewise, we say A_{ij} is **antisymmetric** in i, j iff $A_{ij} = -A_{ji}$ for all i, j. If S is a symmetric bilinear mapping and A is an antisymmetric bilinear mapping then the components of S are symmetric and the components of A are antisymmetric. Why? Simply note:

$$S(e_i, e_j) = S(e_j, e_i) \quad \Rightarrow \quad S_{ij} = S_{ji}$$

and

$$A(e_i, e_j) = -A(e_j, e_i) \quad \Rightarrow \quad A_{ij} = -A_{ji}.$$

You can prove that the sum or scalar multiple of an (anti)symmetric bilinear mapping is once more (anti)symmetric therefore the set of antisymmetric bilinear maps $\Lambda^2(V)$ and the set of symmetric bilinear maps $ST^0_2 V$ are subspaces of $T^0_2 V$. The notation $\Lambda^2(V)$ is part of a larger discussion on the wedge product, we will return to it in a later section.

Finally, if we consider the special case of $V = \mathbb{R}^n$ once more we find that a bilinear mapping $b : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ has a symmetric matrix $[b]^T = [b]$ iff b is symmetric whereas it has an antisymmetric matric $[b]^T = -[b]$ iff b is antisymmetric.

bilinear maps on $V^* \times V^*$

Suppose $h : V^* \times V^* \to \mathbb{R}$ is bilinear then we say $h \in T^2_0 V$. In addition, suppose $\beta = \{e_1, e_2, \dots, e_n\}$ is a basis for V whereas $\beta^* = \{e^1, e^2, \dots, e^n\}$ is a basis of V^* with $e^j(e_i) = \delta_{ij}$. Let $\alpha, \beta \in V^*$

$$\begin{aligned} h(\alpha,\beta) &= h\Big(\sum_{i=1}^n \alpha_i e^i,\ \sum_{j=1}^n \beta_j e^j \Big) \qquad (7.2) \\ &= \sum_{i,j=1}^n h(\alpha_i e^i, \beta_j e^j) \\ &= \sum_{i,j=1}^n \alpha_i \beta_j h(e^i, e^j) \\ &= \sum_{i,j=1}^n h(e^i, e^j)\alpha(e_i)\beta(e_j) \end{aligned}$$

Therefore, if we define $h^{ij} = h(e^i, e^j)$ then we find the nice formula $h(\alpha,\beta) = \sum_{i,j=1}^n h^{ij}\alpha_i\beta_j$. To further refine the formula above we need a new concept.

The dual of the dual is called the double-dual and it is denoted V^{**}. For a finite dimensional vector space there is a cannonical isomorphism of V and V^{**}. In particular, $\Phi : V \to V^{**}$ is defined by $\Phi(v)(\alpha) = \alpha(v)$ for all $\alpha \in V^*$. It is customary to replace V with V^{**} wherever the context allows. For example, to define the tensor product of two vectors $x, y \in V$ as follows:

Definition 7.2.6.

Suppose V is a vector space with dual space V^*. We define the tensor product of vectors x, y as the mapping $x \otimes y : V^* \times V^* \to \mathbb{R}$ by $(x \otimes y)(\alpha, \beta) = \alpha(x)\beta(y)$ for all $x, y \in V$.

We could just as well have defined $x \otimes y = \Phi(x) \otimes \Phi(y)$ where Φ is once more the cannonical isomorphism of V and V^{**}. It's called *cannonical* because it has no particular dependendence on the coordinates used on V. In contrast, the isomorphism of $\mathbb{R}^n$ and $(\mathbb{R}^n)^*$ was built around the dot-product and the standard basis.

All of this said, note that $\alpha(e_i)\beta(e_j) = e_i \otimes e_j(\alpha, \beta)$ thus,

$$h(\alpha,\beta) = \sum_{i,j=1}^n h(e^i, e^j) e_i \otimes e_j(\alpha,\beta) \quad \Rightarrow \quad \boxed{h = \sum_{i,j=1}^n h(e^i, e^j) e_i \otimes e_j}$$

We argue that $\{e_i \otimes e_j\}_{i,j=1}^n$ is a basis[8]

Definition 7.2.7.

Suppose $h : V^* \times V^* \to \mathbb{R}$ is a bilinear mapping then we say:

1. h is **symmetric** iff $h(\alpha,\beta) = h(\beta,\alpha)$ for all $\alpha, \beta \in V^*$
2. h is **antisymmetric** iff $h(\alpha,\beta) = -h(\beta,\alpha)$ for all $\alpha, \beta \in V^*$

The discussion of the preceding subsection transfers to this context, we simply have to switch some vectors to dual vectors and move some indices up or down. I leave this to the reader.

[8] $T^2_0 V$ is a vector space and we've shown $T^2_0(V) \subseteq span\{e_i \otimes e_j\}_{i,j=1}^n$ but we should also show $e_i \otimes e_j \in T^2_0$ and check for LI of $\{e_i \otimes e_j\}_{i,j=1}^n$.

bilinear maps on $V \times V^*$

Suppose $H : V \times V^* \to \mathbb{R}$ is bilinear, we say $H \in T^1_1 V$ (or, if the context demands this detail $H \in T_1{}^1 V$). We define $\alpha \otimes x \in T_1{}^1(V)$ by the natural rule; $(\alpha \otimes x)(y, \beta) = \alpha(x)\beta(x)$ for all $(y, \beta) \in V \times V^*$. We find, by calculations similar to those already given in this section,

$$\boxed{H(y,\beta) = \sum_{i,j=1}^{n} H_i{}^j y^i \beta_j \qquad \text{and} \qquad H = \sum_{i,j=1}^{n} H_i{}^j e^i \otimes e_j}$$

where we defined $H_i{}^j = H(e_i, e^j)$.

bilinear maps on $V^* \times V$

Suppose $G : V^* \times V \to \mathbb{R}$ is bilinear, we say $G \in T^1_1 V$ (or, if the context demands this detail $G \in T^1{}_1 V$). We define $x \otimes \alpha \in T^1{}_1 V$ by the natural rule; $(x \otimes \alpha)(\beta, y) = \beta(x)\alpha(y)$ for all $(\beta, y) \in V^* \times V$. We find, by calculations similar to those already given in this section,

$$\boxed{G(\beta, y) = \sum_{i,j=1}^{n} G^i{}_j \beta_i y^j \qquad \text{and} \qquad G = \sum_{i,j=1}^{n} G^i{}_j e_i \otimes e^j}$$

where we defined $G^i{}_j = G(e^i, e_j)$.

7.2.2 trilinear maps

Definition 7.2.8.

Suppose V_1, V_2, V_3 are vector spaces then $T : V_1 \times V_2 \times V_3 \to \mathbb{R}$ is a **trilinear mapping** on $V_1 \times V_2 \times V_3$ iff for all $u, v \in V_1$, $w, x \in V_2$. $y, z \in V_3$ and $c \in \mathbb{R}$:

$$\begin{array}{llcll} (1.) & T(cu + v, w, y) & = & cT(u, w, y) + T(v, w, y) & \text{(linearity in the first slot)} \\ (2.) & T(u, cw + x, y) & = & cT(u, w, y) + T(u, x, y) & \text{(linearity in the second slot).} \\ (3.) & T(u, w, cy + z) & = & cT(u, w, y) + T(u, w, z) & \text{(linearity in the third slot).} \end{array}$$

If $T : V \times V \times V \to \mathbb{R}$ is trilinear on $V \times V \times V$ then we say T **is a trilinear mapping on** V and we denote the set of all such mappings $T^0_3 V$. The tensor product of three dual vectors is defined much in the same way as it was for two,

$$(\alpha \otimes \beta \otimes \gamma)(x, y, z) = \alpha(x)\beta(y)\gamma(z)$$

Let $\{e_i\}_{i=1}^n$ is a basis for V with dual basis $\{e^i\}_{i=1}^n$ for V^*. If T is trilinear on V it follows

$$T(x, y, z) = \sum_{i,j,k=1}^{n} T_{ijk} x^i y^j z^k \qquad \text{and} \qquad T = \sum_{i,j,k=1}^{n} T_{ijk} e^i \otimes e^j \otimes e^k$$

where we defined $T_{ijk} = T(e_i, e_j, e_k)$ for all $i, j, k \in \mathbb{N}_n$.

Generally suppose that V_1, V_2, V_3 are possibly distinct vector spaces. Moreover, suppose V_1 has basis $\{e_i\}_{i=1}^{n_1}$, V_2 has basis $\{f_j\}_{j=1}^{n_2}$ and V_3 has basis $\{g_k\}_{k=1}^{n_3}$. Denote the dual bases for V_1^*, V_2^*, V_3^* in the usual fashion: $\{e^i\}_{i=1}^{n_1}$, $\{f^j\}_{j=1}^{n_1}$, $\{g^k\}_{k=1}^{n_1}$. With this notation, we can write a trilinear mapping on $V_1 \times V_2 \times V_3$ as follows: (where we define $T_{ijk} = T(e_i, f_j, g_k)$)

$$T(x,y,z) = \sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\sum_{k=1}^{n_3} T_{ijk}x^iy^jz^k \qquad \text{and} \qquad T = \sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\sum_{k=1}^{n_3} T_{ijk}e^i \otimes f^j \otimes g^k$$

However, if V_1, V_2, V_3 happen to be related by duality then it is customary to use up/down indices. For example, if $T : V \times V \times V^* \to \mathbb{R}$ is trilinear then we write[9]

$$T = \sum_{i,j,k=1}^{n} T_{ij}{}^{k} e^i \otimes e^j \otimes e_k$$

and say $T \in T_2{}^1V$. On the other hand, if $S : V^* \times V^* \times V$ is trilinear then we'd write

$$T = \sum_{i,j,k=1}^{n} S^{ij}{}_{k} e_i \otimes e_j \otimes e^k$$

and say $T \in T^2{}_1V$. I'm not sure that I've ever seen this notation elsewhere, but perhaps it could be useful to denote the set of trinlinear maps $T : V \times V^* \times V \to \mathbb{R}$ as $T_1{}^1{}_1V$. Hopefully we will not need such silly notation in what we consider this semester.

There was a natural correspondance between bilinear maps on $\mathbb{R}^n$ and square matrices. For a trilinear map we would need a three-dimensional array of components. In some sense you could picture $T : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ as multiplication by a cube of numbers. Don't think too hard about these silly comments, we actually already wrote the useful formulae for dealing with trilinear objects. Let's stop to look at an example.

Example 7.2.9. *Define $T : \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$ by $T(x,y,z) = det(x|y|z)$. You may not have learned this in your linear algebra course[10] but a nice formula[11] for the determinant is given by the Levi-Civita symbol,*

$$det(A) = \sum_{i,j,k=1}^{3} \epsilon_{ijk}A_{i1}A_{j2}A_{k3}$$

note that $col_1(A) = [A_{i1}], col_2(A) = [A_{i2}]$ and $col_3(A) = [A_{i3}]$. It follows that

$$T(x,y,z) = \sum_{i,j,k=1}^{3} \epsilon_{ijk}x^iy^jz^k$$

[9] we identify e_k with its double-dual hence this tensor product is already defined, but to be safe let me write it out in this context $e^i \otimes e^j \otimes e_k(x,y,\alpha) = e^i(x)e^j(y)\alpha(e_k)$.

[10] maybe you haven't even taken linear yet!

[11] actually, I take this as the definition in linear algebra, it does take considerable effort to recover the expansion by minors formula which I use for concrete examples

Multilinearity follows easily from this formula. For example, linearity in the third slot:

$$\begin{aligned} T(x,y,cz+w) &= \sum_{i,j,k=1}^{3} \epsilon_{ijk}x^iy^j(cz+w)^k && (7.3)\\ &= \sum_{i,j,k=1}^{3} \epsilon_{ijk}x^iy^j(cz^k+w^k) && (7.4)\\ &= c\sum_{i,j,k=1}^{3} \epsilon_{ijk}x^iy^jz^k + \sum_{i,j,k=1}^{3} \epsilon_{ijk}x^iy^jw^k && (7.5)\\ &= cT(x,y,z)+T(x,y,w). && (7.6) \end{aligned}$$

Observe that by properties of determinants, or the Levi-Civita symbol if you prefer, swapping a pair of inputs generates a minus sign, hence:

$$T(x,y,z) = -T(y,x,z) = T(y,z,x) = -T(z,y,x) = T(z,x,y) = -T(x,z,y).$$

If $T : V \times V \times V \to \mathbb{R}$ is a trilinear mapping such that

$$T(x,y,z) = -T(y,x,z) = T(y,z,x) = -T(z,y,x) = T(z,x,y) = -T(x,z,y)$$

for all $x,y,z \in V$ then we say T **is antisymmetric**. Likewise, if $S : V \times V \times V \to \mathbb{R}$ is a trilinear mapping such that

$$S(x,y,z) = -S(y,x,z) = S(y,z,x) = -S(z,y,x) = S(z,x,y) = -S(x,z,y).$$

for all $x,y,z \in V$ then we say T **is symmetric**. Clearly the mapping defined by the determinant is antisymmetric. In fact, many authors define the determinant of an $n \times n$ matrix as the antisymmetric n-linear mapping which sends the identity matrix to 1. It turns out these criteria unquely define the determinant. That is the motivation behind my Levi-Civita symbol definition. That formula is just the nuts and bolts of complete antisymmetry.

You might wonder, can every trilinear mapping can be written as a the sum of a symmetric and antisymmetric mapping? The answer is no. Consider $T : V \times V \times V \to \mathbb{R}$ defined by $T = e^1 \otimes e^2 \otimes e^3$. Is it possible to find constants a, b such that:

$$e^1 \otimes e^2 \otimes e^3 = ae^{[1} \otimes e^2 \otimes e^{3]} + be^{(1} \otimes e^2 \otimes e^{3)}$$

where $[\dots]$ denotes complete antisymmetrization of $1,2,3$ and $(\dots)$ complete symmetrization:

$$e^{[1} \otimes e^2 \otimes e^{3]} = \frac{1}{6}\left[e^{123} + e^{231} + e^{312} - e^{321} - e^{213} - e^{132}\right]$$

For the symmetrization we also have to include all possible permutations of $(1,2,3)$ but all with $+$:

$$e^{(1} \otimes e^2 \otimes e^{3)} = \frac{1}{6}\left[e^{123} + e^{231} + e^{312} + e^{321} + e^{213} + e^{132}\right]$$

As you can see:

$$ae^{[1} \otimes e^2 \otimes e^{3]} + be^{(1} \otimes e^2 \otimes e^{3)} = \frac{a+b}{6}(e^{123} + e^{231} + e^{312}) + \frac{b-a}{6}(e^{321} + e^{213} + e^{132})$$

There is no way for these to give back only $e^1 \otimes e^2 \otimes e^3$. I leave it to the reader to fill the gaps in this argument. Generally, the decomposition of a multilinear mapping into more basic types is a problem which requires much more thought than we intend here. Representation theory does address this problem: how can we decompose a tensor product into irreducible pieces. Their idea of tensor product is not precisely the same as ours, however algebraically the problems are quite intertwined. I'll leave it at that unless you'd like to do an independent study on representation theory. Ideally you'd already have linear algebra and abstract algebra complete before you attempt that study.

7.2.3 multilinear maps

Definition 7.2.10.

Suppose $V_1, V_2, \dots V_k$ are vector spaces then $T : V_1 \times V_2 \times \cdots \times V_k \to \mathbb{R}$ is a k-**multilinear mapping** on $V_1 \times V_2 \times \cdots \times V_k$ iff for each $c \in \mathbb{R}$ and $x_1, y_1 \in V_1$, $x_2, y_2 \in V_2, \dots, x_k, y_k \in V_k$

$$T(x_1, \dots, cx_j + y_j, \dots, x_k) = cT(x_1, \dots, x_j, \dots, x_k) + T(x_1, \dots, y_j, \dots, x_k)$$

for $j = 1, 2, \dots, k$. In other words, we assume T is linear in each of its k-slots. If T is multilinear on $V^r \times (V^*)^s$ then we say that $T \in T^s_r V$ and we say T **is a type** (r, s) **tensor on** V.

The definition above makes a dual vector a type $(1, 0)$ tensor whereas a double dual of a vector a type $(0, 1)$ tensor, a bilinear mapping on V is a type $(2, 0)$ tensor and a bilinear mapping on V^* is a type $(0, 2)$ tensor with respect to V.

We are free to define tensor products in this context in the same manner as we have previously. Suppose $\alpha_1 \in V_1^*, \alpha_2 \in V_2^*, \dots, \alpha_k \in V_k^*$ and $v_1 \in V_1, v_2 \in V_2, \dots, v_k \in V_k$ then

$$\alpha_1 \otimes \alpha_2 \otimes \cdots \otimes \alpha_k(v_1, v_2, \dots, v_k) = \alpha_1(v_1)\alpha_2(v_2) \cdots \alpha_k(v_k)$$

It is easy to show the tensor produce of k-dual vectors as defined above is indeed a k-multilinear mapping. Moreover, the set of all k-multilinear mappings on $V_1 \times V_2 \times \cdots \times V_k$ clearly forms a vector space of dimension $dim(V_1)dim(V_2) \cdots dim(V_k)$ since it naturally takes the tensor product of the dual bases for $V_1^*, V_2^*, \dots, V_k^*$ as its basis. In particular, suppose for $j = 1, 2, \dots, k$ that V_j has basis $\{E_{ji}\}_{i=1}^{n_j}$ which is dual to $\{E_j^i\}_{i=1}^{n_j}$ the basis for V_j^*. Then we can derive that a k-multilinear mapping can be written as

$$T = \sum_{i_1=1}^{n_1} \sum_{i_2=1}^{n_2} \cdots \sum_{i_k=1}^{n_k} T_{i_1 i_2 \dots i_k} E_1^{i_1} \otimes E_2^{i_2} \otimes E_k^{i_k}$$

If T is a type (r, s) tensor on V then there is no need for the ugly double indexing on the basis since we need only tensor a basis $\{e_i\}_{i=1}^n$ for V and its dual $\{e^i\}_{i=1}^n$ for V^* in what follows:

$$T = \sum_{i_1, \dots, i_r=1}^{n} \sum_{j_1, \dots, j_s=1}^{n} T_{i_1 i_2 \dots i_r}^{j_1 j_2 \dots j_s} e^{i_1} \otimes e^{i_2} \otimes \cdots \otimes e^{i_r} \otimes e_{j_1} \otimes e_{j_2} \otimes \cdots \otimes e_{j_s}.$$

permutations

Before I define symmetric and antisymmetric for k-linear mappings on V I think it is best to discuss briefly some ideas from the theory of permutations.

Definition 7.2.11.

A permutation on $\{1, 2, \dots p\}$ is a bijection on $\{1, 2, \dots p\}$. We define the set of permutations on $\{1, 2, \dots p\}$ to be Σ_p. Further, define the sign of a permutation to be $sgn(\sigma) = 1$ if σ is the product of an even number of transpositions whereas $sgn(\sigma) = -1$ if σ is the product of a odd number transpositions.

Let us consider the set of permutations on $\{1, 2, 3, \dots n\}$, this is called S_n the symmetric group, its order is $n!$ if you were wondering. Let me remind[12] you how the cycle notation works since it allows us to explicitly present the number of transpositions contained in a permutation,

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 1 & 5 & 4 & 6 & 3 \end{pmatrix} \quad \Longleftrightarrow \quad \sigma = (12)(356) = (12)(36)(35) \tag{7.7}$$

recall the cycle notation is to be read right to left. If we think about inputing 5 we can read from the matrix notation that we ought to find $5 \mapsto 6$. Clearly that is the case for the first version of σ written in cycle notation; (356) indicates that $5 \mapsto 6$ and nothing else messes with 6 after that. Then consider feeding 5 into the version of σ written with just two-cycles (a.k.a. transpositions), first we note (35) indicates $5 \mapsto 3$, then that 3 hits (36) which means $3 \mapsto 6$, finally the cycle (12) doesn't care about 6 so we again have that $\sigma(5) = 6$. Finally we note that $sgn(\sigma) = -1$ since it is made of 3 transpositions.

It is always possible to write any permutation as a product of transpositions, such a decomposition is not unique. However, if the number of transpositions is even then it will remain so no matter how we rewrite the permutation. Likewise if the permutation is an product of an odd number of transpositions then any other decomposition into transpositions is also comprised of an odd number of transpositions. This is why we can **define** an **even** permutation is a permutation comprised by an even number of transpositions and an **odd** permutation is one comprised of an odd number of transpositions.

Example 7.2.12. Sample cycle calculations: *we rewrite as product of transpositions to determin if the given permutation is even or odd,*

$$\sigma = (12)(134)(152) = (12)(14)(13)(12)(15) \implies sgn(\sigma) = -1$$

$$\lambda = (1243)(3521) = (13)(14)(12)(31)(32)(35) \implies sgn(\lambda) = 1$$

$$\gamma = (123)(45678) = (13)(12)(48)(47)(46)(45) \implies sgn(\gamma) = 1$$

We will not actually write down permutations in the calculations the follow this part of the notes. I merely include this material as to give a logically complete account of antisymmetry. In practice, if you understood the terms as the apply to the bilinear and trilinear case it will usually suffice for concrete examples. Now we are ready to define symmetric and antisymmetric.

[12] or perhaps, more likely, introduce you to this notation

Definition 7.2.13.

A k-linear mapping $L : V \times V \times \cdots \times V \to \mathbb{R}$ is **completely symmetric** if

$$L(x_1, \dots, x, \dots, y, \dots, x_k) = L(x_1, \dots, y, \dots, x, \dots, x_k)$$

for all possible $x, y \in V$. Conversely, if a k-linear mapping on V has

$$L(x_1, \dots, x, \dots, y, \dots, x_p) = -L(x_1, \dots, y, \dots, x, \dots, x_p)$$

for all possible pairs $x, y \in V$ then it is said to be **completely antisymmetric or alternating**. Equivalently a k-linear mapping L is alternating if for all $\pi \in \Sigma_k$

$$L(x_{\pi_1}, x_{\pi_2}, \dots, x_{\pi_k}) = sgn(\pi) L(x_1, x_2, \dots, x_k)$$

The set of alternating multilinear mappings on V is denoted ΛV, the set of k-linear alternating maps on V is denoted $\Lambda^k V$. Often an alternating k-linear map is called a **k-form**. Moreover, we say the **degree** of a k-form is k.

Similar terminology applies to the components of tensors. We say $T_{i_1 i_2 \dots i_k}$ is completely symmetric in $i_1, i_2, \dots, i_k$ iff $T_{i_1 i_2 \dots i_k} = T_{i_{\sigma(1)} i_{\sigma(2)} \cdots i_{\sigma(k)}}$ for all $\sigma \in \Sigma_k$. On the other hand, $T_{i_1 i_2 \dots i_k}$ is completely antisymmetric in $i_1, i_2, \dots, i_k$ iff $T_{i_1 i_2 \dots i_k} = sgn(\sigma) T_{i_{\sigma(1)} i_{\sigma(2)} \cdots i_{\sigma(k)}}$ for all $\sigma \in \Sigma_k$. It is a simple exercise to show that a completely (anti)symmetric tensor[13] has completely (anti)symmetric components.

The tensor product is an interesting construction to discuss at length. To summarize, it is associative and distributive across addition. Scalars factor out and it is not generally commutative. For a given vector space V we can in principle generate by tensor products multilinear mappings of arbitrarily high order. This tensor algebra is infinite dimensional. In contrast, the space ΛV of forms on V is a finite-dimensional subspace of the tensor algebra. We discuss this next.

7.3 wedge product

We assume V is a vector space with basis $\{e_i\}_{i=1}^n$ throughout this section. The dual basis is denoted $\{e^i\}_{i=1}^n$ as is our usual custom. Our goal is to find a basis for the alternating maps on V and explore the structure implicit within its construction. This will lead us to call ΛV the **exterior algebra** of V after the discussion below is complete.

7.3.1 wedge product of dual basis generates basis for ΛV

Suppose $b : V \times V \to \mathbb{R}$ is antisymmetric and $b = \sum_{i.j=1}^n b_{ij} e^i \otimes e^j$, it follows that $b_{ij} = -b_{ji}$ for all $i, j \in \mathbb{N}_n$. Notice this implies that $b_{ii} = 0$ for $i = 1, 2, \dots, n$. For a given pair of indices i, j either

[13] in this context a tensor is simply a multilinear mapping, in physics there is more attached to the term

$i < j$ or $j < i$ or $i = j$ hence,

$$\begin{aligned} b &= \sum_{i<j} b_{ij} e^i \otimes e^j + \sum_{j<i} b_{ij} e^i \otimes e^j + \sum_{i=j} b_{ij} e^i \otimes e^j \\ &= \sum_{i<j} b_{ij} e^i \otimes e^j + \sum_{j<i} b_{ij} e^i \otimes e^j \\ &= \sum_{i<j} b_{ij} e^i \otimes e^j - \sum_{j<i} b_{ji} e^i \otimes e^j \\ &= \sum_{k<l} b_{kl} e^k \otimes e^l - \sum_{k<l} b_{kl} e^l \otimes e^k \\ &= \sum_{k<l} b_{kl} (e^k \otimes e^l - e^l \otimes e^k). \end{aligned} \tag{7.8}$$

Therefore, $\{e^k \otimes e^l - e^l \otimes e^k | l, k \in \mathbb{N}_n \text{ and } l < k\}$ spans the set of antisymmetric bilinear maps on V. Moreover, you can show this set is linearly independent hence it is a basis fo $\Lambda^2 V$. We define the wedge product of $e^k \wedge e^l = e^k \otimes e^l - e^l \otimes e^k$. With this notation we find that the alternating bilinear form b can be written as

$$\boxed{b = \sum_{k<l} b_{kl} e^k \wedge e^l = \sum_{i,j=1}^{n} \frac{1}{2} b_{ij} e^i \wedge e^j}$$

where the summation on the r.h.s. is over all indices[14]. Notice that $e^i \wedge e^j$ is an antisymmetric bilinear mapping because $e^i \wedge e^j (x, y) = -e^i \wedge e^j (y, x)$, however, there is more structure here than just that. It is also true that $e^i \wedge e^j = -e^j \wedge e^i$. This is a conceptually different antisymmetry, it is the antisymmetry of the wedge produce $\wedge$.

Suppose $b : V \times V \times V \to \mathbb{R}$ is antisymmetric and $b = \sum_{i,j,k=1}^{n} b_{ijk} e^i \otimes e^j \otimes e^k$, it follows that $b_{ijk} = b_{jki} = b_{kij}$ and $b_{ijk} = -b_{kji} = -b_{jik} = b_{ikj}$ for all $i, j, k \in \mathbb{N}_n$. Notice this implies that $b_{iii} = 0$ for $i = 1, 2, \dots, n$. A calculation similar to the one just offered for the case of a bilinear map reveals that we can write b as follows:

$$\begin{aligned} b = \sum_{i<j<k} b_{ijk} \Big(& e^i \otimes e^j \otimes e^k + e^j \otimes e^k \otimes e^i + e^k \otimes e^i \otimes e^j \\ & - e^k \otimes e^j \otimes e^i - e^j \otimes e^i \otimes e^k - e^i \otimes e^k \otimes e^j \Big) \end{aligned} \tag{7.9}$$

Define $e^i \wedge e^j \wedge e^k = e^i \otimes e^j \otimes e^k + e^j \otimes e^k \otimes e^i + e^k \otimes e^i \otimes e^j - e^k \otimes e^j \otimes e^i - e^j \otimes e^i \otimes e^k - e^i \otimes e^k \otimes e^j$ thus

$$\boxed{b = \sum_{i<j<k} b_{ijk} e^i \wedge e^j \wedge e^k = \sum_{i,j,k=1}^{n} \frac{1}{3!} b_{ijk} e^i \wedge e^j \wedge e^k} \tag{7.10}$$

and it is clear that $\{e^i \wedge e^j \wedge e^k \mid i, j, k \in \mathbb{N}_n \text{ and } i < j < k\}$ forms a basis for the set of alternating trilinear maps on V.

Following the patterns above, we define the wedge product of p dual basis vectors,

$$e^{i_1} \wedge e^{i_2} \wedge \cdots \wedge e^{i_p} = \sum\nolimits_{\pi \in \Sigma_p} sgn(\pi) e^{i_{\pi(1)}} \otimes e^{i_{\pi(2)}} \otimes \cdots \otimes e^{i_{\pi(p)}} \tag{7.11}$$

[14] yes there is something to work out here, probably in your homework

If $x, y \in V$ we would like to show that

$$e^{i_1} \wedge e^{i_2} \wedge \cdots \wedge e^{i_p}(\dots, x, \dots, y, \dots) = -e^{i_1} \wedge e^{i_2} \wedge \cdots \wedge e^{i_p}(\dots, y, \dots, x, \dots) \tag{7.12}$$

follows from the complete antisymmetrization in the definition of the wedge product. Before we give the general argument, let's see how this works in the trilinear case. Consider, $e^i \wedge e^j \wedge e^k =$

$$= e^i \otimes e^j \otimes e^k + e^j \otimes e^k \otimes e^i + e^k \otimes e^i \otimes e^j - e^k \otimes e^j \otimes e^i - e^j \otimes e^i \otimes e^k - e^i \otimes e^k \otimes e^j.$$

Calculate, noting that $e^i \otimes e^j \otimes e^k(x, y, z) = e^i(x)e^j(y)e^k(z) = x^i y^j z^k$ hence

$$e^i \wedge e^j \wedge e^k(x, y, z) = x^i y^j z^k + x^j y^k z^i + x^k y^i z^j - x^k y^j z^i - x^j y^i z^k - x^i y^k z^j$$

Thus,

$$e^i \wedge e^j \wedge e^k(x, z, y) = x^i z^j y^k + x^j z^k y^i + x^k z^i y^j - x^k z^j y^i - x^j z^i y^k - x^i z^k y^j$$

and you can check that $e^i \wedge e^j \wedge e^k(x, y, z) = -e^i \wedge e^j \wedge e^k(x, z, y)$. Similar tedious calculations prove antisymmetry of the the interchange of the first and second or the first and third slots. Therefore, $e^i \wedge e^j \wedge e^k$ is an alternating trilinear map as it is clearly trilinear since it is built from the sum of tensor products which we know are likewise trilinear.

The multilinear case follows essentially the same argument, note

$$e^{i_1} \wedge e^{i_2} \wedge \cdots \wedge e^{i_p}(\dots, x_j, \dots, x_k, \dots) = \sum\nolimits_{\pi \in \Sigma_p} sgn(\pi) x_1^{i_{\pi(1)}} \cdots x_j^{i_{\pi(j)}} \cdots x_k^{i_{\pi(k)}} \cdots x_p^{i_{\pi(p)}} \tag{7.13}$$

whereas,

$$e^{i_1} \wedge e^{i_2} \wedge \cdots \wedge e^{i_p}(\dots, x_k, \dots, x_j, \dots) = \sum\nolimits_{\sigma \in \Sigma_p} sgn(\sigma) x_1^{i_{\sigma(1)}} \cdots x_k^{i_{\sigma(k)}} \cdots x_j^{i_{\sigma(j)}} \cdots x_p^{i_{\sigma(p)}}. \tag{7.14}$$

Suppose we take each permutation σ and subsitute $\delta \in \Sigma_p$ such that $\sigma(j) = \delta(k)$ and $\sigma(k) = \delta(j)$ and otherwise δ and σ agree. In cycle notation, $\delta(jk) = \sigma$. Substitution δ into Equation 7.14:

$$\begin{aligned} &e^{i_1} \wedge e^{i_2} \wedge \cdots \wedge e^{i_p}(\dots, x_k, \dots, x_j, \dots) \\ &\qquad = \sum\nolimits_{\delta \in \Sigma_p} sgn(\delta(jk)) x_1^{i_{\delta(1)}} \cdots x_k^{i_{\delta(j)}} \cdots x_j^{i_{\delta(k)}} \cdots x_p^{i_{\delta(p)}} \\ &\qquad = -\sum\nolimits_{\delta \in \Sigma_p} sgn(\delta) x_1^{i_{\delta(1)}} \cdots x_j^{i_{\delta(k)}} \cdots x_k^{i_{\delta(j)}} \cdots x_p^{i_{\delta(p)}} \\ &\qquad = -e^{i_1} \wedge e^{i_2} \wedge \cdots \wedge e^{i_p}(\dots, x_j, \dots, x_k, \dots) \end{aligned} \tag{7.15}$$

Here the sgn of a permutation σ is $(-1)^N$ where N is the number of cycles in σ. We observed that $\delta(jk)$ has one more cycle than δ hence $sgn(\delta(jk)) = -sgn(\delta)$. Therefore, we have shown that $e^{i_1} \wedge e^{i_2} \wedge \cdots \wedge e^{i_p} \in \Lambda^p V$.

Recall that $e^i \wedge e^j = -e^j \wedge e^i$ in the $p = 2$ case. There is a generalization of that result to the $p > 2$ case. In words, the wedge product is antisymetric with respect the interchange of any two dual vectors. For $p = 3$ we have the following identities for the wedge product:

$$e^i \wedge e^j \wedge e^k = -\underbrace{e^j \wedge e^i}_{swapped} \wedge e^k = e^j \wedge \underbrace{e^k \wedge e^i}_{swapped} = -\underbrace{e^k \wedge e^j}_{swapped} \wedge e^i = e^k \wedge \underbrace{e^i \wedge e^j}_{swapped} = -\underbrace{e^i \wedge e^k}_{swapped} \wedge e^j$$

I've indicated how these signs are consistent with the $p = 2$ antisymmetry. Any permutation of the dual vectors can be thought of as a combination of several transpositions. In any event, it is sometimes useful to just know that the wedge product of three elements is invariant under **cyclic** permutations of the dual vectors,

$$e^i \wedge e^j \wedge e^k = e^j \wedge e^k \wedge e^i = e^k \wedge e^i \wedge e^j$$

and changes by a sign for **anticyclic** permutations of the given object,

$$e^i \wedge e^j \wedge e^k = -e^j \wedge e^i \wedge e^k = -e^k \wedge e^j \wedge e^i = -e^i \wedge e^k \wedge e^j$$

Generally we can argue that, for any permutation $\pi \in \Sigma_p$:

$$\boxed{e^{i_1} \wedge e^{i_2} \wedge \cdots \wedge e^{i_p} = sgn(\pi) e^{i_{\pi(1)}} \wedge e^{i_{\pi(2)}} \wedge \cdots \wedge e^{i_{\pi(p)}}}$$

This is just a slick formula which says the wedge product generates a minus whenever you flip two dual vectors which are wedged.

7.3.2 the exterior algebra

The careful reader will realize we have yet to define wedge products of anything except for the dual basis. But, naturally you must wonder if we can take the wedge product of other dual vectors or morer generally alternating tensors. The answer is yes. Let us define the general wedge product:

Definition 7.3.1. *Suppose $\alpha \in \Lambda^p V$ and $\beta \in \Lambda^q V$. We define $\mathcal{I}_p$ to be the set of all increasing lists of p-indices, this set can be empty if dim(V) is not sufficiently large. Moreover, if $I = (i_1, i_2, \dots, i_p)$ then introduce notation $e^I = e^{i_1} \wedge e^{i_2} \wedge \cdots \wedge e^{i_p}$ hence:*

$$\alpha = \sum_{i_1, i_2, \dots, i_p = 1}^{n} \frac{1}{p!} \alpha_{i_1 i_2 \dots i_p} e^{i_1} \wedge e^{i_2} \wedge \cdots \wedge e^{i_p} = \sum_I \frac{1}{p!} \alpha_I e^I = \sum_{I \in \mathcal{I}_p} \alpha_I e^I$$

and

$$\beta = \sum_{j_1, j_2, \dots, j_q = 1}^{n} \frac{1}{q!} \beta_{j_1 j_2 \dots j_q} e^{j_1} \wedge e^{j_2} \wedge \cdots \wedge e^{j_q} = \sum_J \frac{1}{q!} \beta_J e^J = \sum_{J \in \mathcal{I}_q} \beta_J e^J$$

Naturally, $e^I \wedge e^J = e^{i_1} \wedge e^{i_2} \wedge \cdots \wedge e^{i_p} \wedge e^{j_1} \wedge e^{j_2} \wedge \cdots \wedge e^{j_q}$ and we defined this carefully in the preceding subsection. Define $\alpha \wedge \beta \in \Lambda^{p+q} V$ as follows:

$$\alpha \wedge \beta = \sum_I \sum_J \frac{1}{p!q!} \alpha_I \beta_J e^I \wedge e^J.$$

Again, but with less slick notation:

$$\alpha \wedge \beta = \sum_{i_1, i_2, \dots, i_p = 1}^{n} \sum_{j_1, j_2, \dots, j_q = 1}^{n} \frac{1}{p!q!} \alpha_{i_1 i_2 \dots i_p} \beta_{j_1 j_2 \dots j_q} e^{i_1} \wedge e^{i_2} \wedge \cdots \wedge e^{i_p} \wedge e^{j_1} \wedge e^{j_2} \wedge \cdots \wedge e^{j_q}$$

All the definition above really says is that we extend the wedge product on the basis to distribute over the addition of dual vectors. What this means calculationally is that the wedge product obeys the usual laws of addition and scalar multiplication. The one feature that is perhaps foreign is the antisymmetry of the wedge product. We must take care to maintain the order of expressions since the wedge product is not generally commutative.

Proposition 7.3.2.

Let α, β, γ be forms on V and $c \in \mathbb{R}$ then

$$\begin{array}{lll} (i) & (\alpha+\beta)\wedge\gamma = \alpha\wedge\gamma + \beta\wedge\gamma & \text{distributes across vector addition} \\ (ii) & \alpha\wedge(\beta+\gamma) = \alpha\wedge\beta + \alpha\wedge\gamma & \text{distributes across vector addition} \\ (iii) & (c\alpha)\wedge\beta = \alpha\wedge(c\beta) = c\,(\alpha\wedge\beta) & \text{scalars factor out} \\ (iv) & \alpha\wedge(\beta\wedge\gamma) = (\alpha\wedge\beta)\wedge\gamma & \text{associativity} \end{array}$$

I leave the proof of this proposition to the reader.

Proposition 7.3.3. *graded commutivity of homogeneous forms.*

Let α, β be forms on V of degree p and q respectively then

$$\alpha \wedge \beta = -(-1)^{pq}\beta \wedge \alpha$$

Proof: suppose $\alpha = \sum_I \frac{1}{p!}e^I$ is a p-form on V and $\beta = \sum_J \frac{1}{q!}e^J$ is a q-form on V. Calculate:

$$\begin{aligned} \alpha\wedge\beta &= \sum_I\sum_J \frac{1}{p!q!}\alpha_I\beta_J e^I \wedge e^J && \text{by defn. of } \wedge, \\ &= \sum_I\sum_J \frac{1}{p!q!}\beta_J\alpha_I e^I \wedge e^J && \text{coefficients are scalars,} \\ &= (-1)^{pq}\sum_I\sum_J \frac{1}{p!q!}\beta_J\alpha_I e^J \wedge e^I && \text{(details on sign given below)} \\ &= (-1)^{pq}\beta\wedge\alpha \end{aligned}$$

Let's expand in detail why $e^J \wedge e^I = (-1)^{pq}e^I \wedge e^J$. Suppose $I = (i_1, i_2, \dots, i_p)$ and $J = (j_1, j_2, \dots, j_q)$, the key is that every swap of dual vectors generates a sign:

$$\begin{aligned} e^I \wedge e^J &= e^{i_1}\wedge e^{i_2}\wedge\cdots\wedge e^{i_p}\wedge e^{j_1}\wedge e^{j_2}\wedge\cdots\wedge e^{j_q} \\ &= (-1)^q e^{i_1}\wedge e^{i_2}\wedge\cdots\wedge e^{i_{p-1}}\wedge e^{j_1}\wedge e^{j_2}\wedge\cdots\wedge e^{j_q}\wedge e^{i_p} \\ &= (-1)^q(-1)^q e^{i_1}\wedge e^{i_2}\wedge\cdots\wedge e^{i_{p-2}}\wedge e^{j_1}\wedge e^{j_2}\wedge\cdots\wedge e^{j_q}\wedge e^{i_{p-1}}\wedge e^{i_p} \\ &\quad\vdots \qquad\quad \vdots \qquad\quad \vdots \\ &= \underbrace{(-1)^q(-1)^q\cdots(-1)^q}_{p-factors} e^{j_1}\wedge e^{j_2}\wedge\cdots\wedge e^{j_q}\wedge e^{i_1}\wedge e^{j_2}\wedge\cdots\wedge e^{i_p} \\ &= (-1)^{pq}e^J\wedge e^I. \end{aligned}$$

□

Example 7.3.4. *Let α be a 2-form defined by*

$$\alpha = ae^1\wedge e^2 + be^2\wedge e^3$$

And let β be a 1-form defined by

$$\beta = 3e^1$$

Consider then,

$$\begin{aligned} \alpha\wedge\beta &= (ae^1\wedge e^2 + be^2\wedge e^3)\wedge(3e^1) \\ &= (3ae^1\wedge e^2\wedge e^1 + 3be^2\wedge e^3\wedge e^1 \\ &= 3be^1\wedge e^2\wedge e^3. \end{aligned} \tag{7.16}$$

whereas,

$$\begin{aligned}\beta\wedge\alpha &= 3e^1\wedge(ae^1\wedge e^2+be^2\wedge e^3)\\ &= (3ae^1\wedge e^1\wedge e^2+3be^1\wedge e^2\wedge e^3\\ &= 3be^1\wedge e^2\wedge e^3.\end{aligned} \tag{7.17}$$

so this agrees with the proposition, $(-1)^{pq}=(-1)^2=1$ *so we should have found that* $\alpha\wedge\beta=\beta\wedge\alpha$*. This illustrates that although the wedge product is antisymmetric on the basis, it is not always antisymmetric, in particular it is commutative for even forms.*

The graded commutivity rule $\alpha\wedge\beta=-(-1)^{pq}\beta\wedge\alpha$ has some suprising implications. This rule is ultimately the reason ΛV is finite dimensional. Let's see how that happens.

Proposition 7.3.5. *linear dependent one-forms wedge to zero:*

If $\alpha,\beta\in V^*$ and $\alpha=c\beta$ for some $c\in\mathbb{R}$ then $\alpha\wedge\beta=0$.

Proof: to begin, note that $\beta\wedge\beta=-\beta\wedge\beta$ hence $2\beta\wedge\beta=0$ and it follows that $\beta\wedge\beta=0$. Note:

$$\alpha\wedge\beta=c\beta\wedge\beta=c(0)=0$$

therefore the proposition is true. □

Proposition 7.3.6.

Suppose that $\alpha_1,\alpha_2,\dots,\alpha_p$ are linearly dependent 1-forms then

$$\alpha_1\wedge\alpha_2\wedge\cdots\wedge\alpha_p=0.$$

Proof: by assumption of linear dependence there exist constants $c_1,c_2,\dots,c_p$ not all zero such that

$$c_1\alpha_1+c_2\alpha_2+\cdots c_p\alpha_p=0.$$

Suppose that c_k is a nonzero constant in the sum above, then we may divide by it and consequently we can write α_k in terms of all the other 1-forms,

$$\alpha_k=\frac{-1}{c_k}\Big(c_1\alpha_1+\cdots+c_{k-1}\alpha_{k-1}+c_{k+1}\alpha_{k+1}+\cdots+c_p\alpha_p\Big)$$

Insert this sum into the wedge product in question,

$$\begin{aligned}\alpha_1\wedge\alpha_2\wedge\ldots\wedge\alpha_p &= \alpha_1\wedge\alpha_2\wedge\cdots\wedge\alpha_k\wedge\cdots\wedge\alpha_p\\ &= (-c_1/c_k)\alpha_1\wedge\alpha_2\wedge\cdots\wedge\alpha_1\wedge\cdots\wedge\alpha_p\\ &\quad +(-c_2/c_k)\alpha_1\wedge\alpha_2\wedge\cdots\wedge\alpha_2\wedge\cdots\wedge\alpha_p+\cdots\\ &\quad +(-c_{k-1}/c_k)\alpha_1\wedge\alpha_2\wedge\cdots\wedge\alpha_{k-1}\wedge\cdots\wedge\alpha_p\\ &\quad +(-c_{k+1}/c_k)\alpha_1\wedge\alpha_2\wedge\cdots\wedge\alpha_{k+1}\wedge\cdots\wedge\alpha_p+\cdots\\ &\quad +(-c_p/c_k)\alpha_1\wedge\alpha_2\wedge\cdots\wedge\alpha_p\wedge\cdots\wedge\alpha_p\\ &= 0.\end{aligned} \tag{7.18}$$

We know all the wedge products are zero in the above because in each there is at least one 1-form repeated, we simply permute the wedge products till they are adjacent and by the previous proposition the term vanishes. The proposition follows. □

Let us pause to reflect on the meaning of the proposition above for a n-dimensional vector space V. The dual space V^* is likewise n-dimensional, this is a general result which applies to all finite-dimensional vector spaces[15]. Thus, any set of more than n dual vectors is necessarily linearly dependent. Consquently, using the proposition above, we find the wedge product of more than n one-forms is trivial. Therefore, while it is possible to construct $\Lambda^k V$ for $k > n$ we should understand that this space only contains zero. The highest degree of a nontrivial form over a vector space of dimension n is an n-form.

Moreover, we can use the proposition to deduce the dimension of a basis for $\Lambda^p V$, it must consist of the wedge product of distinct linearly independent one-forms. The number of ways to choose p distinct objects from a list of n distinct objects is precisely "n choose p",

$$\binom{n}{p} = \frac{n!}{(n-p)!p!} \qquad \text{for } 0 \leq p \leq n. \tag{7.19}$$

Proposition 7.3.7.

If V is an n-dimensional vector space then $\Lambda^k V$ is an $\binom{n}{p}$-dimensional vector space of p-forms. Moreover, the direct sum of all forms over V has the structure

$$\Lambda V = \mathbb{R} \oplus \Lambda^1 V \oplus \cdots \Lambda^{n-1} V \oplus \Lambda^n V$$

and is a vector space of dimension 2^n

Proof: define $\Lambda^0 V = \mathbb{R}$ then it is clear $\Lambda^k V$ forms a vector space for $k = 0, 1, \dots, n$. Moreover, $\Lambda^j V \cap \Lambda^k V = \{0\}$ for $j \neq k$ hence the term "direct sum" is appropriate. It remains to show $dim(\Lambda V) = 2^n$ where $dim(V) = n$. A natural basis β for ΛV is found from taking the union of the bases for each subspace of k-forms,

$$\beta = \{1, e^{i_1}, e^{i_1} \wedge e^{i_2}, \dots, e^{i_1} \wedge e^{i_2} \wedge \cdots \wedge e^{i_n} \mid 1 \leq i_1 < i_2 < \cdots < i_n \leq n\}$$

But, we can count the number of vectors N in the set above as follows:

$$N = 1 + n + \binom{n}{2} + \cdots + \binom{n}{n-1} + \binom{n}{n}$$

Recall the binomial theorem states

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k = a^n + na^{n-1}b + \cdots + nab^{n-1} + b^n.$$

Recognize that $N = (1+1)^n$ and the proposition follows. $\Box$

We should note that in the basis above the space of n-forms is one-dimensional because there is only one way to choose a strictly increasing list of n integers in $\mathbb{N}_n$. In particular, it is useful to note $\Lambda^n V = span\{e^1 \wedge e^2 \wedge \cdots \wedge e^n\}$. The form $e^1 \wedge e^2 \wedge \cdots \wedge e^n$ is sometimes called the the *top-form*[16].

[15] however, in infinite dimensions, the story is not so simple

[16] or *volume form* for reasons we will explain later, other authors begin the discussion of forms from the consideration of volume, see Chapter 4 in Bernard Schutz' *Geometrical methods of mathematical physics*

Example 7.3.8. exterior algebra of $\mathbb{R}^2$ *Let us begin with the standard dual basis $\{e^1, e^2\}$. By definition we take the $p = 0$ case to be the field itself; $\Lambda^0 V \equiv \mathbb{R}$, it has basis 1. Next, $\Lambda^1 V = span(e^1, e^2) = V^*$ and $\Lambda^2 V = span(e^1 \wedge e^2)$ is all we can do here. This makes ΛV a $2^2 = 4$-dimensional vector space with basis*

$$\{1, e^1, e^2, e^1 \wedge e^2\}.$$

Example 7.3.9. exterior algebra of $\mathbb{R}^3$ *Let us begin with the standard dual basis $\{e^1, e^2, e^3\}$. By definition we take the $p = 0$ case to be the field itself; $\Lambda^0 V \equiv \mathbb{R}$, it has basis 1. Next, $\Lambda^1 V = span(e^1, e^2, e^3) = V^*$. Now for something a little more interesting,*

$$\Lambda^2 V = span(e^1 \wedge e^2, e^1 \wedge e^3, e^2 \wedge e^3)$$

and finally,

$$\Lambda^3 V = span(e^1 \wedge e^2 \wedge e^3).$$

This makes ΛV a $2^3 = 8$-dimensional vector space with basis

$$\{1, e^1, e^2, e^3, e^1 \wedge e^2, e^1 \wedge e^3, e^2 \wedge e^3, e^1 \wedge e^2 \wedge e^3\}$$

it is curious that the number of independent one-forms and 2-forms are equal.

Example 7.3.10. exterior algebra of $\mathbb{R}^4$ *Let us begin with the standard dual basis $\{e^1, e^2, e^3, e^4\}$. By definition we take the $p = 0$ case to be the field itself; $\Lambda^0 V \equiv \mathbb{R}$, it has basis 1. Next, $\Lambda^1 V = span(e^1, e^2, e^3, e^4) = V^*$. Now for something a little more interesting,*

$$\Lambda^2 V = span(e^1 \wedge e^2, e^1 \wedge e^3, e^1 \wedge e^4, e^2 \wedge e^3, e^2 \wedge e^4, e^3 \wedge e^4)$$

and three forms,

$$\Lambda^3 V = span(e^1 \wedge e^2 \wedge e^3, e^1 \wedge e^2 \wedge e^4, e^1 \wedge e^3 \wedge e^4, e^2 \wedge e^3 \wedge e^4).$$

and $\Lambda^3 V = span(e^1 \wedge e^2 \wedge e^3)$. Thus ΛV a $2^4 = 16$-dimensional vector space. Note that, in contrast to $\mathbb{R}^3$, we do not have the same number of independent one-forms and two-forms over $\mathbb{R}^4$.

Let's explore how this algebra fits with calculations we already know about determinants.

Example 7.3.11. *Suppose $A = [A_1|A_2]$. I propose the determinant of A is given by the top-form on $\mathbb{R}^2$ via the formula $det(A) = (e^1 \wedge e^2)(A_1, A_2)$. Suppose $A = \left[\begin{array}{cc} a & b \\ c & d \end{array}\right]$ then $A_1 = (a, c)$ and $A_2 = (b, d)$. Thus,*

$$\begin{aligned} det\left[\begin{array}{cc} a & b \\ c & d \end{array}\right] &= (e^1 \wedge e^2)(A_1, A_2) \\ &= (e^1 \otimes e^2 - e^2 \otimes e^1)((a, c), (b, d)) \\ &= e^1(a, c)e^2(b, d) - e^2(a, c)e^1(b, d) \\ &= ad - bc. \end{aligned}$$

I hope this is not surprising!

Example 7.3.12. *Suppose $A = [A_1|A_2|A_3]$. I propose the determinant of A is given by the top-form on $\mathbb{R}^3$ via the formula $det(A) = (e^1 \wedge e^2 \wedge e^3)(A_1, A_2, A_3)$. Let's see if we can find the expansion by cofactors. By the definition we have $e^1 \wedge e^2 \wedge e^3 =$*

$$\begin{aligned} &= e^1 \otimes e^2 \otimes e^3 + e^2 \otimes e^3 \otimes e^1 + e^3 \otimes e^1 \otimes e^2 - e^3 \otimes e^2 \otimes e^1 - e^2 \otimes e^1 \otimes e^3 - e^1 \otimes e^3 \otimes e^2 \\ &= e^1 \otimes (e^2 \otimes e^3 - e^3 \otimes e^2) - e^2 \otimes (e^1 \otimes e^3 - e^3 \otimes e^1) + e^3 \otimes (e^1 \otimes e^2 - e^2 \otimes e^1) \\ &= e^1 \otimes (e^2 \wedge e^3) - e^2 \otimes (e^1 \wedge e^3) + e^3 \otimes (e^1 \wedge e^2). \end{aligned}$$

I submit to the reader that this is precisely the cofactor expansion formula with respect to the first column of A. Suppose $A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$ then $A_1 = (a, d, g)$, $A_2 = (b, e, h)$ and $A_3 = (c, f, i)$. Calculate,

$$\begin{aligned} det(A) &= e^1(A_1)(e^2 \wedge e^3)(A_2, A_3) - e^2(A_1)(e^1 \wedge e^3)(A_2, A_3) + e^3(A_1)(e^1 \wedge w^2)(A_2, A_3) \\ &= a(e^2 \wedge e^3)(A_2, A_3) - d(e^1 \wedge e^3)(A_2, A_3) + g(e^1 \wedge w^2)(A_2, A_3) \\ &= a(ei - fh) - d(bi - ch) + g(bf - ce) \end{aligned}$$

which is precisely my claim.

7.3.3 connecting vectors and forms in $\mathbb{R}^3$

There are a couple ways to connect vectors and forms in $\mathbb{R}^3$. Mainly we need the following maps:

Definition 7.3.13.

Given $v =< a, b, c > \in \mathbb{R}^3$ we can construct a corresponding one-form $\omega_v = ae^1 + be^2 + ce^3$ or we can construct a corresponding two-form $\Phi_v = ae^2 \wedge e^3 + be^3 \wedge e^1 + ce^1 \wedge e^2$

Recall that $dim(\Lambda^1\mathbb{R}^3) = dim(\Lambda^2\mathbb{R}^3) = 3$ hence the space of vectors, one-forms, and also two-forms are isomorphic as vector spaces. It is not difficult to show that $\omega_{v_1+cv_2} = \omega_{v_1} + c\omega_{v_2}$ and $\Phi_{v_1+cv_2} = \Phi_{v_1} + c\Phi_{v_2}$ for all $v_1, v_2 \in \mathbb{R}^3$ and $c \in \mathbb{R}$. Moreover, $\omega_v = 0$ iff $v = 0$ and $\Phi_v = 0$ iff $v = 0$ hence $ker(\omega) = \{0\}$ and $ker(\Phi) = \{0\}$ but this means that ω and Φ are injective and since the dimensions of the domain and codomain are 3 and these are linear transformations[17] it follows ω and Φ are isomorphisms.

It appears we have two ways to represent vectors with forms in $\mathbb{R}^3$. We'll see why this is important as we study integration of forms. It turns out the two-forms go with surfaces whereas the one-forms attach to curves. This corresponds to the fact in calculus III we have two ways to integrate a vector-field, we can either calculate flux or work. Partly for this reason the mapping ω is called the **work-form correspondence** and Φ is called the **flux-form correspondence**. Integration has to wait a bit, for now we focus on algebra.

Example 7.3.14. *Suppose $v =< 2, 0, 3 >$ and $w =< 0, 1, 2 >$ then $\omega_v = 2e^1 + 3e^3$ and $\omega_w = e^2 + 2e^3$. Calculate the wedge product,*

$$\begin{aligned} \omega_v \wedge \omega_w &= (2e^1 + 3e^3) \wedge (e^2 + 2e^3) \\ &= 2e^1 \wedge (e^2 + 2e^3) + 3e^3 \wedge (e^2 + 2e^3) \\ &= 2e^1 \wedge e^2 + 4e^1 \wedge e^3 + 3e^3 \wedge e^2 + 6e^3 \wedge e^3 \\ &= -3e^2 \wedge e^3 - 4e^3 \wedge e^1 + 2e^1 \wedge e^2 \\ &= \Phi_{<-3,-4,2>} \\ &= \Phi_{v \times w} \end{aligned} \tag{7.20}$$

Coincidence? Nope.

[17] this is not generally true, note $f(x) = x^2$ has $f(x) = 0$ iff $x = 0$ and yet f is not injective. The linearity is key.

Proposition 7.3.15.

Suppose $v, w \in \mathbb{R}^3$ then $\omega_v \wedge \omega_w = \Phi_{v \times w}$ where $v \times w$ denotes the cross-product which is defined by $v \times w = \sum_{i,j,k=1}^{3} \epsilon_{ijk} v_i w_j e_k$.

Proof: Suppose $v = \sum_{i=1}^{3} v_i e_i$ and $w = \sum_{j=1}^{3} w_j e_j$ then $\omega_v = \sum_{i=1}^{3} v_i e^i$ and $\omega_w = \sum_{j=1}^{3} w_j e^j$. Calculate,

$$\omega_v \wedge \omega_w = \left(\sum_{i=1}^{3} v_i e^i \right) \wedge \left(\sum_{j=1}^{3} w_j e^j \right) = \sum_{i=1}^{3} \sum_{j=1}^{3} v_i w_j e^i \wedge e^j$$

In invite the reader to show $e^i \wedge e^j = \Phi(\sum_{k=1}^{3} \epsilon_{ijk} e_k)$ where I'm using $\Phi_v = \Phi(v)$ to make the argument of the flux-form mapping easier to read, hence,

$$\omega_v \wedge \omega_w = \underbrace{\sum_{i=1}^{3} \sum_{j=1}^{3} v_i w_j \Phi(\sum_{k=1}^{3} \epsilon_{ijk} e_k) = \Phi(\sum_{i,j,k=1}^{3} v_i w_j \epsilon_{ijk} e_k)}_{linearity\ of\ \Phi} = \Phi(v \times w) = \Phi_{v \times w}$$

Of course, if you don't like my proof you could just work it out like the example that precedes this proposition. I gave the proof to show off the mappings a bit more. □

Is the wedge product just the cross-product generalized? Well, not really. I think they're quite different animals. The wedge product is an associative product which makes sense in any vector space. The cross-product only matches the wedge product after we interpret it through a pair of isomorphisms (ω and ϕ) which are special to $\mathbb{R}^3$. However, there is debate, largely the question comes down to what you think makes the cross-product the cross-product. If you think it must pick a unique perpendicular direction to a pair of given directions then that is only going to work in $\mathbb{R}^3$ since even in $\mathbb{R}^4$ there is a whole plane of perpendicular vectors to a given pair. On the other hand, if you think the cross-product in $\mathbb{R}^4$ should be pick the unique perpendicular to a given triple of vectors then you could set something up. You could define $v \times w \times x = \omega^{-1}(\psi(\omega_v \wedge \omega_w \wedge \omega_x))$ where $\psi : \Lambda^3\mathbb{R}^4 \to \Lambda^1\mathbb{R}^4$ is an isomorphism we'll describe in a upcoming section. But, you see it's no longer a product of two vectors, it's not a binary operation, it's a tertiary operation. In any event, you can read a lot more on this if you wish. We have all the tools we need for this course. The wedge product provides the natural antisymmetric algebra for n-dimensiona and the work and flux-form maps naturally connect us to the special world of three-dimensional mathematics.

There is more algebra for forms on $\mathbb{R}^3$ however we defer it to a later section where we have a few more tools. Chief among those is the Hodge dual. But, before we can discuss Hodge duality we need to generalize our idea of a dot-product just a little.

7.4 bilinear forms and geometry, metric duality

The concept of a metric goes beyond the familar case of the dot-product. If you want a more strict generalization of the dot-product then you should think about an *inner-product*. In contrast to the definition below, the inner-product replaces non-degeneracy with the stricter condition of positive-definite which would read $g(x, x) > 0$ for $x \neq 0$. I included a discussion of inner products at the end of this section for the interested reader although we are probably not going to need all of that material.

7.4.1 metric geometry

A **geometry** is a vector space paired with a metric. For example, if we pair $\mathbb{R}^n$ with the dot-product you get Euclidean space. However, if we pair $\mathbb{R}^4$ with the Minkowski metric then we obtain Minkowski space.

Definition 7.4.1.

If V is a vector space and $g : V \times V \to \mathbb{R}$ is

1. bilinear: $g \in T_2^0 V$,
2. **symmetric**: $g(x, y) = g(y, x)$ for all $x, y \in V$,
3. **nondegenerate**: $g(x, y) = 0$ for all $x \in V$ implies $y = 0$.

the we call g a **metric** on V.

If $V = \mathbb{R}^n$ then we can write $g(x, y) = x^T G y$ where $[g] = G$. Moreover, $g(x, y) = g(y, x)$ implies $G^T = G$. Nondegenerate means that $g(x, y) = 0$ for all $y \in \mathbb{R}^n$ iff $x = 0$. It follows that $Gy = 0$ has no non-trivial solutions hence G^{-1} exists.

Example 7.4.2. *Suppose $g(x, y) = x^T y$ for all $x, y \in \mathbb{R}^n$. This defines a metric for $\mathbb{R}^n$, it is just the dot-product. Note that $g(x, y) = x^T y = x^T I y$ hence we see $[g] = I$ where I denotes the identity matrix in $\mathbb{R}^{\ n\times n}$.*

Example 7.4.3. *Suppose $v = (v^0, v^1, v^2, v^3), w = (w^0, w^1, w^2, w^3) \in \mathbb{R}^4$ then define the* **Minkowski product** *of v and w as follows:*

$$g(v, w) = -v^0 w^0 + v^1 w^1 + v^2 w^2 + v^3 w^3$$

It is useful to write the Minkowski product in terms of a matrix multiplication. Observe that for $x, y \in \mathbb{R}^4$,

$$g(x, y) = -x^0 y^0 + x^1 y^1 + x^2 y^2 + x^3 y^3 = \begin{pmatrix} x^0 & x^1 & x^2 & x^3 \end{pmatrix} \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} y^0 \\ y^1 \\ y^2 \\ y^3 \end{pmatrix} \equiv x^t \eta y$$

where we have introduced η the matrix of the Minkowski product. Notice that $\eta^T = \eta$ and $det(\eta) = -1 \neq 0$ hence $g(x, y) = x^t \eta y$ makes g a symmetric, nondegenerate bilinear form on $\mathbb{R}^4$. The formula is clearly related to the dot-product. Suppose $\bar{v} = (v^0, \vec{v})$ and $\bar{w} = (w^0, \vec{w})$ then note

$$g(v, w) = -v^0 w^0 + \vec{v} \cdot \vec{w}$$

For vectors with zero in the zeroth slot this Minkowski product reduces to the dot-product. However, for vectors which have nonzero entries in both the zeroth and later slots much differs. Recall that any vector's dot-product with itself gives the square of the vectors length. Of course this means that $\vec{x} \cdot \vec{x} = 0$ iff $\vec{x} = 0$. Contrast that with the following: if $v = (1, 1, 0, 0)$ then

$$g(v, v) = -1 + 1 = 0$$

Yet $v \neq 0$. Why study such a strange generalization of length? The answer lies in physics. I'll give you a brief account by defining a few terms: Let $v = (v^0, v^1, v^2, v^3) \in \mathbb{R}^4$ then we say

1. *v is a timelike vector if $< v, v > <\ 0$*

2. *v is a lightlike vector if $< v, v > =\ 0$*

3. *v is a spacelike vector if $< v, v > >\ 0$*

If we consider the trajectory of a massive particle in $\mathbb{R}^4$ that begins at the origin then at any later time the trajectory will be located at a timelike vector. If we consider a light beam emitted from the origin then at any future time it will located at the tip of a lightlike vector. Finally, spacelike vectors point to points in $\mathbb{R}^4$ which cannot be reached by the motion of physical particles that pass throughout the origin. We say that massive particles are confined within their light cones, this means that they are always located at timelike vectors relative to their current position in space time. If you'd like to know more I can reccomend a few books.

At this point you might wonder if there are other types of metrics beyond these two examples. Surprisingly, in a certain sense, no. A rather old theorem of linear algebra due to Sylvester states that we can change coordinates so that the metric more or less resembles either the dot-product or something like it with some sign-flips. We'll return to this in a later section.

7.4.2 metric duality for tensors

Throughout this section we consider a vector space V paired with a metric $g : V \times V \to \mathbb{R}$. Moreover, the vector space V has basis $\{e_i\}_{i=1}^n$ which has a g-dual basis $\{e^i\}_{i=1}^n$. Up to this point we always have used a g-dual basis where the duality was offered by the dot-product. In the context of Minkowski geometry that sort of duality is no longer natural. Instead we must follow the definition below:

Definition 7.4.4.

If V is a vector space with metric g and basis $\{e_i\}_{i=1}^n$ then we say the basis $\{e^i\}_{i=1}^n$ is g-dual iff

Suppose $e^i(e_j) = \delta_{ij}$ and consider $g = \sum_{i,j=1}^n g_{ij} e^i \otimes e^j$. Furthermore, suppose g^{ij} are the components of the inverse matrix to (g_{ij}) this means that $\sum_{k=1}^n g_{ik} g^{kj} = \delta_{ij}$. We use the components of the metric and its inverse to raise and lower indices on tensors. Here are the basic conventions: given an object A^j which has the *contravariant* index j we can lower it to be covariant by contracting against the metric components as follows:

$$\boxed{A_i = \sum_j g_{ij} A^j}$$

On the other hand, given an object B_j which has a *covariant* index j we can raise it to be contravariant by contracting against the inverse components of the metric:

$$\boxed{B^i = \sum_j g^{ij} B_j}$$

I like to think of this as some sort of conservation of indices. Strict adherence to the notation drives us to write things such as $\sum_{k=1}^n g_{ik} g^{kj} = \delta_i^j$ just to keep up the up/down index pattern. I should mention that these formulas are much more beautiful in the physics literature, you can look at my old Math 430 notes from NCSU if you'd like a healthy dose of that notation[18]. I use

[18] just a taste: $v_\mu = \eta_{\mu\nu} v^\nu$ or $v^\mu = \eta^{\mu\nu} v_\nu$ or $v^\mu v_\mu = \eta^{\mu\nu} v_\nu v_\mu = \eta_{\mu\nu} v^\mu v^\nu$

Einstein's implicit summation notation throughout those notes and I discuss this index calculation more in the way a physicist typically approaches it. Here I am trying to be careful enough that these equations are useful to mathematicians. Let me show you some examples:

Example 7.4.5. *Specialize for this example to $V = \mathbb{R}^4$ with $g(x,y) = x^T\eta y$. Suppose $x = \sum_{\mu=0}^4 x^\mu e_\mu$ the components x^μ are called* **contravariant components***. The metric allows us to define* **covariant components** *by*

$$x_\nu = \sum_{\mu=0}^{4} \eta_{\nu\mu}x^\mu.$$

For the minkowski metric this just adjoins a minus to the zeroth component: if $(x^\mu) = (a,b,c,d)$ then $x_\mu = (-a,b,c,d)$.

Example 7.4.6. *Suppose we are working on $\mathbb{R}^n$ with the Euclidean metric $g_{ij} = \delta_{ij}$ and it follows that $g^{ij} = \delta_{ij}$ or to be a purist for a moment $\sum_k g_{ik}g^{kj} = \delta_i^j$. In this case $v^i = \sum_j g^{ij}v_j = \sum_j \delta_{ij}v_j = v_i$. The covariant and contravariant components are the same. This is why is was ok to ignore up/down indices when we work with a dot-product exclusively.*

What if we raise an index and the lower it back down once more? Do we really get back where we started? Given x^μ we lower the index by $x_\nu = \sum_\mu g_{\nu\mu}x^\mu$ then we raise it once more by

$$x^\alpha = \sum_\nu g^{\alpha\nu}x_\nu = \sum_\nu g^{\alpha\nu}\sum_\mu g_{\nu\mu}x^\mu = \sum_{\mu,\nu} g^{\alpha\nu}g_{\nu\mu}x^\mu = \sum_\mu \delta^\alpha_\mu x^\mu$$

and the last summation squishes down to x^α once more. It would seem this procedure of raising and lowering indices is at least consistent.

Example 7.4.7. *Suppose we raise the index on the basis $\{e_i\}$ and formally obtain $\{e^j = \sum_k g^{jk}e_k\}$ on the other hand suppose we lower the index on the dual basis $\{e^l\}$ to formally obtain $\{e_m = \sum_l g_{ml}e^l\}$. I'm curious, are these consistent? We should get $e^j(e_m) = \delta^j_m$, I'll be nice an look at $e_m(e^j)$ in the following sense:*

$$\sum_l g_{ml}e^l\Big(\sum_k g^{jk}e_k\Big) = \sum_{l,k} g_{ml}g^{jk}e^l(e_k) = \sum_{l,k} g_{ml}g^{jk}\delta^l_k = \sum_k g_{mk}g^{jk} = \delta^j_m$$

Interesting, but what does it mean?

I used the term *formal* in the preceding example to mean that the example makes sense in as much as you accept the equations which are written. If you think harder about it then you'll find it was rather meaningless. That said, this index notation is rather forgiving.

Ok, but what are we doing? Recall that I insisted on using lower indices for forms and upper indices for vectors? The index conventions I'm toying with above are the reason for this strange notation. When we lower an index we might be changing a vector to a dual vector, or vice-versa when we raise an index we might be changing a dual vector into a vector. Let me be explicit.

1. given $v \in V$ we create $\alpha_v \in V^*$ by the rule $\alpha_v(x) = g(x,v)$.

2. given $\alpha \in V^*$ we create $v_\alpha \in V^{**}$ by the rule $v_\alpha(\beta) = g^{-1}(\alpha,\beta)$ where $g^{-1}(\alpha,\beta) = \sum_{ij}\alpha_i\beta_j g^{ij}$.

Recall we at times identify V and V^{**}. Let's work out the component structure of α_v and see how it relates to v,

$$\alpha_v(e_i) = g(v, e_i) = g(\sum_j v^j e_j, e_i) = \sum_j v^j g(e_j, e_i) = \sum_j v^j g_{ji}$$

Thus, $\alpha_v = \sum_i v_i e^i$ where $v_i = \sum_j v^j g_{ji}$. When we lower the index we're actually using an isomorphism which is provided by the metric to map vectors to forms. The process of raising the index is just the inverse of this isomorphism.

$$v_\alpha(e^i) = g^{-1}(\alpha, e^i) = g^{-1}(\sum_j \alpha_j e^j, e^i) = \sum_j \alpha_j g^{ji}$$

thus $v_\alpha = \sum_i \alpha^i e_i$ where $\alpha^i = \sum_j \alpha_j g^{ji}$.

Suppose we want to change a type $(0, 2)$ tensor to a type $(2, 0)$ tensor. We're given $T : V^* \times V^*$ where $T = \sum_{ij} T^{ij} e_i \otimes e_j$. Define $\tilde{T} : V \times V \to \mathbb{R}$ as follows:

$$\tilde{T}(v, w) = T(\alpha_v, \alpha_w)$$

What does this look like in components? Note $\alpha_{e_i}(e_j) = g(e_i, e_j) = g_{ij}$ hence $\alpha_{e_i} = \sum_j g_{ij} e^j$ and

$$\tilde{T}(e_i, e_j) = T(\alpha_{e_i}, \alpha_{e_j}) = T\left(\sum_k g_{ik} e^k, \ \sum_l g_{jl} e^l\right) = \sum_{k,l} g_{ki} g_{lj} T(e^k, e^l) = \sum_{k,l} g_{ki} g_{lj} T^{kl}$$

Or, as is often customary, we could write $T_{ij} = \sum_{k,l} g_{ik} g_{jl} T^{kl}$. However, this is an abuse of notation since T_{ij} are not technically components for T. If we have a metric we can recover either T from $\tilde{T}$ or vice-versa. Generally, if we are given two tensors, say T_1 of rank (r, s) and the T_2 of rank (r', s'), then these might be equilvalent if $r + s = r' + s'$. It may be that through raising and lowering indices (a.k.a. appropriately composing with the vector↔dual vector isomorphisms) we can convert T_1 to T_2. If you read *Gravitation* by Misner, Thorne and Wheeler you'll find many more thoughts on this equivalence. Challenge: can you find the explicit formulas like $\tilde{T}(v, w) = T(\alpha_v, \alpha_w)$ which back up the index calculations below?

$$T_{ij}{}^{k} = \sum_{a,b} g_{ia} g_{jb} T^{abk} \qquad \text{or} \qquad S^{ij} = \sum_{a,b} g^{ia} g^{jb} S_{ab}$$

I hope I've given you enough to chew on in this section to put these together.

7.4.3 inner products and induced norm

There are generalized dot-products on many abstract vector spaces, we call them **inner-products**.

Definition 7.4.8.

Suppose V is a vector space. If $<, >: V \times V \to \mathbb{R}$ is a function such that for all $x, y, z \in V$ and $c \in \mathbb{R}$:

1. $< x, y >=< y, x >$ (symmetric)
2. $< x + y, z >=< x, z > + < y, z >$ (additive in the first slot)
3. $< cx, y >= c < x, y >$ (homogeneity in the first slot)
4. $< x, x > \geq \ 0$ and $< x, x >= 0$ iff $x = 0$

then we say $(V, <, >)$ is an **inner-product space** with inner product $<, >$.

Given an inner-product space $(V, <, >)$ we can easily induce a norm for V by the formula $||x|| = \sqrt{< x, x >}$ for all $x \in V$. Properties (1.), (3.) and (4.) in the definition of the norm are fairly obvious for the induced norm. Let's think throught the triangle inequality for the induced norm:

$$\begin{aligned} ||x+y||^2 &= < x+y, x+y > && \text{def. of induced norm} \\ &= < x, x+y > + < y, x+y > && \text{additive prop. of inner prod.} \\ &= < x+y, x > + < x+y, y > && \text{symmetric prop. of inner prod.} \\ &= < x, x > + < y, x > + < x, y > + < y, y > && \text{additive prop. of inner prod.} \\ &= \ ||x||^2 + 2 < x, y > + ||y||^2 \end{aligned}$$

At this point we're stuck. A nontrivial identity[19] called the **Cauchy-Schwarz** identity helps us proceed; $< x, y > \leq ||x||||y||$. It follows that $||x+y||^2 \leq ||x||^2 + 2||x||||y|| + ||y||^2 = (||x|| + ||y||)^2$. However, the induced norm is clearly positive[20] so we find $||x+y|| \leq ||x|| + ||y||$.

Most linear algebra texts have a whole chapter on inner-products and their applications, you can look at my notes for a start if you're curious. That said, this is a bit of a digression for this course.

7.5 hodge duality

We can prove that $\binom{n}{p} = \binom{n}{n-p}$. This follows from explicit computation of the formula for $\binom{n}{p}$ or from the symmetry of Pascal's triangle if you prefer. In any event, this equality suggests there is some isomorphism between p and $(n-p)$-forms. When we are given a metric g on a vector space V (and the notation of the preceding section) it is fairly simple to construct the isomorphism. Suppose we are given $\alpha \in \Lambda^p V$ and following our usual notation:

$$\alpha = \sum_{i_1, i_2, \dots, i_p = 1}^{n} \frac{1}{p!} \alpha_{i_1 i_2 \dots i_p} e^{i_1} \wedge e^{i_2} \wedge \cdots \wedge e^{i_p}$$

Then, define $*\alpha$ the **hodge dual** to be the $(n-p)$-form given below:

$$\boxed{*\alpha = \sum_{i_1, i_2, \dots, i_n = 1}^{n} \frac{1}{p!(n-p)!} \alpha^{i_1 i_2 \dots i_p} \epsilon_{i_1 i_2 \dots i_p j_1 j_2 \dots j_{n-p}} e^{j_1} \wedge e^{j_2} \wedge \cdots \wedge e^{j_{n-p}}}$$

I should admit, to prove this is a reasonable definition we'd need to do some work. It's clearly a linear transformation, but bijectivity and coordinate invariance of this definition might take a little work. I intend to omit those details and instead focus on how this works for $\mathbb{R}^3$ or $\mathbb{R}^4$. My advisor taught a course on fiber bundles and there is a much more general and elegant presentation of the hodge dual over a manifold. Ask if interested, I think I have a pdf.

7.5.1 hodge duality in euclidean space $\mathbb{R}^3$

To begin, consider a scalar 1, this is a 0-form so we expect the hodge dual to give a 3-form:

$$*1 = \sum_{i,j,k} \frac{1}{0!3!} \epsilon_{ijk} e^i \wedge e^j \wedge e^k = e^1 \wedge e^2 \wedge e^3$$

[19] I prove this for the dot-product in my linear notes, however, the proof is written in such a way it equally well applies to a general inner-product

[20] note: if you have $(-5)^2 < (-7)^2$ it does not follow that $-5 < -7$, in order to take the squareroot of the inequality we need positive terms squared

Interesting, the hodge dual of 1 is the top-form on $\mathbb{R}^3$. Conversely, calculate the dual of the top-form, note $e^1 \wedge e^2 \wedge e^3 = \sum_{ijk} \frac{1}{6}\epsilon_{ijk} e^i \wedge e^j \wedge e^k$ reveals the components of the top-form are precisely ϵ_{ijk} thus:

$$*(e^1 \wedge e^2 \wedge e^3) = \sum_{i,j,k=1}^{3} \frac{1}{3!(3-3)!}\epsilon^{ijk}\epsilon_{ijk} = \frac{1}{6}(1+1+1+(-1)^2+(-1)^2+(-1)^2) = 1.$$

Next, consider e^1, note that $e^1 = \sum_k \delta_{k1} e^k$ hence the components are δ_{k1}. Thus,

$$*e^1 = \sum_{i,j,k} \frac{1}{1!2!}\epsilon_{ijk}\delta^{k1} e^i \wedge e^j = \frac{1}{2}\sum_{i,j} \epsilon_{ij1} e^i \wedge e^j = \frac{1}{2}(\epsilon_{231} e^2 \wedge e^3 + \epsilon_{321} e^3 \wedge e^2) = e^2 \wedge e^3$$

Similar calculations reveal $*e^2 = e^3 \wedge e^1$ and $*e^3 = e^1 \wedge e^2$. What about the duals of the two-forms? Begin with $\alpha = e^1 \wedge e^2$ note that $e^1 \wedge e^2 = e^1 \otimes e^2 - e^2 \otimes e^1$ thus we can see the components are $\alpha_{ij} = \delta_{i1}\delta_{j2} - \delta_{i2}\delta_{j1}$. Thus,

$$*(e^1 \wedge e^2) = \sum_{i,j,k} \frac{1}{2!1!}\epsilon_{ijk}(\delta_{i1}\delta_{j2} - \delta_{i2}\delta_{j1})e^k = \frac{1}{2}\left(\sum_k \epsilon_{12k} e^k - \sum_k \epsilon_{21k} e^k\right) = \frac{1}{2}(e^3 - (-e^3)) = e^3.$$

Similar calculations show that $*(e^2 \wedge e^3) = e^1$ and $*(e^3 \wedge e^1) = e^2$. Put all of this together and we find that

$$*(ae^1 + be^2 + ce^3) = ae^2 \wedge e^3 + be^3 \wedge e^1 + ce^1 \wedge e^2$$

and

$$*(ae^2 \wedge e^3 + be^3 \wedge e^1 + ce^1 \wedge e^2) = ae^1 + be^2 + ce^3$$

Which means that $\boxed{*\omega_v = \Phi_v}$ and $\boxed{*\Phi_v = \omega_v}$. Hodge duality links the two different form-representations of vectors in a natural manner. Moveover, for $\mathbb{R}^3$ we should also note that $**\alpha = \alpha$ for all $\alpha \in \Lambda\mathbb{R}^3$. In general, for other metrics, we can have a change of signs which depends on the degree of α.

We can summarize hodge duality for three-dimensional Euclidean space as follows:
A simple rule to calculate the hodge dual of a basis form is as follows

1. begin with the top-form $e^1 \wedge e^2 \wedge e^3$
2. permute the forms until the basis form you wish to hodge dual is to the left of the expression, whatever remains to the right is the hodge dual.

For example, to calculate the dual of $e^2 \wedge e^3$ note

$$e^1 \wedge e^2 \wedge e^3 = \underbrace{e^2 \wedge e^3}_{to\ be\ dualed} \wedge \underbrace{e^1}_{the\ dual} \quad \Rightarrow \quad *(e^2 \wedge e^3) = e^1.$$

Consider what happens if we calculate $**\alpha$, since the dual is a linear operation it suffices to think about the basis forms. Let me sketch the process of $**e^I$ where I is a multi-index:

1. begin with $e^1 \wedge e^2 \wedge e^3$
2. write $e^1 \wedge e^2 \wedge e^3 = (-1)^N e^I \wedge e^J$ and identify $*e^I = (-1)^N e^J$.
3. then to calculate the second dual once more begin with $e^1 \wedge e^2 \wedge e^3$ and note

$$e^1 \wedge e^2 \wedge e^3 = (-1)^N e^J \wedge e^I$$

since the same N transpositions are required to push e^I to the left or e^J to the right.

4. It follows that $**e^I = e^I$ for any multi-index hence $\boxed{**\alpha = \alpha}$ for all $\alpha \in \Lambda\mathbb{R}^3$.

I hope that once you get past the index calculation you can see the hodge dual is not a terribly complicated construction. Some of the index calculation in this section was probably gratutious, but I would like you to be aware of such techniques. Brute-force calculation has it's place, but a well-thought index notation can bring far more insight with much less effort.

7.5.2 hodge duality in minkowski space $\mathbb{R}^4$

The logic here follows fairly close to the last section, however the wrinkle is that the metric here demands more attention. We must take care to raise the indices on the forms when we Hodge dual them. First let's list the basis forms, we have to add time to the mix (again $c = 1$ so $x^0 = ct = t$ if you worried about it) Remember that the Greek indices are defined to range over $0, 1, 2, 3$. Here the top form is degree four since in four dimensions we can have at most four dual-basis vectors without a repeat. Wedge products work the same as they have before, just now we have e^0 to play with. Hodge duality may offer some surprises though.

Definition 7.5.1. *The antisymmetric symbol in* **flat** $\mathbb{R}^4$ *is denoted* $\epsilon_{\mu\nu\alpha\beta}$ *and it is defined by the value*

$$\epsilon_{0123} = 1$$

plus the demand that it be completely antisymmetric.

We must not assume that this symbol is invariant under a cyclic exhange of indices. Consider,

$$\begin{aligned} \epsilon_{0123} &= -\epsilon_{1023} && \text{flipped (01)} \\ &= +\epsilon_{1203} && \text{flipped (02)} \\ &= -\epsilon_{1230} && \text{flipped (03).} \end{aligned} \tag{7.21}$$

In four dimensions we'll use antisymmetry directly and forego the cyclicity shortcut. Its not a big deal if you notice it before it confuses you.

Example 7.5.2. *Find the Hodge dual of* $\gamma = e^1$ *with respect to the Minkowski metric* $\eta_{\mu\nu}$*, to begin notice that* dx *has components* $\gamma_\mu = \delta^1_\mu$ *as is readily verified by the equation* $e^1 = \sum_\mu \delta^1_\mu e^\mu$*. Lets raise the index using* η *as we learned previously,*

$$\gamma^\mu = \sum_\nu \eta^{\mu\nu}\gamma_\nu = \sum_\nu \eta^{\mu\nu}\delta^1_\nu = \eta^{1\mu} = \delta^{1\mu}$$

Starting with the definition of Hodge duality we calculate

$$\begin{aligned} {}^*(e^1) &= \textstyle\sum_{\alpha,\beta,\mu,\nu} \frac{1}{p!}\frac{1}{(n-p)!}\gamma^\mu \epsilon_{\mu\nu\alpha\beta} e^\nu \wedge e^\alpha \wedge e^\beta \\ &= \textstyle\sum_{\alpha,\beta,\mu,\nu} (1/6)\delta^{1\mu}\epsilon_{\mu\nu\alpha\beta} e^\nu \wedge e^\alpha \wedge e^\beta \\ &= \textstyle\sum_{\alpha,\beta,\nu} (1/6)\epsilon_{1\nu\alpha\beta} e^\nu \wedge e^\alpha \wedge e^\beta \\ &= (1/6)[\epsilon_{1023} e^0 \wedge e^2 \wedge e^3 + \epsilon_{1230} e^2 \wedge e^3 \wedge e^0 + \epsilon_{1302} e^3 \wedge e^0 \wedge e^2 \\ &\qquad + \epsilon_{1320} e^3 \wedge e^2 \wedge e^0 + \epsilon_{1203} e^2 \wedge e^0 \wedge e^3 + \epsilon_{1032} e^0 \wedge e^3 \wedge e^2] \\ &= (1/6)[-e^0 \wedge e^2 \wedge e^3 - e^2 \wedge e^3 \wedge e^0 - e^3 \wedge e^0 \wedge e^2 \\ &\qquad + e^3 \wedge e^2 \wedge e^0 + e^2 \wedge e^0 \wedge e^3 + e^0 \wedge e^3 \wedge e^2] \\ &= -e^2 \wedge e^3 \wedge e^0 = -e^0 \wedge e^2 \wedge e^3. \end{aligned} \tag{7.22}$$

the difference between the three and four dimensional Hodge dual arises from two sources, for one we are using the Minkowski metric so indices up or down makes a difference, and second the antisymmetric symbol has more possibilities than before because the Greek indices take four values.

I suspect we can calculate the hodge dual by the following pattern: suppose we wish to find the dual of α where α is a basis form for $\Lambda\mathbb{R}^4$ with the Minkowski metric

1. begin with the top-form $e^0 \wedge e^1 \wedge e^2 \wedge e^3$
2. permute factors as needed to place α to the left,
3. the form which remains to the right will be the hodge dual of α if no e^0 is in α otherwise the form to the right multiplied by -1 is $*\alpha$.

Note this works for the previous example as follows:

1. begin with $e^0 \wedge e^1 \wedge e^2 \wedge e^3$
2. note $e^0 \wedge e^1 \wedge e^2 \wedge e^3 = -e^1 \wedge e^0 \wedge e^2 \wedge e^3 = e^1 \wedge (-e^0 \wedge e^2 \wedge e^3)$
3. identify $*e^1 = -e^0 \wedge e^2 \wedge e^3$ (no extra sign since no e^0 appears in e^1)

Follow the algorithm for finding the dual of e^0,

1. begin with $e^0 \wedge e^1 \wedge e^2 \wedge e^3$
2. note $e^0 \wedge e^1 \wedge e^2 \wedge e^3 = e^0 \wedge (e^1 \wedge e^2 \wedge e^3)$
3. identify $*e^0 = -e^1 \wedge e^2 \wedge e^3$ (added sign since e^0 appears in form being hodge dualed)

Let's check from the definition if my algorithm worked out right.

Example 7.5.3. *Find the Hodge dual of $\gamma = e^0$ with respect to the Minkowski metric $\eta_{\mu\nu}$, to begin notice that e^0 has components $\gamma_\mu = \delta^0_\mu$ as is readily verified by the equation $e^0 = \sum_\mu \delta^0_\mu e^\mu$. Lets raise the index using η as we learned previously,*

$$\gamma^\mu = \sum_\nu \eta^{\mu\nu}\gamma_\nu = \sum_\nu \eta^{\mu\nu}\delta^0_\nu = \eta^{\mu 0} = -\delta^{0\mu}$$

the minus sign is due to the Minkowski metric. Starting with the definition of Hodge duality we calculate

$$\begin{aligned}
{}^*(e^0) &= \textstyle\sum_{\alpha,\beta,\mu,\nu} \frac{1}{p!}\frac{1}{(n-p)!}\gamma^\mu \epsilon_{\mu\nu\alpha\beta} e^\nu \wedge e^\alpha \wedge e^\beta \\
&= \textstyle\sum_{\alpha,\beta,\mu,\nu} -(1/6)\delta^{0\mu}\epsilon_{\mu\nu\alpha\beta} e^\nu \wedge e^\alpha \wedge e^\beta \\
&= \textstyle\sum_{\alpha,\beta,\nu} -(1/6)\epsilon_{0\nu\alpha\beta} e^\nu \wedge e^\alpha \wedge e^\beta \qquad (7.23) \\
&= \textstyle\sum_{i,j,k} -(1/6)\epsilon_{0ijk} e^i \wedge e^j \wedge e^k \\
&= \textstyle\sum_{i,j,k} -(1/6)\epsilon_{ijk}\epsilon_{ijk} e^1 \wedge e^2 \wedge e^3 \qquad \leftarrow \textit{sneaky step} \\
&= -e^1 \wedge e^2 \wedge e^3.
\end{aligned}$$

Notice I am using the convention that Greek indices sum over $0, 1, 2, 3$ whereas Latin indices sum over $1, 2, 3$.

Example 7.5.4. *Find the Hodge dual of $\gamma = e^0 \wedge e^1$ with respect to the Minkowski metric $\eta_{\mu\nu}$, to begin notice the following identity, it will help us find the components of γ*

$$e^0 \wedge e^1 = \sum_{\mu,\nu} \frac{1}{2} 2\delta^0_\mu \delta^1_\nu e^\mu \wedge e^\nu$$

now we antisymmetrize to get the components of the form,

$$e^0 \wedge e^1 = \sum_{\mu,\nu} \frac{1}{2} \delta^0_{[\mu} \delta^1_{\nu]} dx^\mu \wedge dx^\nu$$

where $\delta^0_{[\mu}\delta^1_{\nu]} = \delta^0_\mu\delta^1_\nu - \delta^0_\nu\delta^1_\mu$ and the factor of two is used up in the antisymmetrization. Lets raise the index using η as we learned previously,

$$\gamma^{\alpha\beta} = \sum_{\mu,\nu} \eta^{\alpha\mu}\eta^{\beta\nu}\gamma_{\mu\nu} = \sum_{\mu,\nu} \eta^{\alpha\mu}\eta^{\beta\nu}\delta^0_{[\mu}\delta^1_{\nu]} = -\eta^{\alpha 0}\eta^{\beta 1} + \eta^{\beta 0}\eta^{\alpha 1} = -\delta^{[\alpha 0}\delta^{\beta]1}$$

the minus sign is due to the Minkowski metric. Starting with the definition of Hodge duality we calculate

$$\begin{aligned} {}^*(e^0 \wedge e^1) &= \tfrac{1}{p!}\tfrac{1}{(n-p)!}\gamma^{\alpha\beta}\epsilon_{\alpha\beta\mu\nu}e^\mu \wedge e^\nu \\ &= (1/4)(-\delta^{[\alpha 0}\delta^{\beta]1})\epsilon_{\alpha\beta\mu\nu}e^\mu \wedge e^\nu \\ &= -(1/4)(\epsilon_{01\mu\nu}e^\mu \wedge e^\nu - \epsilon_{10\mu\nu}e^\mu \wedge e^\nu) \\ &= -(1/2)\epsilon_{01\mu\nu}e^\mu \wedge e^\nu \\ &= -(1/2)[\epsilon_{0123}e^2 \wedge e^3 + \epsilon_{0132}e^3 \wedge e^2] \\ &= -e^2 \wedge e^3 \end{aligned} \tag{7.24}$$

Note, the algorithm works out the same,

$$e^0 \wedge e^1 \wedge e^2 \wedge e^3 = \underbrace{e^0 \wedge e^1}_{has\ e^0} \wedge (e^2 \wedge e^3) \quad \Rightarrow \quad *(e^0 \wedge e^1) = -e^2 \wedge e^3$$

The other Hodge duals of the basic two-forms calculate by almost the same calculation. Let us make a table of all the basic Hodge dualities in Minkowski space, I have grouped the terms to emphasize the isomorphisms between the one-dimensional $\Lambda^0\mathbb{R}^4$ and $\Lambda^4\mathbb{R}^4$, the four-dimensional $\Lambda^1\mathbb{R}^4$ and $\Lambda^3\mathbb{R}^4$, the six-dimensional $\Lambda^2\mathbb{R}^4$ and itself. Notice that the dimension of $\Lambda\mathbb{R}^4$ is 16 which we have explained in depth in the previous section. Finally, it is useful to point out the three-dimensional work and flux form mappings to provide some useful identities in this $1+3$-dimensional setting.

$$\boxed{*\omega_{\vec{v}} = -e^0 \wedge \Phi_{\vec{v}} \qquad *\,\Phi_{\vec{v}} = e^0 \wedge \omega_{\vec{v}} \qquad *\left(e^0 \wedge \Phi_{\vec{v}}\right) = \omega_{\vec{v}}}$$

I leave verification of these formulas to the reader (use the table). Finally let us analyze the process of taking two hodge duals in succession. In the context of $\mathbb{R}^3$ we found that $**\alpha = \alpha$, we seek to discern if a similar formula is available in the context of $\mathbb{R}^4$ with the minkowksi metric. We can calculate one type of example with the identities above:

$$*\omega_{\vec{v}} = -e^0 \wedge \Phi_{\vec{v}} \quad \Rightarrow \quad **\,\omega_{\vec{v}} = -*(e^0 \wedge \Phi_{\vec{v}}) = -\omega_{\vec{v}} \quad \Rightarrow \quad **\,\omega_{\vec{v}} = -\omega_{\vec{v}}$$

Perhaps this is true in general?

If we accept my algorithm then it's not too hard to sort through using multi-index notation: since hodge duality is linear it suffices to consider a basis element e^I where I is a multi-index,

1. transpose dual vectors so that $e^0 \wedge e^1 \wedge e^2 \wedge e^3 = (-1)^N e^I \wedge e^J$

2. if $0 \notin I$ then $*e^I = (-1)^N e^J$ and $0 \in J$ since $I \cup J = \{0,1,2,3\}$. Take a second dual by writing $e^0 \wedge e^1 \wedge e^2 \wedge e^3 = (-1)^N e^J \wedge e^I$ but note $*((-1)^N e^J) = -e^I$ since $0 \in J$. We find $**e^I = -e^I$ for all I not containing the 0-index.

3. if $0 \in I$ then $*e^I = -(-1)^N e^J$ and $0 \notin J$ since $I \cup J = \{0,1,2,3\}$. Take a second dual by writing $e^0 \wedge e^1 \wedge e^2 \wedge e^3 = -(-1)^N e^J \wedge (-e^I)$ and hence $*(-(-1)^N e^J) = -e^I$ since $0 \notin J$. We find $**e^I = -e^I$ for all I containing the 0-index.

4. it follows that $\boxed{**\alpha = -\alpha}$ for all $\alpha \in \Lambda\mathbb{R}^4$ with the minkowski metric.

To conclude, I would warn the reader that the results in this section pertain to our choice of notation for $\mathbb{R}^4$. Some other texts use a metric which is $-\eta$ relative to our notation. This modifies many signs in this section. See Misner, Thorne and Wheeler's *Gravitation* or Bertlmann's *Anomalies in Field Theory* for future reading on Hodge duality and a more systematic explanation of how and when these signs arise from the metric.

7.6 coordinate change

Suppose V has two bases $\bar{\beta} = \{\bar{f}_1, \bar{f}_2, \dots, \bar{f}_n\}$ and $\beta = \{f_1, f_2, \dots, f_n\}$. If $v \in V$ then we can write v in as a linear combination of the $\bar{\beta}$ basis or the β basis:

$$v = x^1 f_1 + x^2 f_2 + \cdots + x^n f_n \quad \text{and} \quad v = \bar{x}^1 \bar{f}_1 + \bar{x}^2 \bar{f}_2 + \cdots + \bar{x}^n \bar{f}_n$$

given the notation above, we define **coordinate maps** as follows:

$$\Phi_\beta(v) = (x^1, x^2, \dots, x^n) = x \quad \text{and} \quad \Phi_{\bar{\beta}}(v) = (\bar{x}^1, \bar{x}^2, \dots, \bar{x}^n) = \bar{x}$$

We sometimes use the notation $\Phi_\beta(v) = [v]_\beta = x$ whereas $\Phi_{\bar{\beta}}(v) = [v]_{\bar{\beta}} = \bar{x}$. A coordinate map takes an abstract vector v and maps it to a particular representative in $\mathbb{R}^n$. A natural question to ask is how do different representatives compare? How do x and $\bar{x}$ compare in our current notation? Because the coordinate maps are isomorphisms it follows that $\Phi_\beta \circ \Phi_{\bar{\beta}}^{-1} : \mathbb{R}^n \to \mathbb{R}^n$ is an isomorphism and given the domain and codomain we can write its formula via matrix multiplication:

$$\Phi_\beta \circ \Phi_{\bar{\beta}}^{-1}(u) = Pu \quad \Rightarrow \quad \Phi_\beta \circ \Phi_{\bar{\beta}}^{-1}(\bar{x}) = P\bar{x}$$

However, $\Phi_{\bar{\beta}}^{-1}(\bar{x}) = v$ hence $\Phi_\beta(v) = P\bar{x}$ and consequently, $\boxed{x = P\bar{x}}$. Conversely, to switch to barred coordinates we multiply the coordinate vectors by P^{-1}; $\boxed{\bar{x} = P^{-1}x}$.

Continuing this discussion we turn to the dual space. Suppose $\bar{\beta}^* = \{\bar{f}^j\}_{j=1}^n$ is dual to $\bar{\beta} = \{\bar{f}_j\}_{j=1}^n$ and $\beta^* = \{f^j\}_{j=1}^n$ is dual to $\beta = \{f_j\}_{j=1}^n$. By definition we are given that $f^j(f_i) = \delta_{ij}$ and $\bar{f}^j(\bar{f}_i) = \delta_{ij}$ for all $i, j \in \mathbb{N}_n$. Suppose $\alpha \in V^*$ is a dual vector with components α_j with respect to the β^* basis and components $\bar{\alpha}_j$ with respect to the $\bar{\beta}^*$ basis. In particular this means we can either write $\alpha = \sum_{j=1}^n \alpha_j f^j$ or $\alpha = \sum_{j=1}^n \bar{\alpha}_j \bar{f}^j$. Likewise, given a vector $v \in V$ we can either write $v = \sum_{i=1}^n x^i f_i$ or $v = \sum_{i=1}^n \bar{x}^i \bar{f}_i$. With these notations in mind calculate:

$$\alpha(v) = \left(\sum_{j=1}^n \alpha_j f^j\right)\left(\sum_{i=1}^n x^i f_i\right) = \sum_{i,j=1}^n \alpha_j x^i f^j(f_i) = \sum_{i=1}^n \sum_{j=1}^n \alpha_j x^i \delta_{ij} = \sum_{i=1}^n \alpha_i x^i$$

and by the same calculation in the barred coordinates we find, $\alpha(v) = \sum_{i=1}^n \bar{\alpha}_i \bar{x}^i$. Therefore,

$$\sum_{i=1}^n \alpha_i x^i = \sum_{i=1}^n \bar{\alpha}_i \bar{x}^i.$$

Recall, $x = P\bar{x}$. In components, $x^i = \sum_{k=1}^n P^i_k \bar{x}^k$. Substituting,

$$\sum_{i=1}^n \sum_{k=1}^n \alpha_i P^i_k \bar{x}^k = \sum_{i=1}^n \bar{\alpha}_i \bar{x}^i.$$

But, this formula holds for all possible vectors v and hence all possible coordinate vectors $\bar{x}$. If we consider $v = \bar{f}_j$ then $\bar{x}^i = \delta_{ij}$ hence $\sum_{i=1}^n \bar{\alpha}_i \bar{x}^i = \sum_{i=1}^n \bar{\alpha}_i \delta_{ij} = \bar{\alpha}_j$. Moreover, $\sum_{i=1}^n \sum_{k=1}^n \alpha_i P^i_k \bar{x}^k = \sum_{i=1}^n \sum_{k=1}^n \alpha_i P^i_k \delta_{kj} = \sum_{i=1}^n \alpha_i P^i_j$. Thus, $\bar{\alpha}_j = \sum_{i=1}^n P^i_j \alpha_i$. Compare how vectors and dual vectors transform:

$$\boxed{\bar{\alpha}_j = \sum_{i=1}^n P^i_j \alpha_i \qquad \text{verses} \qquad \bar{x}^j = \sum_{i=1}^n (P^{-1})^j_i x^i.}$$

It is customary to use lower-indices on the components of dual-vectors and upper-indices on the components of vectors: we say $x = \sum_{i=1}^n x^i e_i \in \mathbb{R}^n$ has **contravariant components** whereas $\alpha = \sum_{j=1}^n \alpha_j e^j \in (\mathbb{R}^n)^*$ has **covariant components**. These terms arise from the coordinate change properties we derived in this section. The convenience of the up/down index notation will be more apparent as we continue our study to more complicated objects. It is interesting to note the basis elements tranform inversely:

$$\boxed{\bar{f}^j = \sum_{i=1}^n (P^{-1})^j_i f^i \qquad \text{verses} \qquad \bar{f}_j = \sum_{i=1}^n P^i_j f_i.}$$

The formulas above can be derived by arguments similar to those we already gave in this section, however I think it may be more instructive to see how these rules work in concert:

$$\begin{aligned} x = \sum_{i=1}^n \bar{x}^i \bar{f}_i &= \sum_{i=1}^n \sum_{j=1}^n (P^{-1})^i_j x^j \bar{f}_i \qquad (7.25) \\ &= \sum_{i=1}^n \sum_{j=1}^n (P^{-1})^i_j x^j \sum_{k=1}^n P^k_i f_k \\ &= \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n (P^{-1})^i_j P^k_i x^j f_k \\ &= \sum_{j=1}^n \sum_{k=1}^n \delta^k_j x^j f_k \\ &= \sum_{k=1}^n x^k f_k. \end{aligned}$$

7.6.1 coordinate change for $T^0_2(V)$

For an abstract vector space, or for $\mathbb{R}^n$ with a nonstandard basis, we have to replace v, w with their coordinate vectors. If V has basis $\bar{\beta} = \{\bar{f}_1, \bar{f}_2, \dots, \bar{f}_n\}$ with dual basis $\beta^* = \{\bar{f}^1, \bar{f}^2, \dots, \bar{f}^n\}$ and

v, w have coordinate vectors $\bar{x}, \bar{y}$ (which means $v = \sum_{i=1}^{n} \bar{x}^i \bar{f}_i$ and $w = \sum_{i=1}^{n} \bar{y}^i \bar{f}_i$) then,

$$b(v, w) = \sum_{i,j=1}^{n} \bar{x}^i \bar{y}^j \bar{B}_{ij} = \bar{x}^T \bar{B} \bar{y}$$

where $\bar{B}_{ij} = b(\bar{f}_i, \bar{f}_j)$. If $\beta = \{f_1, f_2, \ldots, f_n\}$ is another basis on V with dual basis β^* then we define $B_{ij} = b(f_i, f_j)$ and we have

$$b(v, w) = \sum_{i,j=1}^{n} x^i y^j B_{ij} = x^T B y.$$

Recall that $\bar{f}_i = \sum_{k=1}^{n} P_i^k f_k$. With this in mind calculate:

$$\bar{B}_{ij} = b(\bar{f}_i, \bar{f}_j) = b\left(\sum_{k=1}^{n} P_i^k f_k, \sum_{l=1}^{n} P_j^l f_l\right) = \sum_{k,l=1}^{n} P_i^k P_j^l b(f_k, f_l) = \sum_{k,l=1}^{n} P_i^k P_j^l B_{kl}$$

We find the components of a bilinear map transform as follows:

$$\boxed{\bar{B}_{ij} = \sum_{k,l=1}^{n} P_i^k P_j^l B_{kl}}$$

Chapter 8

calculus with differential forms

A manifold is an abtract space which allows for local calculus. We discuss how coordinate charts cover the a manifold and how we use them to define smoothness in the abstract. For example, a function is smooth if all its local coordinate representations are smooth. The local coordinate representative is a function from $\mathbb{R}^m$ to $\mathbb{R}^n$ thus we may quantify its smoothness in terms of ordinary partial derivatives of the component functions. On the other hand, while the concept of a coordinate chart is at first glance abstract the usual theorems of advanced calculus all lift naturally to the abstract manifold. For example, we see how partial derivatives with respect to manifold coordinates hold to all the usual linearity, product and chain-rules hold in $\mathbb{R}^n$. We prove a number of these results in considerably more detail than Lecture will bear.

The differential gains a deeper meaning than we found in advanced calculus. In the manifold context, the differential acts on tangent space which is **not** identified as some subset of the manifold itself. So, in a sense we lose the direct approximating concept for the differential. One could always return to the best linear approximation ideal as needed, but the path ahead is quite removed from pure numerical approximation. First step towards this abstract picture of tangent space is the realization that tangent vectors themselves should be identified as **derivations**. We show how partial derivatives give derivations and we sketch a technical result which also provides a converse in the smooth category[1] Once we've settled how to study the tangent space to a manifold we find the natural extension of the differential as the **push-forward** induced from a smooth map. It turns out that you have probably already calculated a few push-forwards in multivariate calculus. We attempt to obtain some intuition for this abstract push-forward. We also pause to note how the push-forward might allow us to create new vectors fields from old (or not).

The cotangent space is the dual space to the tangent space. The basis which is dual to the partial derivative basis is naturally identified with the differentials of the coordinate maps themselves. We then have a basis and dual basis for the tangent and cotangent space at each point on the manifold. We use these to build vector fields and differential forms just as we used f_i and f^i in the previous chapter. Multilinear algebra transfers over point by point on the manifold. However, something new happens. Differential forms permit a differentiation called the **exterior derivative**. This natural operation takes a p-form and generates a new $(p+1)$-form. We examine how the exterior derivative recovers **all** the interesting vector-calculus derivatives from vector calculus on $\mathbb{R}^3$. Of course, it goes much deeper as the exterior derivative provides part of the machinery to write co-

[1] it's actually false for C^k manifolds which have an infinite-dimensional space of derivations. The tangent space to a n-dimensional manifold is an n-dimensional vector space so we need an n-dimensional space of derivations to make the identification.

homology on spaces of arbitrarily high dimension. Ultimately, this theory of cohomology detects topological aspects of the space. A basic example of this is the Poincare Lemma.

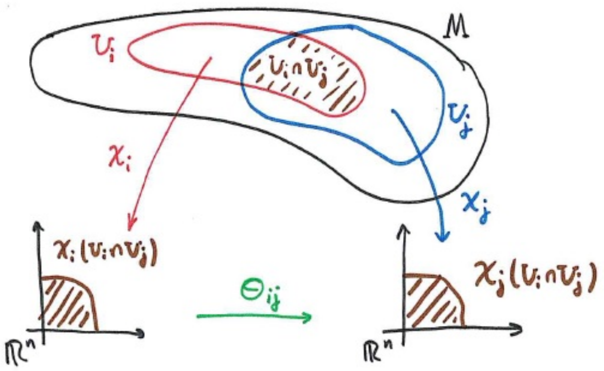
To understand the Poincare Lemma as well as a number of other interesting calculations we find it necessary to introduce a dual operation to the push-foward; the **pull-back** gives us a natural method to take differential forms in the range of a mapping and transport them back to the domain of the map. Moreover, this pull-back operation plays nicely with the wedge product and the exterior derivative. Several pages are devotes towards understanding some intuitive method of calculating the pull-back. We are indebted to Harold M. Edwards' *Advanced Calculus: A Differential Form Approach* which encouraged us to search for intuition. I also attempted to translate a differential forms version of the implict mapping theorem from the same text, I'm fairly certain there is some conceptual error in that section as it stands. It is a work in progress.

The abstract calculus of forms is interesting in it's own right, but we are happy to find how it reduces to the familar calculus on $\mathbb{R}^3$. We state[2] the Generalized Stokes Theorem and see how the **flux-form** and **work-form** mappings produce the usual theorems of vector calculus as corollaries to the Generalized Stokes Theorem. The definition of integrals of differential forms is accomplished by pulling-back the forms to euclidean space where an ordinary integral quantifies the result. In some sense, this discussion may help answer the question *what is a differential form?*. We spend some effort attempting to understand how the form integration interfaces with ordinary surface or line integration of vector fields.

Finally, the fact that $d^2 = 0$ paired with the nice properties of the pull-back and wedge product proves to give a technique for study of exact differential equations and partial differential equations. This final section opens a door to the vast topic of exterior differential systems. In this study, solutions to PDEs are manifolds and the PDE itself is formulated in terms of wedge products and differential forms. Here I borrow wisdom from *Cartan for Beginners* by Ivey and Landsberg as well as *Equivalence, Invariance, and Symmetry* by Olver. Please keep in mind, we're just dipping are toes in the pond here.

8.1 an informal introduction to manifolds

A manifold $\mathcal{M}$ is space which locally resembles $\mathbb{R}^n$. This requirement is implemented by the existence of coordinate charts (χ, U) which cover $\mathcal{M}$ and allow us to do calculus locally. Let me sketch the idea with a picture:



[2] the proof is found in Edwards and many other places

The charts (χ_j, U_j) have to cover the manifold $\mathcal{M}$ and their **transition functions** θ_{ij} must be smooth mappings on $\mathbb{R}^n$. Rather going on about the proper definition[3], I'll show a few examples.

Example 8.1.1. *Let $\mathcal{M} = \{(x,y) \in \mathbb{R}^2 \mid x^2+y^2 = 1\}$. The usual chart on the unit-circle is the angle-chart $\chi = \theta$. Given $(\cos t, \sin t) = p \in \mathcal{M}$ we define $\theta(p) = t$. If $x > 0$ and $x^2+y^2 = 1$ then $(x,y) \in \mathcal{M}$ and we have $\theta(x,y) = \tan^{-1}(y/x)$.*

Example 8.1.2. *Let $\mathcal{M} = \mathbb{R}^2$. The usual chart is simply the cartesian coordinate system $\chi_1 = (x,y)$ with $U_1 = \mathbb{R}^2$. If $(a,b) \in \mathbb{R}^2$ then $x(a,b) = a$ and $y(a,b) = b$. In practice the symbols x, y are used both as maps and variables so one must pay attention to context. A second coordinate system on $\mathcal{M}$ is given by the polar coordinate chart $\chi_2 = (r,\theta)$ with domain $U_2 = (0,\infty)\times\mathbb{R}$. I'll just take their domain to be the right half-plane for the sake of having a nice formula: $(r,\theta)(a,b) = (\sqrt{a^2+b^2}, \tan^{-1}(b/a))$. You can extend these to most of the plane, but you have to delete the origin and you must lose a ray since the angle-chart is not injective if we go full-circle. That said, the coordinate systems $(\chi_1 = (x,y), U_1)$ and $(\chi_2 = (r,\theta), U_2)$ are* **compatible** *because they have smooth transition functions. One can calculate $\chi_1 \circ \chi_2^{-1}$ is a smooth mapping on $\mathbb{R}^2$. Explicitly:*

$$\chi_1 \circ \chi_2^{-1}(u,v) = \chi_1(\chi_2^{-1}(u,v)) = \chi_1(u\cos v, u\sin v) = (u\cos(v), u\sin(v))$$

Technically, the use of the term **coordinate system** in calculus III is less strict than the concept which appears in manifold theory. Departure from injectivity in a geometrically tractible setting is manageable, but for the abstract setting, injectivity of the coordinate charts is important to many arguments.

Example 8.1.3. *Define $\chi_{spherical}(x,y,z) = (r,\theta,\phi)$ implicitly by the coordinate transformations*

$$x = r\cos(\theta)\sin(\phi), \quad y = r\sin(\theta)\sin(\phi), \quad z = r\cos(\phi)$$

These can be inverted,

$$r = \sqrt{x^2+y^2+z^2}, \quad \theta = \tan^{-1}\left[\frac{y}{x}\right], \quad \phi = \cos^{-1}\left[\frac{z}{\sqrt{x^2+y^2+z^2}}\right]$$

To show compatibility with the standard Cartesian coordinates we would need to select a subset of $\mathbb{R}^3$ for which $\chi_{spherical}$ is 1-1 and the since $\chi_{Cartesian} = Id$ the transition functions are just $\chi^{-1}_{spherical}$.

Example 8.1.4. *Define $\chi_{cylindrical}(x,y,z) = (s,\theta,z)$ implicitly by the coordinate transformations*

$$x = s\cos(\theta), \quad y = s\sin(\theta), \quad z = z$$

These can be inverted,

$$s = \sqrt{x^2+y^2}, \quad \theta = \tan^{-1}\left[\frac{y}{x}\right], \quad z = z$$

You can take dom$(\chi_{cylindrical}) = \{(x,y,z) \mid 0 < \theta < 2\pi, \} - \{(0,0,0)\}$

Example 8.1.5. *Let $\mathcal{M} = V$ where V is an n-dimensional vector space over $\mathbb{R}$. If β_1 is a basis for V then $\Phi_{\beta_1} : V \to \mathbb{R}^n$ gives a global coordinate chart. Moreover, if β_2 is another basis for V then $\Phi_{\beta_2} : V \to \mathbb{R}^n$ also gives a global coordinate chart. The transition function $\theta_{12} = \Phi_{\beta_2} \circ \Phi_{\beta_1}^{-1}$ is linear hence it is clearly smooth. In short, an n-dimensional vector space is an n-dimensional manifold. We could put non-linear charts on V if we wished, that freedom is new here, in linear algebra all the coordinate systems considered are linear.*

[3] see my 2011 notes, or better yet study Loring Tu's *An Introduction to Manifolds*

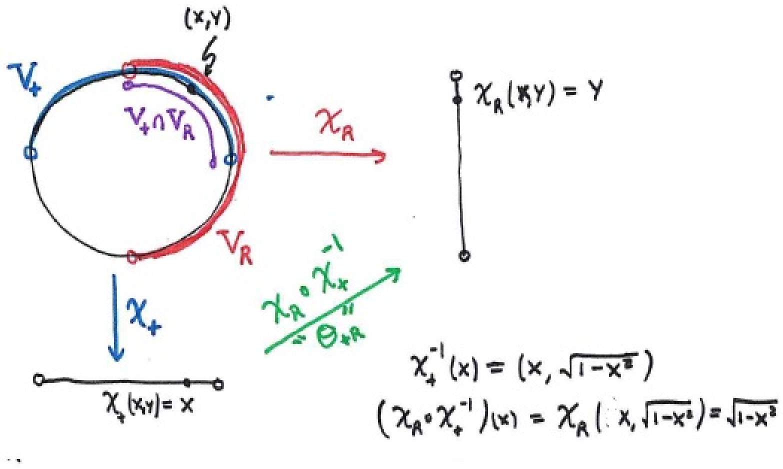
Example 8.1.6. *Let $\mathcal{M} = \{(x,y) \in \mathbb{R}^2 \mid x^2+y^2=1\}$.*

1. *Let $V_+ = \{(x,y) \in \mathcal{M} \mid y > 0\} = dom(\chi_+)$ and define $\chi_+(x,y) = x$*
2. *Let $V_- = \{(x,y) \in \mathcal{M} \mid y < 0\} = dom(\chi_-)$ and define $\chi_-(x,y) = x$*
3. *Let $V_R = \{(x,y) \in \mathcal{M} \mid x > 0\} = dom(\chi_R)$ and define $\chi_R(x,y) = y$*
4. *Let $V_L = \{(x,y) \in \mathcal{M} \mid x < 0\} = dom(\chi_L)$ and define $\chi_L(x,y) = y$*

The set of charts $\mathcal{A} = \{(V_+,\chi_+),(V_-,\chi_-),(V_R,\chi_R),(V_L,\chi_L)\}$ forms an atlas on $\mathcal{M}$ which gives the circle a differentiable structure[4]. It is not hard to show the transition functions are smooth on the image of the intersection of their respect domains. For example, $V_+ \cap V_R = W_{+R} = \{(x,y) \in \mathcal{M} \mid x,y > 0\}$, it's easy to calculate that $\chi_+^{-1}(x) = (x, \sqrt{1-x^2})$ hence

$$(\chi_R \circ \chi_+^{-1})(x) = \chi_R(x, \sqrt{1-x^2}) = \sqrt{1-x^2}$$

for each $x \in \chi_R(W_{+R})$. Note $x \in \chi_R(W_{+R})$ implies $0 < x < 1$ hence it is clear the transition function is smooth.


(x,y)
V_+
$V_+ \cap V_R$
χ_R
$\chi_R(x,y) = y$
V_R
$\chi_R \circ \chi_+^{-1}$
$= \theta_{+R}$
χ_+
$\chi_+(x,y) = x$
$\chi_+^{-1}(x) = (x, \sqrt{1-x^2})$
$(\chi_R \circ \chi_+^{-1})(x) = \chi_R(x, \sqrt{1-x^2}) = \sqrt{1-x^2}$

Similar calculations hold for all the other overlapping charts. This manifold is usually denoted $\mathcal{M} = S_1$.

A cylinder is the Cartesian product of a line and a circle. In other words, we can create a cylinder by gluing a copy of a circle at each point along a line. If all these copies line up and don't twist around then we get a cylinder. The example that follows here illustrates a more general pattern, we can take a given manifold an paste a copy at each point along another manifold by using a Cartesian product.

Example 8.1.7. *Let $\mathcal{P} = \{(x,y,z) \in \mathbb{R}^3 \mid x^2+y^2=1\}$.*

1. *Let $V_+ = \{(x,y,z) \in \mathcal{P} \mid y > 0\} = dom(\chi_+)$ and define $\chi_+(x,y,z) = (x,z)$*
2. *Let $V_- = \{(x,y,z) \in \mathcal{P} \mid y < 0\} = dom(\chi_-)$ and define $\chi_-(x,y,z) = (x,z)$*
3. *Let $V_R = \{(x,y,z) \in \mathcal{P} \mid x > 0\} = dom(\chi_R)$ and define $\chi_R(x,y,z) = (y,z)$*
4. *Let $V_L = \{(x,y,z) \in \mathcal{P} \mid x < 0\} = dom(\chi_L)$ and define $\chi_L(x,y,z) = (y,z)$*

[4] meaning that if we adjoin the infinity of likewise compatible charts that defines a differentiable structure on $\mathcal{M}$

The set of charts $\mathcal{A} = \{(V_+, \chi_+), (V_-, \chi_-), (V_R, \chi_R), (V_L, \chi_L)\}$ forms an atlas on $\mathcal{P}$ which gives the cylinder a differentiable structure. It is not hard to show the transition functions are smooth on the image of the intersection of their respective domains. For example, $V_+ \cap V_R = W_{+R} = \{(x, y, z) \in \mathcal{P} \mid x, y > 0\}$, it's easy to calculate that $\chi_+^{-1}(x, z) = (x, \sqrt{1-x^2}, z)$ hence

$$(\chi_R \circ \chi_+^{-1})(x, z) = \chi_R(x, \sqrt{1-x^2}, z) = (\sqrt{1-x^2}, z)$$

for each $(x, z) \in \chi_R(W_{+R})$. Note $(x, z) \in \chi_R(W_{+R})$ implies $0 < x < 1$ hence it is clear the transition function is smooth. Similar calculations hold for all the other overlapping charts.

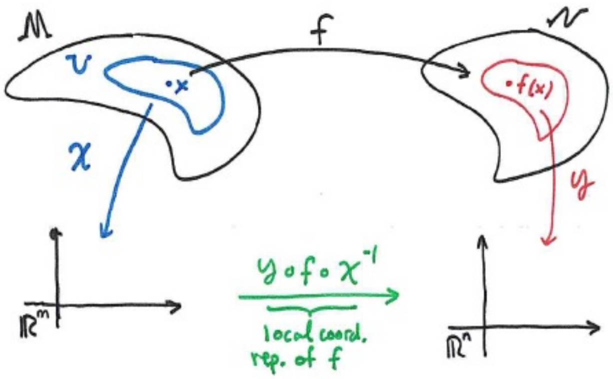
Generally, given two manifolds $\mathcal{M}$ and $\mathcal{N}$ we can construct $\mathcal{M} \times \mathcal{N}$ by taking the Cartesian product of the charts. Suppose $\chi_{\mathcal{M}} : V \subseteq \mathcal{M} \to U \subseteq \mathbb{R}^m$ and $\chi_{\mathcal{N}} : V' \subseteq \mathcal{N} \to U' \subseteq \mathbb{R}^n$ then you can define the product chart $\chi : V \times V' \to U \times U'$ as $\chi = \chi_{\mathcal{M}} \times \chi_{\mathcal{N}}$. The Cartesian product $\mathcal{M} \times \mathcal{N}$ together with all such product charts naturally is given the structure of an $(m+n)$-dimensional manifold. For example, in the preceding example we took $\mathcal{M} = S_1$ and $\mathcal{N} = \mathbb{R}$ to consruct $\mathcal{P} = S_1 \times \mathbb{R}$.

Example 8.1.8. *The 2-torus, or donut, is constructed as $T_2 = S_1 \times S_1$. The n-torus is constructed by taking the product of n-circles:*

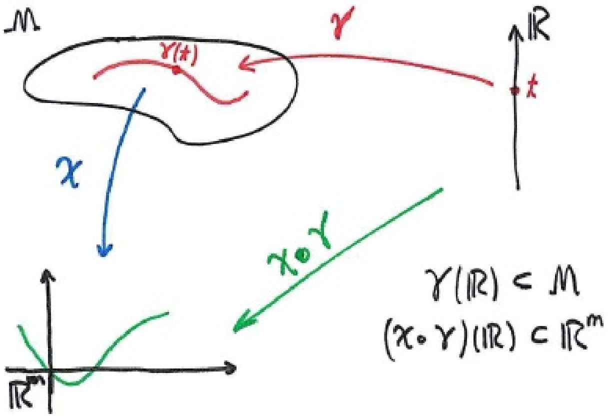
$$T_n = \underbrace{S_1 \times S_1 \times \cdots \times S_1}_{n \text{ copies}}$$

The atlas on this space can be obtained by simply taking the product of the S_1 charts n-times.

Recall from our study of linear algebra that vector space structure is greatly elucidated by the study of linear transformations. In our current context, the analogous objects are smooth maps. These are natural mappings between manifolds. In particular, suppose $\mathcal{M}$ is an m-fold and $\mathcal{N}$ is an n-fold with $f : U \subseteq \mathcal{M} \to \mathcal{N}$ is a function. Then, we say f is **smooth** iff all local coordinate representatives of f are smooth mappings from $\mathbb{R}^m$ to $\mathbb{R}^n$. See the diagram below:



This definition allows us to discuss smooth curves on manifolds. A curve $\gamma : I \subseteq \mathbb{R} \to \mathcal{M}$ is smooth iff $\chi \circ \gamma : I \to \mathbb{R}^n$ is a smooth curve in $\mathbb{R}^n$.



Finally, just for your information, if a bijection between manifolds is smooth with smooth inverse then the manifolds are said to be **diffeomorphic**. One fascinating result of recent mathematics is that $\mathbb{R}^4$ permits distinct differentiable structures in the sense that there does not exist a diffeomorphism between certain atlases[5]. Curiously, up to diffeomorphism, there is just one differentiable structure on $\mathbb{R}^n$ for $n \neq 4$. Classifying possible differentiable structures for a given point set is an interesting and ongoing problem.

8.2 vectors as derivations

To begin, let us define the set of locally smooth functions at $p \in \mathcal{M}$:

$$C^\infty(p) = \{f : \mathcal{M} \to \mathbb{R} \mid f \text{ is smooth on an open set containing } p\}$$

In particular, we suppose $f \in C^\infty(p)$ to mean there exists a patch $\phi : U \to V \subseteq \mathcal{M}$ such that f is smooth on V. Since we use Cartesian coordinates on $\mathbb{R}$ by convention it follows that $f : V \to \mathbb{R}$ smooth indicates the local coordinate representative $f \circ \phi : U \to \mathbb{R}$ is smooth (it has continuous partial derivatives of all orders).

Definition 8.2.1.

Suppose $X_p : C^\infty(p) \to \mathbb{R}$ is a linear transformation which satisfies the Leibniz rule then we say X_p is a **derivation** on $C^\infty(p)$. Moreover, we denote $X_p \in \mathcal{D}_p\mathcal{M}$ iff $X_p(f + cg) = X_p(f) + cX_p(g)$ and $X_p(fg) = f(p)X_p(g) + X_p(f)g(p)$ for all $f, g \in C^\infty(p)$ and $c \in \mathbb{R}$.

Example 8.2.2. *Let $\mathcal{M} = \mathbb{R}$ and consider $X_{t_o} = d/dt|_{t_o}$. Clearly X is a derivation on smooth functions near t_o.*

Example 8.2.3. *Consider $\mathcal{M} = \mathbb{R}^2$. Pick $p = (x_o, y_o)$ and define $X_p = \frac{\partial}{\partial x}\big|_p$ and $Y_p = \frac{\partial}{\partial y}\big|_p$. Once more it is clear that $X_p, Y_p \in \mathcal{D}(p)\mathbb{R}^2$. These derivations action is accomplished by partial differentiation followed by evaluation at p.*

Example 8.2.4. *Suppose $\mathcal{M} = \mathbb{R}^m$. Pick $p \in \mathbb{R}^m$ and define $X = \frac{\partial}{\partial x^j}\big|_p$. Clearly this is a derivation for any $j \in \mathbb{N}_m$.*

[5] there is even a whole book devoted to this exotic chapter of mathematics, see the *The Wild World of 4-Manifolds* by Alexandru Scorpan. This is on my list of books I "need" to buy.

Are the other types of derivations? Is the only thing a derivation is is a partial derivative operator? Before we can explore this question we need to define partial differentiation on a manifold. We should hope the definition is consistent with the langauge we already used in multivariate calculus (and the preceding pair of examples) and yet is also general enough to be stated on any abstract smooth manifold.

Definition 8.2.5.

Let $\mathcal{M}$ be a smooth m-dimensional manifold and let $\phi : U \to V$ be a local parametrization with $p \in V$. The **j-th coordinate function** $x^j : V \to \mathbb{R}$ is the j-component function of $\phi^{-1} : V \to U$. In other words:

$$\phi^{-1}(p) = x(p) = (x^1(p), x^2(p), \dots, x^m(p))$$

These x^j are **manifold coordinates**. In constrast, we will denote the standard Cartesian coordinates in $U \subseteq \mathbb{R}^m$ via u^j so a typical point has the form $(u^1, u^2, \dots, u^m)$ and viewed as functions $u^j : \mathbb{R}^m \to \mathbb{R}$ where $u^j(v) = e^j(v) = v^j$. We define the **partial derivative with respect to** x^j at p for $f \in C^\infty(p)$ as follows:

$$\frac{\partial f}{\partial x^j}(p) = \frac{\partial}{\partial u^j}\left[(f \circ \phi)(u)\right]\Big|_{u=\phi^{-1}(p)} = \frac{\partial}{\partial u^j}\left[f \circ x^{-1}\right]\Big|_{x(p)}.$$

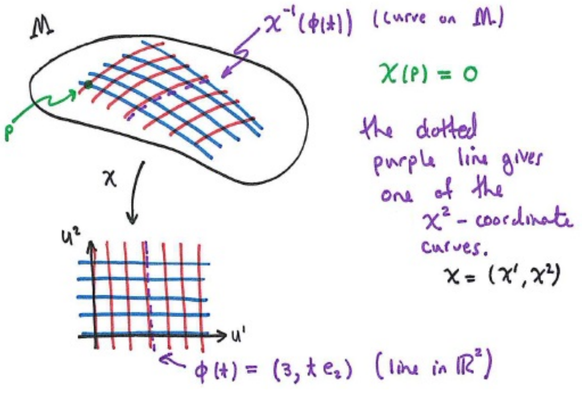
The idea of the definition is simply to take the function f with domain in $\mathcal{M}$ then pull it back to a function $f \circ x^{-1} : U \subseteq \mathbb{R}^m \to V \to \mathbb{R}$ on $\mathbb{R}^m$. Then we can take partial derivatives of $f \circ x^{-1}$ in the same way we did in multivariate calculus. In particular, the partial derivative w.r.t. u^j is calculated by:

$$\frac{\partial f}{\partial x^j}(p) = \frac{d}{dt}\left[\big(f \circ \phi\big)(x(p) + te_j)\right]_{t=0}$$

which is precisely the directional derivative of $f \circ x^{-1}$ in the j-direction at $x(p)$. In fact, Note

$$\big(f \circ \phi\big)(x(p) + te_j) = f(x^{-1}(x(p) + te_j)).$$

The curve $t \to x^{-1}(x(p) + te_j)$ is the curve on $\mathcal{M}$ through p where all coordinates are fixed except the j-coordinate. It is a *coordinate curve* on $\mathcal{M}$.


M
$x^{-1}(\phi(t))$ (curve on M)
$x(p) = 0$
p
the dotted purple line gives one of the x^2-coordinate curves.
$x = (x^1, x^2)$
x
u^2
u^1
$\phi(t) = (3, te_2)$ (line in $\mathbb{R}^2$)

Notice in the case that $\mathcal{M} = \mathbb{R}^m$ is given Cartesian coordinate $\phi = Id$ then $x^{-1} = Id$ as well and the $t \to x^{-1}(x(p) + te_j)$ reduces to $t \to p + te_j$ which is just the j-th coordinate curve through p on

$\mathbb{R}^m$. It follows that the partial derivative defined for manifolds naturally reduces to the ordinary partial derivative in the context of $\mathcal{M} = \mathbb{R}^m$ with Cartesian coordinates. The beautiful thing is that almost everything we know for ordinary partial derivatives equally well transfers to $\frac{\partial}{\partial x^j}\Big|_p$.

Theorem 8.2.6. *Partial differentiation on manifolds*

Let $\mathcal{M}$ be a smooth m-dimensional manifold with coordinates $x^1, x^2, \dots, x^m$ near p. Furthermore, suppose coordinates $y^1, y^2, \dots, y^m$ are also defined near p. Suppose $f, g \in C^\infty(p)$ and $c \in \mathbb{R}$ then:

1. $\frac{\partial}{\partial x^j}\Big|_p [f+g] = \frac{\partial f}{\partial x^j}\Big|_p + \frac{\partial g}{\partial x^j}\Big|_p$
2. $\frac{\partial}{\partial x^j}\Big|_p [cf] = c\frac{\partial f}{\partial x^j}\Big|_p$
3. $\frac{\partial}{\partial x^j}\Big|_p [fg] = f(p)\frac{\partial g}{\partial x^j}\Big|_p + \frac{\partial f}{\partial x^j}\Big|_p g(p)$
4. $\frac{\partial x^i}{\partial x^j}\Big|_p = \delta_{ij}$
5. $\sum_{k=1}^m \frac{\partial x^k}{\partial y^j}\Big|_p \frac{\partial y^i}{\partial x^k}\Big|_p = \delta_{ij}$
6. $\frac{\partial f}{\partial y^j}\Big|_p = \sum_{k=1}^m \frac{\partial x^k}{\partial y^j}\Big|_p \frac{\partial f}{\partial x^k}\Big|_p$

Proof: The proof of (1.) and (2.) follows from the calculation below:

$$\begin{aligned}
\frac{\partial(f+cg)}{\partial x^j}(p) &= \frac{\partial}{\partial u^j}\Big[(f+cg)\circ x^{-1}\Big]\Big|_{x(p)} \\
&= \frac{\partial}{\partial u^j}\Big[f\circ x^{-1} + cg\circ x^{-1}\Big]\Big|_{x(p)} \\
&= \frac{\partial}{\partial u^j}\Big[f\circ x^{-1}\Big]\Big|_{x(p)} + c\frac{\partial}{\partial u^j}\Big[g\circ x^{-1}\Big]\Big|_{x(p)} \\
&= \frac{\partial f}{\partial x^j}(p) + c\frac{\partial g}{\partial x^j}(p) \qquad (8.1)
\end{aligned}$$

The key in this argument is that composition $(f+cg)\circ x^{-1} = f\circ x^{-1} + cg\circ x^{-1}$ along side the linearity of the partial derivative. Item (3.) follows from the identity $(fg)\circ x^{-1} = (f\circ x^{-1})(g\circ x^{-1})$ in tandem with the product rule for a partial derivative on $\mathbb{R}^m$. The reader may be asked to complete the argument for (3.) in the homework. Continuing to (4.) we calculate from the definition:

$$\frac{\partial x^i}{\partial x^j}\Big|_p = \frac{\partial}{\partial u^j}\Big[(x^i\circ x^{-1})(u)\Big]\Big|_{x(p)} = \frac{\partial u^i}{\partial u^j}\Big|_{x(p)} = \delta_{ij}.$$

where the last equality is known from multivariate calculus. In invite the reader to prove it from the definition if unaware of this fact. Before we prove (5.) it helps to have a picture and a bit more notation in mind. Near the point p we have two coordinate charts $x : V \to U \subseteq \mathbb{R}^m$ and $y : V \to W \subseteq \mathbb{R}^m$, we take the chart domain V to be small enough so that both charts are defined. Denote Cartesian coordinates on U by $u^1, u^2, \dots, u^m$ and for W we likewise use Cartesian coordinates $w^1, w^2, \dots, w^m$. Let us denote patches ϕ, ψ as the inverses of these charts; $\phi^{-1} = x$ and $\psi^{-1} = y$. Transition functions $\psi^{-1}\circ\phi = y\circ x^{-1}$ are mappings from $U \subseteq \mathbb{R}^m$ to $W \subseteq \mathbb{R}^m$ and we note

$$\frac{\partial}{\partial u^j}\Big[(y^i\circ x^{-1})(u)\Big] = \frac{\partial y^i}{\partial x^j}$$

Likewise, the inverse transition functions $\phi^{-1} \circ \psi = x \circ y^{-1}$ are mappings from $W \subseteq \mathbb{R}^m$ to $U \subseteq \mathbb{R}^m$

$$\frac{\partial}{\partial w^j}\bigg[(x^i \circ y^{-1})(w)\bigg] = \frac{\partial x^i}{\partial y^j}$$

Recall that if $F, G : \mathbb{R}^m \to \mathbb{R}^m$ and $F \circ G = Id$ then $F'G' = I$ by the chainrule, hence $(F')^{-1} = G'$. Apply this general fact to the transition functions, we find their derivative matrices are inverses. Item (5.) follows. In matrix notation we item (5.) reads $\frac{\partial x}{\partial y}\frac{\partial y}{\partial x} = I$. Item (6.) follows from:

$$\begin{aligned}
\frac{\partial f}{\partial y^j}\bigg|_p &= \frac{\partial}{\partial w^j}\bigg[(f \circ y^{-1})(w)\bigg]\bigg|_{y(p)} \\
&= \frac{\partial}{\partial w^j}\bigg[(f \circ x^{-1} \circ x \circ y^{-1})(w)\bigg]\bigg|_{y(p)} \\
&= \frac{\partial}{\partial w^j}\bigg[\big(f \circ x^{-1}\big)(u^1(w), \dots, u^m(w))\bigg]\bigg|_{y(p)} \quad : \ \text{where } u^k(w) = (x \circ y^{-1})^k(w) \\
&= \sum_{k=1}^m \frac{\partial (x \circ y^{-1})^k}{\partial w^j}\bigg|_{y(p)} \frac{\partial (f \circ x^{-1})}{\partial u^k}\bigg|_{(x \circ y^{-1})(y(p))} \quad : \ \text{chain rule} \\
&= \sum_{k=1}^m \frac{\partial (x^k \circ y^{-1})}{\partial w^j}\bigg|_{y(p)} \frac{\partial (f \circ x^{-1})}{\partial u^k}\bigg|_{x(p)} \\
&= \sum_{k=1}^m \frac{\partial x^k}{\partial y^j}\Big|_p \frac{\partial f}{\partial x^k}\Big|_p
\end{aligned}$$

The key step was the multivariate chain rule. □

This theorem proves we can lift calculus on $\mathbb{R}^m$ to $\mathcal{M}$ in a natural manner. Moreover, we should note that items (1.), (2.) and (3.) together show $\frac{\partial}{\partial x^i}\big|_p$ is a derivation at p. Item (6.) should remind the reader of the contravariant vector discussion. Removing the f from the equation reveals that

$$\frac{\partial}{\partial y^j}\bigg|_p = \sum_{k=1}^m \frac{\partial x^k}{\partial y^j}\bigg|_p \frac{\partial}{\partial x^k}\bigg|_p$$

A notation convenient to the current discussion is that a contravariant transformation is $(p, v_x) \to (p, v_y)$ where $v_y = Pv_x$ and $P = (y \circ x^{-1})'(x(p)) = \frac{\partial y}{\partial x}\big|_{x(p)}$. Notice this is the inverse of what we see in (6.). This suggests that the partial derivatives change coordinates like as a basis for the tangent space. To complete this thought we need a few well-known propositions for derivations.

Proposition 8.2.7. *derivations on constant function gives zero.*

If $f \in C^\infty(p)$ is a constant function and $X_p \in \mathcal{D}_p\mathcal{M}$ then $X_p(f) = 0$.

Proof: Suppose $f(x) = c$ for all $x \in V$, define $g(x) = 1$ for all $x \in V$ and note $f = fg$ on V. Since X_p is a derivation is satisfies the Leibniz rule hence

$$X_p(f) = X_p(fg) = f(p)X_p(g) + X(f)g(p) = cX_p(g) + X_p(f) \quad \Rightarrow \quad cX_p(g) = 0.$$

Moreover, by homogeneity of X_p, note $cX_p(g) = X_p(cg) = X_p(f)$. Thus, $X_p(f) = 0$. □

Proposition 8.2.8.

If $f, g \in C^\infty(p)$ and $f(x) = g(x)$ for all $x \in V$ and $X_p \in \mathcal{D}_p\mathcal{M}$ then $X_p(f) = X_p(g)$.

Proof: Note that $f(x) = g(x)$ implies $h(x) = f(x) - g(x) = 0$ for all $x \in V$. Thus, the previous proposition yields $X_p(h) = 0$. Thus, $X_p(f - g) = 0$ and by linearity $X_p(f) - X_p(g) = 0$. The proposition follows. □

Proposition 8.2.9.

Suppose $X_p \in \mathcal{D}_p\mathcal{M}$ and x is a chart defined near p,

$$X_p = \sum_{j=1}^{m} X_p(x^j)\frac{\partial}{\partial x^j}\bigg|_p$$

Proof: this is a less trivial proposition. We need a standard lemma before we begin.

Lemma 8.2.10.

Let p be a point in smooth manifold $\mathcal{M}$ and let $f : \mathcal{M} \to \mathbb{R}$ be a smooth function. If $x : V \to U$ is a chart with $p \in V$ and $x(p) = 0$ then there exist smooth functions $g_j : \mathcal{M} \to \mathbb{R}$ whose values at p satisfy $g_j(p) = \frac{\partial f}{\partial x^j}(p)$. In addition, for all q near enough to p we have $f(q) = f(p) + \sum_{k=1}^{m} x^j(q)g_j(q)$

Proof: follows from proving a similar identity on $\mathbb{R}^m$ then lifting to the manifold. I leave this as a nontrivial exercise for the reader. This can be found in many texts, see Burns and Gidea page 29 for one source. It should be noted that the manifold must be smooth for this construction to hold. It turns out the set of derivations on a C^k-manifold forms an infinite-dimensional vector space over $\mathbb{R}$, see Lawrence Conlon's *Differentiable Manifolds* page 49. ▽

Consider $f \in C^\infty(p)$, and use the lemma, we assume $x(p) = 0$ and $g_j(p) = \frac{\partial f}{\partial x^j}(p)$:

$$\begin{aligned}
X_p(f) &= X_p\bigg(f(p) + \sum_{k=1}^{m} x^j(q)g_j(q)\bigg) \\
&= X_p(f(p)) + \sum_{k=1}^{m} X_p\big(x^j(q)g_j(q)\big) \\
&= \sum_{k=1}^{m}\Big[X_p(x^j)g_j(q) + x^j(p)X_p(g_j(q))\Big] \\
&= \sum_{k=1}^{m} X_p(x^j)\frac{\partial f}{\partial x^j}(p).
\end{aligned}$$

The calculation above holds for arbitrary $f \in C^\infty(p)$ hence the proposition follows. □

We've answered the question posed earlier in this section. It is true that every derivation of a manifold is simply a linear combination of partial derivatives. We can say more. The set of derivations at p naturally forms a vector space under the usual addition and scalar multiplication of operators: if $X_p, Y_p \in \mathcal{D}_p\mathcal{M}$ then we define $X_p + Y_p$ by $(X_p + Y_p)(f) = X_p(f) + Y_p(f)$ and cX_p by

$(cX_p)(f) = cX_p(f)$ for all $f, g \in C^\infty(p)$ and $c \in \mathbb{R}$. It is easy to show $\mathcal{D}_p\mathcal{M}$ is a vector space under these operations. Moreover, the preceding proposition shows that $\mathcal{D}_p\mathcal{M} = span\{\frac{\partial f}{\partial x^j}\big|_p\}_{j=1}^m$ hence $\mathcal{D}_p\mathcal{M}$ is an m-dimensional vector space[6].

Finally, let's examine coordinate change for derivations. Given two coordinate charts x, y at $p \in \mathcal{M}$ we have two ways to write the derivation X_p:

$$X_p = \sum_{j=1}^m X_p(x^j)\frac{\partial}{\partial x^j}\bigg|_p \qquad \text{or} \qquad X_p = \sum_{k=1}^m X_p(y^k)\frac{\partial}{\partial y^k}\bigg|_p$$

It is simple to connect these formulas. Whereas, for y-coordinates,

$$X_p(y^k) = \sum_{j=1}^m X_p(x^j)\frac{\partial y^k}{\partial x^j}\bigg|_p \tag{8.2}$$

This is the contravariant transformation rule. In contrast, recall $\frac{\partial}{\partial y^j}\big|_p = \sum_{k=1}^m \frac{\partial x^k}{\partial y^j}\big|_p \frac{\partial}{\partial x^k}\big|_p$. We should have anticipated this pattern since from the outset it is clear there is no coordinate dependence in the definition of a derivation.

Definition 8.2.11. *tangent space*

> We denote $T_p\mathcal{M} = derT_p\mathcal{M}$.

8.2.1 concerning the geometry of derivations

An obvious question we should ask:

> How are derivations related to tangent vectors geometrically?

To give a proper answer, we focus our attention to $\mathbb{R}^3$ and I follow Barret Oneil's phenomenal text on *Elementary Differential Geometry*, second edition. Consider a function $f : \mathbb{R}^3 \to \mathbb{R}$ and calculate the directional derivative in the $v = \langle a, b, c\rangle$-direction at the point p. We make no assumption that $||v|| = 1$, this is the same directional derivative for which we discussed the relation with the Frechet derivative in an earlier chapter. If f is smooth,

$$Df(p)(v) = v \bullet (\nabla f)(p) = a\frac{\partial f}{\partial x}(p) + b\frac{\partial f}{\partial y}(p) + \frac{\partial f}{\partial z}(p) = \left(a\frac{\partial}{\partial x} + b\frac{\partial}{\partial y} + c\frac{\partial}{\partial z}\right)\bigg|_p f$$

Therefore, we find the following interpretation of the derivation $X_p = (a\partial_x + b\partial_y + c\partial_z)\,|_p$:

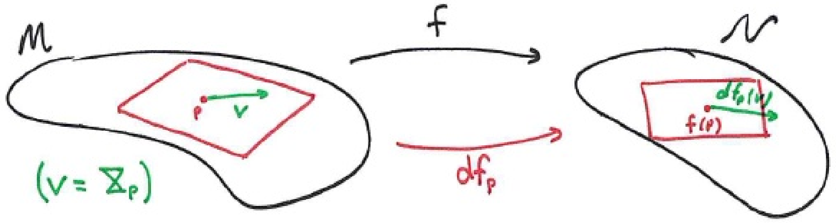
> When a derivation $X_p = (a\partial_x + b\partial_y + c\partial_z)\,|_p$ acts on a smooth function f it describes the rate at which f changes at p in the $\langle a, b, c\rangle$ direction. In other words, a derivation at p generates directional derivatives of functions at p.

Therefore, our work over the last few pages can be interpreted as abstracting the directional derivative to manifolds. In particular, the partial derivatives with respect to manifold coordinate x^j measure the rate of change of functions along the curve on the manifold which allows x^j to vary while all the other coordinates are held fixed.

[6] technically, we should show the coordinate derivations $\frac{\partial}{\partial x^j}\big|_p$ are linearly independent to make this conclusion. I don't suppose we've done that directly at this juncture. You might find this as a homework

8.3 differential for manifolds, the push-forward

In this section we generalize the concept of the differential to the context of manifolds. Recall that for $F : U \subseteq \mathbb{R}^m \to V \subseteq \mathbb{R}^n$ the differential $d_pF : \mathbb{R}^m \to \mathbb{R}^n$ was a linear transformation which best approximated the change in F near p. Notice that while the domain of F could be a mere subset of $\mathbb{R}^m$ the differential always took all of $\mathbb{R}^m$ as its domain. This suggests we should really think of the differential as a mapping which transports tangent vectors to U to tangent vectors at V. I find the following picture helpful at times, here I picture the tangent space as if it is attached to the point p on the manifold. Keep in mind this is an abstract picture:



Often d_pf is called the **push-forward** by f at p because it pushes tangent vectors in the same direction as the mapping transports points.

Definition 8.3.1. *differential for manifolds.*

Suppose $\mathcal{M}$ and $\mathcal{N}$ are smooth manifolds of dimension m and n respective. Furthermore, suppose $f : \mathcal{M} \to \mathcal{N}$ is a smooth mapping. We define $d_pf : T_p\mathcal{M} \to T_{f(p)}\mathcal{N}$ as follows: for each $X_p \in T_p\mathcal{M}$ and $g \in C^\infty(f(p))$

$$d_pf(X_p)(g) = X_p(g \circ f).$$

Notice that $g : dom(g) \subseteq \mathcal{N} \to \mathbb{R}$ and consequently $g \circ f : \mathcal{M} \to \mathcal{N} \to \mathbb{R}$ and it follows $g \circ f \in C^\infty(p)$ and it is natural to find $g \circ f$ in the domain of X_p. In addition, it is not hard to show $d_pf(X_p) \in \mathcal{D}_{f(p)}\mathcal{N}$. Observe:

1. $d_pf(X_p)(g+h) = X_p((g+h) \circ f) = X_p(g \circ f + h \circ f) = X_p(g \circ f) + X_p(h \circ f)$
2. $d_pf(X_p)(cg) = X_p((cg) \circ f) = X_p(cg \circ f)) = cX_p(g \circ f)) = cd_pf(X_p)(g)$

The proof of the Leibniz rule is similar. In this section we generalize the concept of the differential to the context of manifolds. Recall that for $F : U \subseteq \mathbb{R}^m \to V \subseteq \mathbb{R}^n$ the differential $d_pF : \mathbb{R}^m \to \mathbb{R}^n$ was a linear transformation which best approximated the change in F near p. Notice that while the domain of F could be a mere subset of $\mathbb{R}^m$ the differential always took all of $\mathbb{R}^m$ as its domain. This suggests we should really think of the differential as a mapping which transports tangent vectors to U to tangent vectors at V. Therefore, d_pf is called the **push-forward**.

Definition 8.3.2. *differential for manifolds.*

Suppose $\mathcal{M}$ and $\mathcal{N}$ are smooth manifolds of dimension m and n respective. Furthermore, suppose $f : \mathcal{M} \to \mathcal{N}$ is a smooth mapping. We define $d_pf : T_p\mathcal{M} \to T_{f(p)}\mathcal{N}$ as follows: for each $X_p \in T_p\mathcal{M}$ and $g \in C^\infty(f(p))$

$$d_pf(X_p)(g) = X_p(g \circ f).$$

Notice that $g : dom(g) \subseteq \mathcal{N} \to \mathbb{R}$ and consequently $g \circ f : \mathcal{M} \to \mathcal{N} \to \mathbb{R}$ and it follows $g \circ f \in C^\infty(p)$ and it is natural to find $g \circ f$ in the domain of X_p. In addition, it is not hard to show $d_pf(X_p) \in \mathcal{D}_{f(p)}\mathcal{N}$. Observe:

1. $d_pf(X_p)(g+h) = X_p((g+h)\circ f) = X_p(g\circ f + h \circ f) = X_p(g\circ f) + X_p(h\circ f)$
2. $d_pf(X_p)(cg) = X_p((cg)\circ f) = X_p(cg \circ f)) = cX_p(g\circ f)) = cd_pf(X_p)(g)$

The proof of the Leibniz rule is similar.

8.3.1 intuition for the push-forward

Perhaps it will help to see how the push-forward appears in calculus III. I'll omit the point-dependence to reduce some clutter. Consider $f : \mathbb{R}^2_{r\theta} \to \mathbb{R}^2_{xy}$ defined by $f(r,\theta) = (r\cos\theta, r\sin\theta)$. Consider, for g a smooth function on $\mathbb{R}^2_{xy}$, using the chain-rule,

$$df\left(\frac{\partial}{\partial r}\right)(g) = \frac{\partial}{\partial r}(g\circ f) = \frac{\partial g}{\partial x}\frac{\partial(x\circ f)}{\partial r} + \frac{\partial g}{\partial y}\frac{\partial(y\circ f)}{\partial r}$$

Note $x\circ f = r\cos\theta$ and $y \circ f = r\sin\theta$. Hence:

$$df\left(\frac{\partial}{\partial r}\right)(g) = \cos\theta\frac{\partial g}{\partial x} + \sin\theta\frac{\partial g}{\partial y}$$

This holds for all g hence we derive,

$$df\left(\frac{\partial}{\partial r}\right) = \frac{x}{\sqrt{x^2+y^2}}\frac{\partial}{\partial x} + \frac{y}{\sqrt{x^2+y^2}}\frac{\partial}{\partial y}$$

A similar calculation shows

$$df\left(\frac{\partial}{\partial \theta}\right) = -y\frac{\partial}{\partial x} + x\frac{\partial}{\partial y}$$

We could go through the reverse calculation for f^{-1} and derive that:

$$df^{-1}\left(\frac{\partial}{\partial x}\right) = \cos\theta\frac{\partial}{\partial r} - \frac{\sin\theta}{r}\frac{\partial}{\partial\theta} \qquad \text{and} \qquad df^{-1}\left(\frac{\partial}{\partial y}\right) = \sin\theta\frac{\partial}{\partial r} + \cos\theta\frac{\partial}{\partial\theta}$$

These formulas go to show that $\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0$ transforms to $\frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r} + \frac{1}{r^2}\frac{\partial^2 u}{\partial \theta^2} = 0$.

8.3.2 a pragmatic formula for the push-forward

In practice I quote the following result as the definition. However, Definition 8.3.2 is prefered by many for its coordinate independence. As we've seen in the intuition example, coordinates tend to arise in actual calculations.

Proposition 8.3.3.

If $F : M \to N$ is a differentiable mapping of manifolds and (x^i) are coordinates$(i = 1, \dots, m)$ on M and (y^j) are coordinates $(j = 1, \dots, n)$ on N which contain $p \in M$ and $F(p) \in N$ respective then:

$$(dF)_p\left(\sum_{i=1}^m X^i \frac{\partial}{\partial x^i}\bigg|_p\right) = \sum_{i=1}^m\sum_{j=1}^n X^i \frac{\partial(y^j \circ F)}{\partial x^i}\frac{\partial}{\partial y^j}\bigg|_{F(p)}.$$

Proof: Notice the absence of the f in this formula as compared to Defnition 8.3.2 $dF(X)(f) = X(f \circ F)$. To show equivalence, we can expand on our definition, we assume here that $X = \sum_{i=1}^m X^i \frac{\partial}{\partial x^i}\Big|_p$ thus:

$$X(f \circ F) = \sum_{i=1}^m X^i \frac{\partial}{\partial x^i}\bigg|_p (f \circ F)$$

note f is a function on N hence the chain rule gives:

$$X(f \circ F) = \sum_{i=1}^m X^i \sum_{j=1}^n \frac{\partial f}{\partial y^j}(F(p)) \frac{\partial (y^j \circ F)}{\partial x^i}$$

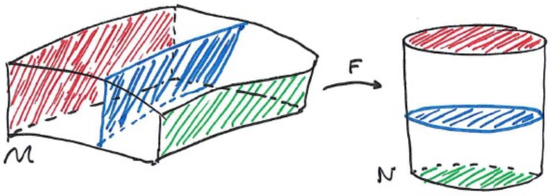
But, we can write this as:

$$X(f \circ F) = \left(\sum_{i=1}^m X^i \sum_{j=1}^n \frac{\partial (y^j \circ F)}{\partial x^i} \frac{\partial}{\partial y^j}\bigg|_{F(p)} \right)(f).$$

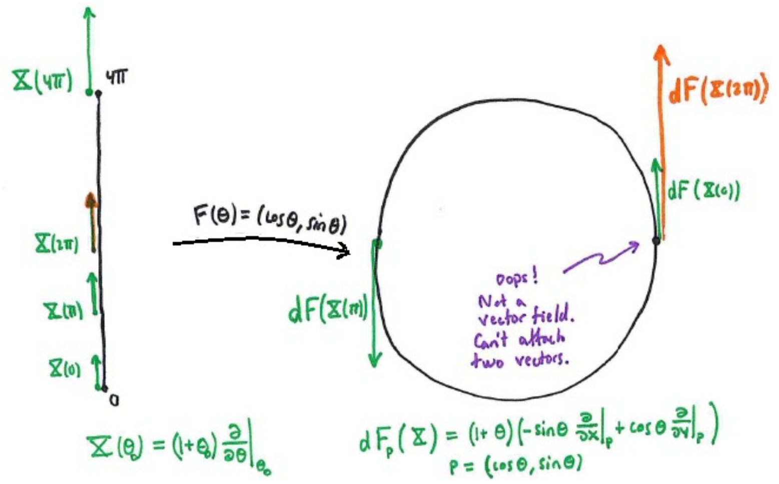
Therefore, as f is arbitrary, we have shown the claim of the proposition. □

Observe that the difference between Definition 8.3.2 and Proposition 8.3.3 is merely an application of the chain-rule.

The push-forward is more than just coordinate change. If we consider a mapping between spaces of disparate dimension then the push-forward captures something about the mapping in question and the domain and range spaces. For example, the existence of a nontrivial vector field on the whole of a manifold implies the existence of a foliation of the manifold.



If the mapping is a diffeomorphism then we expect it will carry the nontrivial vector field to the range space. However, if the mapping is not injective then there is no assurance a vector field even maps to a vector field. We could attach two vectors to a point in the range for a two-to-one map.For example, this mapping wraps around the circle and when it hits the circle the second time the vector pushed-forward does not match what was pushed forward the first time. It follows that push-forward of the vector field does not form a vector field in this case:



The pull-back (introducted in Section 8.7) is also an important tool to compare geometry of different spaces. We'll see in Section 8.10 how the pull-back even allows us to write a general formula for calculating the potential energy function of a flux-form of arbitrary degree. This captures the electric and magnetic potentials of electromagnetism and much more we have yet to discover experimentally. That said, the pull-back is formulated in terms of the push-forward we consider here thus the importance of the push-forward is hard to overstate.

Example 8.3.4. *Suppose $F : \mathbb{R}^2 \to \mathbb{R}^{2\times 2}$ is defined by*

$$F(x,y) = e^x \begin{bmatrix} \cos y & -\sin y \\ \sin y & \cos y \end{bmatrix}$$

Let $\mathbb{R}^2$ have the usual (x,y)-coordinate chart and let Z^{ij} defined by

$$Z^{ij}(A) = A_{ij}$$

for $\mathbb{R}^{2\times 2}$ form the global coordinate chart for 2×2 matrices. Let us calculate the push-forward of the coordinate vector $\partial_x|_p$:

$$dF_p(\partial_x|_p) = \sum_{i,j=1}^{2} \frac{\partial(Z^{ij}\circ F)}{\partial x} \frac{\partial}{\partial Z^{ij}}\bigg|_{F(p)} \quad \& \quad dF_p(\partial_y|_p) = \sum_{i,j=1}^{2} \frac{\partial(Z^{ij}\circ F)}{\partial y} \frac{\partial}{\partial Z^{ij}}\bigg|_{F(p)}$$

Observe that:

$$Z^{11}(F(x,y)) = e^x \cos y = Z^{22}(F(x,y)), \qquad Z^{21}(F(x,y)) = e^x \sin y = -Z^{12}(F(x,y)).$$

From which we derive,

$$dF_p(\partial_x) = e^x \cos y(\partial_{11} + \partial_{22}) + e^x \sin y(\partial_{21} - \partial_{12})$$

$$dF_p(\partial_y) = -e^x \sin y(\partial_{11} + \partial_{22}) + e^x \cos y(\partial_{21} - \partial_{12})$$

Here ∂_x, ∂_y are at $p \in \mathbb{R}^2$ whereas ∂_{ij} is at $F(p) \in \mathbb{R}^{2\times 2}$. However, these are constant vector fields so the point-dependence is not too interesting.

8.4 cotangent space

The tangent space to a smooth manifold $\mathcal{M}$ is a vector space of derivations and we denote it by $T_p\mathcal{M}$. The dual space to this vector space is called the **cotangent space** and the typical elements are called **covectors**.

Definition 8.4.1. *cotangent space $T_p\mathcal{M}^*$*

Suppose $\mathcal{M}$ is a smooth manifold and $T_p\mathcal{M}$ is the tangent space at $p \in \mathcal{M}$. We define, $T_p\mathcal{M}^* = \{\alpha_p : T_p\mathcal{M} \to \mathbb{R} \mid \alpha \textit{ is linear}\}$.

If x is a local coordinate chart at p and $\frac{\partial}{\partial x^1}\big|_p, \frac{\partial}{\partial x^2}\big|_p, \dots, \frac{\partial}{\partial x^m}\big|_p$ is a basis for $T_p\mathcal{M}$ then we denote the dual basis $d_px^1, d_px^2, \dots, d_px^m$ where $d_px^i\big(\partial_k\big|_p\big) = \delta_{ik}$. Moreover, if α is a covector at p then[7]:

$$\alpha = \sum_{k=1}^{m} \alpha_k dx^k$$

where $\alpha_k = \alpha\left(\frac{\partial}{\partial x^k}\Big|_p\right)$ and dx^k is a short-hand for d_px^k. We should understand that covectors are defined at a point even if the point is not explicitly indicated in a particular context. This does lead to some ambiguity in the same way that the careless identification of the function f and it's value $f(x)$ does throughout calculus. That said, an abbreviated notation is often important to help us see through more difficult patterns without getting distracted by the minutia of the problem.

You might worry the notation used for the differential and our current notation for the dual basis of covectors is not consistent. After all, we have two rather different meanings for d_px^k at this time:

1. $x^k : V \to \mathbb{R}$ is a smooth function hence $d_px^k : T_p\mathcal{M} \to T_{x^k(p)}\mathbb{R}$
   is defined as a *push-forward*, $d_px^k(X_p)(g) = X_p(g \circ x^k)$
2. $d_px^k : T_p\mathcal{M} \to \mathbb{R}$ where $d_px^k\big(\partial_j\big|_p\big) = \delta_{jk}$

It is customary to identify $T_{x^k(p)}\mathbb{R}$ with $\mathbb{R}$ hence there is no trouble. Let us examine how the dual-basis condition can be *derived* for the differential, suppose $g : \mathbb{R} \to \mathbb{R}$ hence $g \circ x^k : V \to \mathbb{R}$,

$$d_px^k\left(\frac{\partial}{\partial x^j}\bigg|_p\right)(g) = \frac{\partial}{\partial x^j}\bigg|_p (g(x^k)) = \underbrace{\frac{\partial x^k}{\partial x^j}\bigg|_p \frac{dg}{dt}\bigg|_{x^k(p)}}_{chain\ rule} = \delta_{jk}\frac{d}{dt}\bigg|_{x^k(p)}(\ g\ ) = \delta_{jk}g$$

Where, we've made the identification $1 = \frac{d}{dt}\Big|_{x^k(p)}$ (which is the nut and bolts of $T_{x^k(p)}\mathbb{R} = \mathbb{R}$) and hence have the beautiful identity:

$$\boxed{d_px^k\left(\frac{\partial}{\partial x^j}\bigg|_p\right) = \delta_{jk}.}$$

In contrast, there is no need to derive this for case (2.) since in that context this serves as the definition for the object. Personally, I find the multiple interpretations of objects in manifold theory is one of the most difficult aspects of the theory. On the other hand, the notation is really neat

[7] we explained this for an arbitrary vector space V and its dual V^* in a previous chapter, we simply apply those results once more here in the particular context $V = T_p\mathcal{M}$

once you understand how subtly it assumes many theorems. You should understand the notation we enjoy at this time is the result of generations of mathematical thought. Following a similar derivation for an arbitrary vector $X_p \in T_p\mathcal{M}$ and $f : \mathcal{M} \to \mathbb{R}$ we find

$$\boxed{d_pf(X_p) = X_p(f)}$$

This notation is completely consistent with the *total differential* as commonly discussed in multivariate calculus. Recall that if $f : \mathbb{R}^m \to \mathbb{R}$ then we defined

$$df = \frac{\partial f}{\partial x^1}dx^1 + \frac{\partial f}{\partial x^2}dx^2 + \cdots + \frac{\partial f}{\partial x^m}dx^m.$$

Notice that the j-th component of df is simply $\frac{\partial f}{\partial x^j}$. Notice that the identity $d_pf(X_p) = X_p(f)$ gives us the same component if we simply evaluate the covector d_pf on the coordinate basis $\frac{\partial}{\partial x^j}\Big|_p$,

$$d_pf\left(\frac{\partial}{\partial x^j}\Big|_p\right) = \frac{\partial f}{\partial x^j}\Big|_p$$

8.5 differential forms

In this section we apply the results of Chapter 7 on exterior algebra to the vector space $V = T_pM$ and its dual space $V^* = T_pM^*$. In short, this involves making the identifications:

$$e_i = \frac{\partial}{\partial x^i}\Big|_p \qquad \text{and} \qquad e^j = d_px^j$$

however, we often omit the p-dependence of d_px^j and just write dx^j. Observe that:

$$dx^j(\partial_i|_p) = \partial_i(x^j) = \delta_{ij}$$

Therefore, the coordinate basis $\{\partial_1|_p, \dots, \partial_m|_p\}$ for T_pM is indeed dual to the differentials of the coordinates $\{dx^1, \dots, dx^m\}$ basis for the cotangent space T_pM^*.

In contrast to Chapter 7 we have a point-dependence to our basis. We establish the following terminology: for a manifold M and $p \in M$,

(0.) $p \mapsto \mathbb{R}$ is a 0-form, or function on M

(1.) $p \mapsto dx^j$ is a 1-form on M

(2.) $p \mapsto dx^i \wedge dx^j$ is a 2-form on M

(3.) $p \mapsto dx^i \wedge dx^j \wedge dx^k$ is a 3-form on M

(4.) $p \mapsto dx^{i_1} \wedge \cdots \wedge dx^{i_k}$ is a k-form on M

Generally a k-form is formed from taking sums of the basic differential forms given above with coefficients which are smooth functions. [8]

(1.) a one-form $\alpha = \sum_{j=1}^m \alpha_j dx^j$ has smooth coefficient functions α_j.

[8] Technically, this means the exterior algebra of differential forms is a module over the ring of smooth functions. However, the exterior algebra at a point is a vector space.

(2.) a two-form $\beta = \sum_{i,j=1}^{m} \beta_{ij} dx^i \wedge dx^j$ has smooth coefficient functions β_{ij}.

(3.) a k-form $\gamma = \sum_{i_1,\dots,i_k}^{n} \gamma_{i_1,\dots,i_k} dx^{i_1} \wedge \cdots \wedge dx^{i_k}$ has smooth coefficient functions $\gamma_{i_1,\dots,i_k}$.

The algebra of differential forms follows the same rules as the exterior algebra we previously discussed. However, instead of having scalars as numbers we now consider scalars as functions. This comment is made explicit in the theorem to follow:

Theorem 8.5.1.

If α is a p-form, β is a k-form, and γ is a l-form on M then

1. $\alpha \wedge (\beta \wedge \gamma) = (\alpha \wedge \beta) \wedge \gamma$
2. $\alpha \wedge \beta = (-1)^{pk} (\beta \wedge \alpha)$
3. $\alpha \wedge (a\beta + b\gamma) = a(\alpha \wedge \beta) + b(\alpha \wedge \gamma) \quad a, b \in \mathbb{R}$

Notice that in $\mathbb{R}^3$ the set of differential forms

$$\mathcal{B} = \{1, dx, dy, dz, dy \wedge dz, dz \wedge dx, dx \wedge dy, dx \wedge dy \wedge dz\}$$

is a basis of the space of differential forms in the sense that every form on $\mathbb{R}^3$ is a linear combination of the forms in $\mathcal{B}$ with smooth real-valued functions on $\mathbb{R}^3$ as coefficients.

Example 8.5.2. *Let $\alpha = fdx + gdy$ and let $\beta = 3dx + dz$ where f, g are functions. Find $\alpha \wedge \beta$, write the answer in terms of the basis defined in the Remark above,*

$$\begin{aligned} \alpha \wedge \beta &= (fdx + gdy) \wedge (3dx + dz) \\ &= fdx \wedge (3dx + dz) + gdy \wedge (3dx + dz) \\ &= 3fdx \wedge dx + fdx \wedge dz + 3gdy \wedge dx + gdy \wedge dz \\ &= -gdy \wedge dz - fdz \wedge dx - 3gdx \wedge dy \end{aligned} \tag{8.3}$$

Example 8.5.3. Top form: *Let $\alpha = dx \wedge dy \wedge dz$ and let β be any other form with degree $p > 0$. We argue that $\alpha \wedge \beta = 0$. Notice that if $p > 0$ then there must be at least one differential inside β so if that differential is dx^k we can rewrite $\beta = dx^k \wedge \gamma$ for some γ. Then consider,*

$$\alpha \wedge \beta = dx \wedge dy \wedge dz \wedge dx^k \wedge \gamma \tag{8.4}$$

now k has to be either $1, 2$ or 3 therefore we will have dx^k repeated, thus the wedge product will be zero. (can you prove this?).

The proposition below and its proof are included here to remind the reader on the structure of the $\otimes$ and $\wedge$ products and components. One distinction, the components are functions now whereas they were scalars in the previous chapter.

Proposition 8.5.4.

If ω is a p-form in an n-dimensional space is written in the coordinate coframe $dx^1, \dots, dx^n$ at p then the components of ω are given by evaluation on the coordinate frame $\partial_1, \dots, \partial_n$. at p.

Proof: Suppose ω has component functions $\omega_{i_1 i_2 \dots i_p}$ with respect to the tensor basis $dx^{i_1} \otimes \cdots \otimes dx^{i_p}$ for type $(0, p)$ tensors.

$$\omega = \sum_{i_1 i_2 \dots i_p}^{n} \omega_{i_1 i_2 \dots i_p} dx^{i_1} \otimes dx^{i_2} \otimes \cdots \otimes dx^{i_p}$$

the functions $\omega_{i_1 i_2 \dots i_p}$ are called the **tensor components** of ω. Consider evaluation of ω on a p-tuple of coordinate vector fields,

$$\begin{aligned}\omega(\partial_{j_1}, \partial_{j_2} \dots \partial_{j_p}) &= \sum_{i_1, i_2, \dots, i_p = 1}^{n} \omega_{i_1 i_2 \dots i_p} dx^{i_1} \otimes dx^{i_2} \otimes \cdots \otimes dx^{i_p}(\partial_{j_1}, \partial_{j_2}, \dots, \partial_{j_p}) \\ &= \sum_{i_1, i_2, \dots, i_p = 1}^{n} \omega_{i_1 i_2 \dots i_p} \delta_{i_1 j_1} \delta_{i_2 j_2} \cdots \delta_{i_p j_p} \\ &= \omega_{j_1 j_2 \dots j_p}\end{aligned}$$

If ω is a p-form that indicated ω is a completely antisymmetric tensor and by a calculation similar to those near Equation 7.10 we can express ω by a sum over all p-forms (this is not a basis expansion)

$$\omega = \sum_{i_1, i_2, \dots, i_p = 1}^{n} \frac{1}{p!} \omega_{i_1 i_2 \dots i_p} dx^{i_1} \wedge dx^{i_2} \wedge \cdots \wedge dx^{i_p}.$$

But, we just found $\omega(\partial_{j_1}, \partial_{j_2} \dots \partial_{j_p}) = \omega_{i_1 i_2 \dots i_p}$ hence:

$$\omega = \sum_{i_1 i_2 \dots i_p}^{n} \frac{1}{p!} \omega(\partial_{i_1}, \partial_{i_2} \dots \partial_{i_p}) dx^{i_1} \wedge dx^{i_2} \wedge \cdots \wedge dx^{i_p}.$$

which proves the assertion of the proposition. □

Note, if we work with an expansion of linearly independent p-vectors then we can write the conclusion of the proposition:

$$\omega = \sum_{i_1 < i_2 < \cdots < i_p}^{n} \omega(\partial_{i_1}, \partial_{i_2} \dots \partial_{i_p}) dx^{i_1} \wedge dx^{i_2} \wedge \cdots \wedge dx^{i_p}.$$

8.6 the exterior derivative

The operation $\wedge$ depends only on the values of the forms point by point. We define an operator d on differential forms which depends not only on the value of the differential form at a point but on its value in an entire neighborhood of the point. Thus if β ia k-form then to define $d\beta$ at a point p we need to know not only the value of β at p but we also need to know its value at every q in a neighborhood of p.

You might note the derivative below does not directly involve the construction of differential forms from tensors. Also, the rule given below is easily taken as a starting point for formal calculations. In other words, even if you don't understand the nuts and bolts of manifold theory you can still calculate with differential forms. In the same sense that highschool students "do" calculus, you can "do" differential form calculations. I don't believe this is a futile exercise so long as you understand you have more to learn. Which is not to say we don't know some things!

Definition 8.6.1. *the exterior derivative.*

If β is a k-form and $x = (x^1, x^2, \cdots, x^n)$ is a chart and $\beta = \sum_I \frac{1}{k!}\beta_I dx^I$ and we define a $(k+1)$-form $d\beta$ to be the form

$$d\beta = \sum_I \frac{1}{k!} d\beta_I \wedge dx^I \ .$$

Where $d\beta_I$ is defined as it was in calculus III,

$$d\beta_I = \sum_{j=1}^{n} \frac{\partial \beta_I}{\partial x_j} dx^j.$$

Note that $d\beta_I$ is well-defined as

$$\beta_I = \beta_{i_1 i_2 \cdots i_k}$$

is just a real-valued function on $dom(x)$. The definition in an expanded form is given by

$$d_p\beta = \frac{1}{k!}\sum_{i_1=1}^{n}\sum_{i_2=1}^{n}\cdots\sum_{i_k=1}^{n}(d_p\beta_{i_1 i_2\cdots i_k}) \wedge d_p x^{i_1} \wedge \cdots \wedge d_p x^{i_k}$$

where

$$\beta_q = \frac{1}{k!}\sum_{i_1=1}^{n}\sum_{i_2=1}^{n}\cdots\sum_{i_k=1}^{n}\beta_{i_1 i_2\cdots i_k}(q) d_q x^{i_1} \wedge \cdots \wedge d_p x^{i_k} \ .$$

Consequently we see that for each k the operator d maps $\wedge^k(M)$ into $\wedge^{k+1}(M)$. Also:

Theorem 8.6.2. *properties of the exterior derivative.*

If $\alpha \in \wedge^k(M)$, $\beta \in \wedge^l(M)$ and $a, b \in \mathbf{R}$ then

1. $d(a\alpha + b\beta) = a(d\alpha) + b(d\beta)$
2. $d(\alpha \wedge \beta) = (d\alpha \wedge \beta) + (-1)^k(\alpha \wedge d\beta)$
3. $d(d\alpha) = 0$

Remark 8.6.3.

Warning: I use Einstein's repeated index notation in the proof that follows. In fact, it's a bit worse, I use I to denote a **multi-index**. This means a repeated I indicates an implicit sum over all increasing strings of indices of a particular length. This is just a brief notation to sum over the basis of coordinate p-forms. Indeed, from this point on in the notes there is occasional use of Einstein's convention.

Proof: The proof of (1) is obvious. To prove (2), let $x = (x^1, \cdots, x^n)$ be a chart on M then suppose $\alpha = \alpha_I dx^I$ and $\beta = \beta_J dx^J$

$$\begin{aligned} d(\alpha \wedge \beta) = d(\alpha_I \beta_J) \wedge dx^I \wedge dx^J \quad &= (\alpha_I d\beta_J + \beta_J d\alpha_I) \wedge dx^I \wedge dx^J \\ &= \alpha_I(d\beta_J \wedge dx^I \wedge dx^J) \\ &\quad\ + \beta_J(d\alpha_I \wedge dx^I \wedge dx^J) \\ &= \alpha_I(dx^I \wedge (-1)^k(d\beta_J \wedge dx^J)) \\ &\quad\ + \beta_J((d\alpha_I \wedge dx^I) \wedge dx^J) \\ &= (\alpha \wedge (-1)^k d\beta) + \beta_J(d\alpha \wedge dx^J) \\ &= d\alpha \wedge \beta + (-1)^k(\alpha \wedge d\beta) \ . \end{aligned}$$

To prove (3.) we could resort to a beautiful tensor calculation (see Equation 8.11) or:

$$d\alpha = d\alpha_I \wedge dx^I$$

hence

$$d(d\alpha) = d(d\alpha_I \wedge dx^I) = d(d\alpha_I) \wedge dx^I + \alpha_I \wedge d(dx^I).$$

Notice $d(dx^I) = d(dx^{i_1} \wedge \cdots \wedge dx^{i_k}) = d(1) \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_k} = 0$. Therefore, we have reduced the problem to showing $d(d\alpha_I) = 0$ for a function α_I. I leave that problem to the reader. $\Box$.

8.6.1 exterior derivatives on $\mathbb{R}^3$

We begin by noting that vector fields may correspond either to a one-form or to a two-form.

Definition 8.6.4. *dictionary of vectors verses forms on $\mathbb{R}^3$.*

Let $\vec{A} = (A^1, A^2, A^3)$ denote a vector field in $\mathbb{R}^3$. Define then,

$$\omega_A = \delta_{ij} A^i dx^j = A_i dx^i$$

which we will call the **work-form** of $\vec{A}$. Also define

$$\Phi_A = \frac{1}{2}\delta_{ik} A^k \epsilon_{ijk}(dx^i \wedge dx^j) = \frac{1}{2} A_i \epsilon_{ijk}(dx^i \wedge dx^j)$$

which we will call the **flux-form** of $\vec{A}$.

If you accept the primacy of differential forms, then you can see that vector calculus confuses two separate objects. Apparently there are two types of vector fields. In fact, if you have studied coordinate change for vector fields deeply then you will encounter the qualifiers **axial** or **polar** vector fields. Those fields which are axial correspond directly to two-forms whereas those correspondant to one-forms are called polar. Note, the magnetic field is axial whereas the electric field is polar.

Example 8.6.5. Gradient: *Consider three-dimensional Euclidean space. Let $f : \mathbb{R}^3 \to \mathbb{R}$ then*

$$df = \frac{\partial f}{\partial x^i} dx^i = \omega_{\nabla f}$$

which gives the one-form corresponding to ∇f.

Example 8.6.6. Curl: *Consider three-dimensional Euclidean space. Let $\vec{F}$ be a vector field and let $\omega_F = F_i dx^i$ be the corresponding one-form then*

$$\begin{aligned} d\omega_F &= dF_i \wedge dx^i \\ &= \partial_j F_i dx^j \wedge dx^i \\ &= \partial_x F_y dx \wedge dy + \partial_y F_x dy \wedge dx + \partial_z F_x dz \wedge dx + \partial_x F_z dx \wedge dz + \partial_y F_z dy \wedge dz + \partial_z F_y dz \wedge dy \\ &= (\partial_x F_y - \partial_y F_x) dx \wedge dy + (\partial_z F_x - \partial_x F_z) dz \wedge dx + (\partial_y F_z - \partial_z F_y) dy \wedge dz \\ &= \Phi_{\nabla \times \vec{F}}. \end{aligned}$$

Thus we recover the curl.

Example 8.6.7. Divergence: *Consider three-dimensional Euclidean space. Let $\vec{G}$ be a vector field and let $\Phi_G = \frac{1}{2}\epsilon_{ijk} G_i dx^j \wedge dx^k$ be the corresponding two-form then*

$$\begin{aligned} d\Phi_G &= d(\tfrac{1}{2}\epsilon_{ijk} G_i) \wedge dx^j \wedge dx^k \\ &= \tfrac{1}{2}\epsilon_{ijk}(\partial_m G_i) dx^m \wedge dx^j \wedge dx^k \\ &= \tfrac{1}{2}\epsilon_{ijk}(\partial_m G_i)\epsilon_{mjk} dx \wedge dy \wedge dz \\ &= \tfrac{1}{2} 2\delta_{im}(\partial_m G_i) dx \wedge dy \wedge dz \\ &= \partial_i G_i dx \wedge dy \wedge dz \\ &= (\nabla \cdot \vec{G}) dx \wedge dy \wedge dz \end{aligned}$$

and in this way we recover the divergence.

8.6.2 coordinate independence of exterior derivative

The Einstein summation convention is used in this section and throughout the remainder of this chapter, please feel free to email me if it confuses you somewhere. When an index is repeated in a single summand it is implicitly assumed there is a sum over all values of that index

It must be shown that this definition is independent of the chart used to define $d\beta$. Suppose for example, that

$$\beta_q = \overline{\beta}_J(q)(d_q\overline{x}^{j_1} \wedge \cdots \wedge d_q\overline{x}^{j_k})$$

for all q in the domain of a chart $(\overline{x}^1, \overline{x}^2, \cdots \overline{x}^n)$ where

$$dom(x) \cap dom(\overline{x}), \neq \emptyset \ .$$

We assume, of course that the coefficients $\{\overline{\beta}_J(q)\}$ are skew-symmetric in J for all q. We will have defined $d\beta$ in this chart by

$$d\beta = d\overline{\beta}_J \wedge d\overline{x}^J \ .$$

We need to show that $d_p\overline{\beta}_J \wedge d_p\overline{x}^J = d_p\beta_I \wedge d_p x^I$ for all $p \in dom(x) \cap dom(\overline{x})$ if this definition is to be meaningful. Since β is given to be a well-defined form we know

$$\beta_I(p) d_p x^I = \beta_p = \overline{\beta}_J(p) d_p\overline{x}^J \ .$$

Using the identities

$$d\overline{x}^j = \frac{\partial \overline{x}^j}{\partial x^i} dx^i$$

we have

$$\beta_I dx^I = \overline{\beta}_J \frac{\partial \overline{x}^{j_1}}{\partial x^{i_1}} \frac{\partial \overline{x}^{j_k}}{\partial x^{i_k}} \cdots \frac{\partial \overline{x}^{j_k}}{\partial x^{i_k}} \ dx^I$$

so that

$$\beta_I = \overline{\beta}_J \left(\frac{\partial \overline{x}^{j_1}}{\partial x^{i_1}} \frac{\partial \overline{x}^{j_2}}{\partial x^{i_2}} \cdots \frac{\partial \overline{x}^{j_k}}{\partial x^{i_k}} \right) \ .$$

Consequently,

$$\begin{aligned} d\beta_J \wedge dx^J = \tfrac{\partial \beta_J}{\partial x^\lambda}(dx^\lambda \wedge dx^J) \quad &= \tfrac{\partial}{\partial x^\lambda}[\overline{\beta}_I \left(\tfrac{\partial \overline{x}^{i_1}}{\partial x^{j_1}} \cdots \tfrac{\partial \overline{x}^{i_k}}{\partial x^{j_k}} \right)](dx^\lambda \wedge dx^J) \\ &\overset{*}{=} \tfrac{\partial \overline{\beta}_I}{\partial x^\lambda} \left(\tfrac{\partial \overline{x}^{i_1}}{\partial x^{j_1}} \cdots \tfrac{\partial \overline{x}^{i_k}}{\partial x^{j_k}} \right)(dx^\lambda \wedge dx^J) \\ &\quad + \overline{\beta}_I \textstyle\sum_r \left(\tfrac{\partial \overline{x}^{i_1}}{\partial x^{j_1}} \cdots \tfrac{\partial^2 \overline{x}^{i_r}}{\partial x^\lambda \partial x^{j_r}} \cdots \tfrac{\partial \overline{x}^{i_k}}{\partial x^{j_k}} \right)(dx^\lambda \wedge dx^J) \\ &= \tfrac{\partial \overline{\beta}_I}{\partial x^\lambda} dx^\lambda \wedge \left(\tfrac{\partial \overline{x}^{i_1}}{\partial x^{j_1}} dx^{j_1} \right) \wedge \cdots \wedge \left(\tfrac{\partial \overline{x}^{i_k}}{\partial x^{j_k}} dx^{j_k} \right) \\ &= \tfrac{\partial \overline{\beta}_I}{\partial \overline{x}^p} \left[\left(\tfrac{\partial \overline{x}^p}{\partial x^\lambda} dx^\lambda \right) \wedge d\overline{x}^{i_1} \wedge \cdots \wedge d\overline{x}^{i_k} \right] \\ &= d\overline{\beta}_I \wedge d\overline{x}^I \end{aligned}$$

where in (*) the sum $\sum\limits_r$ is zero since:

$$\frac{\partial^2 \overline{x}^{i_r}}{\partial x^\lambda \partial x^{j_r}}(dx^\lambda \wedge dx^J) = \pm \frac{\partial^2 \overline{x}^{i_r}}{\partial x^\lambda \partial x^{j_r}} [(dx^\lambda \wedge dx^{j_r}) \wedge dx^{j_1} \wedge \cdots \wedge \widehat{dx^{j_r}} \wedge \cdots dx^{j_k}] = 0.$$

It follows that $d\beta$ is independent of the coordinates used to define it.

8.7 the pull-back

Another important operation one can perform on differential forms is the "pull-back" of a form under a map[9]. The definition is constructed in large part by a sneaky application of the push-forward (aka differential) discussed in the preceding chapter. If you are impatient for intuition, skip ahead to the end of this section and later return to the careful calculations at the outset.

Definition 8.7.1. *pull-back of a differential form.*

If $f : \mathcal{M} \to \mathcal{N}$ is a smooth map and $\omega \in \wedge^k(N)$ then $f^*\omega$ is the form on $\mathcal{M}$ defined by

$$(f^*\omega)_p(X_1, \cdots, X_k) = \omega_{f(p)}(d_pf(X_1), d_pf(X_2), \cdots, d_pf(X_k)) \ .$$

for each $p \in \mathcal{M}$ and $X_1, X_2, \dots, X_k \in T_p\mathcal{M}$. Moreover, in the case $k = 0$ we have a smooth function $\omega : \mathcal{N} \to \mathbb{R}$ and the pull-back is accomplished by composition $(f^*\omega)(p) = (\omega \circ f)(p)$ for all $p \in \mathcal{M}$.

This operation is linear on forms and commutes with the wedge product and exterior derivative:

Theorem 8.7.2. *properties of the pull-back.*

If $f : M \to N$ is a C^1-map and $\omega \in \wedge^k(N)$, $\tau \in \wedge^l(N)$ then

1. $f^*(a\omega + b\tau) = a(f^*\omega) + b(f^*\tau) \qquad a, b \in \mathbb{R}$
2. $f^*(\omega \wedge \tau) = f^*\omega \wedge (f^*\tau)$
3. $f^*(d\omega) = d(f^*\omega)$

Proof: The proof of (1) is clear. We now prove (2).

$$\begin{aligned} f^*(\omega\wedge\tau)]_p(X_1, \cdots, X_{k+l}) &= (\omega\wedge\tau)_{f(p)}(d_pf(X_1), \cdots, d_pf(X_{k+l})) \\ &= \textstyle\sum_\sigma(\text{sgn}\sigma)(\omega\otimes\tau)_{f(p)}(d_pf(X_{\sigma_1}), \cdots, d_pf(X_{\sigma(k+l)})) \\ &= \textstyle\sum_\sigma \text{sgn}(\sigma)\omega(d_pf(X_{\sigma(1)}), \cdots d_pf(X_{\sigma(k)}))\tau(df(X_{\sigma(k+1)}\cdots dfX_{\sigma(k+l)}) \\ &= \textstyle\sum_\sigma \text{sgn}(\sigma)(f^*\omega)_p(X_{\sigma(1)}, \cdots, X_{\sigma(k)})(f^*\tau_p)(X_{\sigma(k+1)}, \cdots, X_{\sigma(k+l)}) \\ &= [(f^*\omega)\wedge(f^*\tau)]_p(X_1, X_2, \cdots, X_{(k+l)}) \end{aligned}$$

Finally we prove (3).

$$\begin{aligned} f^*(d\omega)]_p(X_1, X_2\cdots, X_{k+1}) &= (d\omega)_{f(p)}(df(X_1), \cdots df(X_{k+1})) \\ &= (d\omega_I\wedge dx^I)_{f(p)}(df(X_1), \cdots, df(X_{k+1})) \\ &= \left(\frac{\partial\omega_I}{\partial x^\lambda}\bigg|_{f(p)}\right)(dx^\lambda\wedge dx^I)_{f(p)}(df(X_1), \cdots, df(X_{k+1})) \\ &= \left(\frac{\partial\omega_I}{\partial x^\lambda}\bigg|_{f(p)}\right)[d_p(x^\lambda\circ f)\wedge d_p(x^I\circ f)](X_1, \cdots, X_{k+1}) \\ &= [d(\omega_I\circ f)\wedge d(x^I\circ f)](X_1, \cdots, X_{k+1}) \\ &= d[(\omega_I\circ f)_p d_p(x^I\circ f)](X_1, \cdots, X_{k+1}) \\ &= d(f^*\omega)_p(X_1, \cdots, X_{k+1}) \ . \end{aligned}$$

The theorem follows. □.

[9] thanks to my advisor R.O. Fulp for the arguments that follow

We saw that one important application of the push-forward was to change coordinates for a given vector. Similar comments apply here. If we wish to change coordinates on a given differential form then we can use the pull-back. However, given the direction of the operation we need to use the inverse coordinate transformation to pull forms forward. Let me mirror the example from the last chapter for forms on $\mathbb{R}^2$. We wish to convert from r, θ to x, y notation.

Example 8.7.3. *Suppose $F : \mathbb{R}^2_{r,\theta} \rightarrow \mathbb{R}^2_{x,y}$ is the polar coordinate transformation. In particular,*

$$F(r, \theta) = (r\cos\theta, r\sin\theta)$$

The inverse transformation, at least for appropriate angles, is given by

$$F^{-1}(x, y) = \bigg(\sqrt{x^2 + y^2},\ \tan^{-1}(y/x) \bigg).$$

Let calculate the pull-back of dr under F^{-1}: let $p = F^{-1}(q)$

$$F^{-1*}(dr)_q = d_p r(\partial_x|p) d_p x + d_p r(\partial_y|p) d_p y$$

Again, drop the annoying point-dependence to see this clearly:

$$F^{-1*}(dr) = dr(\partial_x)dx + dr(\partial_y)dy = \frac{\partial r}{\partial x}dx + \frac{\partial r}{\partial y}dy$$

Likewise,

$$F^{-1*}(d\theta) = d\theta(\partial_x)dx + d\theta(\partial_y)dy = \frac{\partial \theta}{\partial x}dx + \frac{\partial \theta}{\partial y}dy$$

Note that $r = \sqrt{x^2 + y^2}$ and $\theta = \tan^{-1}(y/x)$ have the following partial derivatives:

$$\frac{\partial r}{\partial x} = \frac{x}{\sqrt{x^2+y^2}} = \frac{x}{r} \qquad \text{and} \qquad \frac{\partial r}{\partial y} = \frac{y}{\sqrt{x^2+y^2}} = \frac{y}{r}$$

$$\frac{\partial \theta}{\partial x} = \frac{-y}{x^2+y^2} = \frac{-y}{r^2} \qquad \text{and} \qquad \frac{\partial \theta}{\partial y} = \frac{x}{x^2+y^2} = \frac{x}{r^2}$$

Of course the expressions using r are pretty, but to make the point, we have changed into x, y-notation via the pull-back of the inverse transformation as advertised. We find:

$$\boxed{dr = \frac{xdx + ydy}{\sqrt{x^2+y^2}} \qquad \text{and} \qquad d\theta = \frac{-ydx + xdy}{x^2+y^2}.}$$

Once again we have found results with the pull-back that we might previously have chalked up to substitution in multivariate calculus. That's often the idea behind an application of the pull-back. It's just a formal langauge to be precise about a substitution. It takes us past simple symbol pushing and gives us a rigorous notation for substutions. It's a bit more than that though, the *substitution* we discuss here takes us from one space to another in general.

8.7.1 intuitive computation of pull-backs

Consider the definition below: how can we understand this computationally?

$$(f^*\omega)_p(X_1, \cdots, X_k) = \omega_{f(p)}(d_pf(X_1), d_pf(X_2), \cdots, d_pf(X_k))\ .$$

In particular, let us consider $f(u,v) = (x,y,z)$. Furthermore, let us consider a one-form to begin $\omega = adx + bdy + cdz$ the pull-back will be formed by a suitable linear combination of du and dv. We calculate,

$$df(\partial_u) = \frac{\partial x}{\partial u}\frac{\partial}{\partial x} + \frac{\partial y}{\partial u}\frac{\partial}{\partial y} + \frac{\partial z}{\partial u}\frac{\partial}{\partial z} \qquad \& \qquad df(\partial_v) = \frac{\partial x}{\partial v}\frac{\partial}{\partial x} + \frac{\partial y}{\partial v}\frac{\partial}{\partial y} + \frac{\partial z}{\partial v}\frac{\partial}{\partial z}$$

Therefore, as $\omega = adx + bdy + cdz$,

$$\omega(df(\partial_u)) = a\frac{\partial x}{\partial u} + b\frac{\partial y}{\partial u} + c\frac{\partial z}{\partial u} \qquad \& \qquad \omega(df(\partial_v)) = a\frac{\partial x}{\partial v} + b\frac{\partial y}{\partial v} + c\frac{\partial z}{\partial v}$$

From which we deduce: by Proposition 8.5.4,

$$\begin{aligned} f^*\omega &= \left[a\frac{\partial x}{\partial u} + b\frac{\partial y}{\partial u} + c\frac{\partial z}{\partial u}\right] du + \left[a\frac{\partial x}{\partial v} + b\frac{\partial y}{\partial v} + c\frac{\partial z}{\partial v}\right] dv \\ &= a\underbrace{\left[\frac{\partial x}{\partial u}du + \frac{\partial x}{\partial v}dv\right]}_{dx \text{ for } x=x(u,v)} + b\underbrace{\left[\frac{\partial y}{\partial u}du + \frac{\partial y}{\partial v}dv\right]}_{dy \text{ for } y=y(u,v)} + c\underbrace{\left[\frac{\partial z}{\partial u}du + \frac{\partial z}{\partial v}dv\right]}_{dz \text{ for } z=z(u,v)}. \end{aligned} \tag{8.5}$$

This shows the pull-back of ω is accomplished by taking $\omega = adx + bdy + cdz$ and substituting the total differentials of $x = x(u,v), y = y(u,v)$ and $z = z(u,v)$ into dx, dy and dz respectively.

Continuing in our somewhat special context, consider $\Omega = ady \wedge dz + bdz \wedge dx + cdx \wedge dy$ and use Theorem 8.7.2 to simplify our life. We already worked out the formula for the one-form case so we can use it to find the intuitive formula for the two-form:

$$\begin{aligned} f^*\Omega &= f^*\left[ady \wedge dz + bdz \wedge dx + cdx \wedge dy\right] \\ &= f^*(ady) \wedge f^*dz + f^*(bdz) \wedge f^*dx + f^*(cdx) \wedge f^*dy \quad \text{(by Theorem 8.7.2)} \\ &= a\left[\frac{\partial y}{\partial u}du + \frac{\partial y}{\partial v}dv\right] \wedge \left[\frac{\partial z}{\partial u}du + \frac{\partial z}{\partial v}dv\right] + b\left[\frac{\partial z}{\partial u}du + \frac{\partial z}{\partial v}dv\right] \wedge \left[\frac{\partial x}{\partial u}du + \frac{\partial x}{\partial v}dv\right] \\ &\quad + c\left[\frac{\partial x}{\partial u}du + \frac{\partial x}{\partial v}dv\right] \wedge \left[\frac{\partial y}{\partial u}du + \frac{\partial y}{\partial v}dv\right] \quad \text{(by Equation 8.5)} \end{aligned}$$

The formula above simplifies considerable as certain terms vanish due to $du \wedge du = 0$ and $dv \wedge dv = 0$ and $dv \wedge du = -du \wedge dv$. Furthermore, the following standard notation[10]

$$\frac{\partial(y,z)}{\partial(u,v)} = \frac{\partial y}{\partial u}\frac{\partial z}{\partial v} - \frac{\partial y}{\partial v}\frac{\partial z}{\partial u} \quad \& \quad \frac{\partial(x,y)}{\partial(u,v)} = \frac{\partial x}{\partial u}\frac{\partial y}{\partial v} - \frac{\partial x}{\partial v}\frac{\partial y}{\partial u}, \quad \& \quad \frac{\partial(z,x)}{\partial(u,v)} = \frac{\partial z}{\partial u}\frac{\partial x}{\partial v} - \frac{\partial z}{\partial v}\frac{\partial x}{\partial u}.$$

Generally, $\frac{\partial(x_i,x_j)}{\partial(u,v)} = \det\begin{bmatrix} \partial_u x_i & \partial_v x_i \\ \partial_u x_j & \partial_v x_j \end{bmatrix}$. Ugly equations aside, this allows us to express the pull-back of the two-form in a way which is easy to remember and is readily generalized.

$$f^*\Omega = \left[a\frac{\partial(y,z)}{\partial(u,v)} + b\frac{\partial(z,x)}{\partial(u,v)} + c\frac{\partial(x,y)}{\partial(u,v)}\right] du \wedge dv$$

I should mention, throughout this calculation I have suppressed the point-dependence. Technically, a, b, c in the expression above should be understood as $a \circ f, b \circ f, c \circ f$ which are functions of u, v. The pull-back once complete trades a form in the range coordinates (x,y,z) for a new form in the

[10] H.M. Edwards Advanced Calculus a Differential Forms approach spends dozens of pages explaining this through intuitive geometric arguments which we do not pursue here for brevity

domain coordinates (u, v).

The discussion thus far is somewhat limiting since the domain only supports a nontrivial two-form. To continue, let's consider pull-backs for the mapping $G : \mathbb{R}^3_{uvw} \to \mathbb{R}^4_{txyz}$. In this case, we can consider the pull-back of

$$\gamma = adt \wedge dx \wedge dz + bdx \wedge dy \wedge dz + cdy \wedge dz \wedge dt + mdz \wedge dt \wedge dy$$

and we will obtain:

$$G^*\gamma = \left[a\frac{\partial(t,x,y)}{\partial(u,v,w)} + b\frac{\partial(x,y,z)}{\partial(u,v,w)} + c\frac{\partial(y,z,t)}{\partial(u,v,w)} + m\frac{\partial(z,t,x)}{\partial(u,v,w)}\right] du \wedge dv \wedge dw$$

Where we define the coefficients by the natural generalization of the second-order case. Multiplying out the wedge product of the pull-back of the three one-forms will produce the signs of the following determinants

$$\frac{\partial(x_i,x_j,x_k)}{\partial(u,v,w)} = \det\begin{bmatrix} \partial_u x_i & \partial_v x_i & \partial_w x_i \\ \partial_u x_j & \partial_v x_j & \partial_w x_j \\ \partial_u x_k & \partial_v x_k & \partial_w x_k \end{bmatrix}$$

In words, the pull-back is formed with coefficients taken from determinants of submatrices of the Jacobian matrix of G. That probably doesn't help you much. Perhaps the following two-form pull-back with respect to G will inspire: an arbitrary two-form on $\mathbb{R}^4_{txyz}$ has the form:

$$\Xi = A_{01}dt \wedge dx + A_{02}dt \wedge dy + A_{03}dt \wedge dz + A_{23}dy \wedge dz + A_{31}dz \wedge dx + A_{12}dx \wedge dy$$

when we pull this back under G to u, v, w-space we will get a form in $dv \wedge dw$, $dw \wedge du$ and $du \wedge dv$. The coefficients are again obtained by appropriate determinants of submatrices of the Jacobian: in particular: $G^*\Xi =$

$$\begin{aligned}
&\left[A_{01}\frac{\partial(t,x)}{\partial(v,w)} + A_{02}\frac{\partial(t,y)}{\partial(v,w)} + A_{03}\frac{\partial(t,z)}{\partial(v,w)} + A_{23}\frac{\partial(y,z)}{\partial(v,w)} + A_{31}\frac{\partial(z,x)}{\partial(v,w)} + A_{12}\frac{\partial(x,y)}{\partial(v,w)}\right] dv \wedge dw \\
&+ \left[A_{01}\frac{\partial(t,x)}{\partial(w,u)} + A_{02}\frac{\partial(t,y)}{\partial(w,u)} + A_{03}\frac{\partial(t,z)}{\partial(w,u)} + A_{23}\frac{\partial(y,z)}{\partial(w,u)} + A_{31}\frac{\partial(z,x)}{\partial(w,u)} + A_{12}\frac{\partial(x,y)}{\partial(w,u)}\right] dw \wedge du \\
&+ \left[A_{01}\frac{\partial(t,x)}{\partial(u,v)} + A_{02}\frac{\partial(t,y)}{\partial(u,v)} + A_{03}\frac{\partial(t,z)}{\partial(u,v)} + A_{23}\frac{\partial(y,z)}{\partial(u,v)} + A_{31}\frac{\partial(z,x)}{\partial(u,v)} + A_{12}\frac{\partial(x,y)}{\partial(u,v)}\right] du \wedge dv.
\end{aligned}$$

The arbitrary case is perhaps a bit tiresome. Let's consider a particular example

$$\Gamma = (t^2 + x^2)dy \wedge dz$$

and the mapping $G(u, v, w) = (u - 1, v^2 + 1, w^3, uvw) = (t, x, y, z)$. Observe:

$$dt = du, \quad dx = 2vdv, \quad dy = 3w^2dw, \quad dz = vwdu + uwdv + uvdw$$

the pull-back of Γ under G is thus:

$$\begin{aligned}
G^*\Gamma &= \big((u-1)^2 + (v^2+1)^2\big)3w^2dw \wedge [vwdu + uwdv + uvdw] \\
&= \big((u-1)^2 + (v^2+1)^2\big)3w^2\,[vw\,dw \wedge du + uw\,dw \wedge dv] \\
&= A_{23}(u,v)(\underbrace{-3w^2uw}_{\frac{\partial(y,z)}{\partial(v,w)}})dv \wedge dw + A_{23}(u,v)(\underbrace{3w^2vw}_{\frac{\partial(y,z)}{\partial(w,u)}})dw \wedge du \qquad (8.6)
\end{aligned}$$

In the above I let $A_{23}(u,v) = (u-1)^2 + (v^2+1)^2$. Let's check the coefficients of Equation 8.6 in view of the general claim of $G^*\Xi$. Are the coefficients in fact the Jacobians indicated? We calculate:

$$\frac{\partial(y,z)}{\partial(v,w)} = \det\left[\begin{array}{cc} y_v & y_w \\ z_v & z_w \end{array}\right] = \det\left[\begin{array}{cc} 0 & 3w^2 \\ uw & uv \end{array}\right] = -3w^2uw$$

and

$$\frac{\partial(y,z)}{\partial(w,u)} = \det\left[\begin{array}{cc} y_w & y_u \\ z_w & z_u \end{array}\right] = \det\left[\begin{array}{cc} 3w^2 & 0 \\ uv & vw \end{array}\right] = 3w^2vw.$$

Well, that's a relief. We can either approach the calculation of a pull-back in terms of Jacobian coefficients or we can just plug in the pull-backs of each coordinate function and multiply it out. I suppose both techinques have their place. Moreover, when faced with many abstract questions I much prefer Definition 8.7.1. The thought of sorting through the proof of Theorem 8.7.2 in Jacobian notation seems hopeless.

8.7.2 implicit function theorem in view of forms

Previously we saw this theorem without the benefit of the calculus of differential forms, I hope this brings new light to the topic. This is the implicit function theorem as presented in H.M. Edwards' *Advanced Calculus: a Differential Forms Approach.*

Proposition 8.7.4.

Let us consider for $i = 1, \dots m$,

$$y_i = f_i(x_1, \dots, x_n)$$

where f_i are continuously differentiable near a point $(\bar{x}_1, \dots, \bar{x}_n)$. Furthermore, denote $\bar{y}_i = f_i(\bar{x}_1, \dots, \bar{x}_n)$ for $i = 1, \dots m$. Suppose the following conditions are satisfied:

1. $\frac{\partial(y_1,\dots,y_r)}{\partial(x_1,\dots,x_r)} \neq 0$ at $(\bar{x}_1, \dots, \bar{x}_n)$
2. the pull-back of every k-form for $k > r$ is identically zero near $(\bar{x}_1, \dots, \bar{x}_n)$

Then near $(\bar{x}_1, \dots, \bar{x}_n, \bar{y}_1, \dots, \bar{y}_m) \in \mathbb{R}^{n+m}$ there exist differentiable functions g_i, h_i which solve $y_i = f_i(x_1, \dots, x_n)$ provided we suppose:

$$x_i = g_i(y_1, \dots, y_r, x_{r+1}, \dots, x_n) \qquad (\text{for } i = 1, \dots, r)$$

$$y_i = h_i(y_1, \dots, y_r) \qquad (\text{for } i = r+1, \dots, m).$$

Condition (1.) implies the rank is at least r then condition (2.) assures us the rank is at most r hence the rank is r. If the graph $y = f(x)$ is viewed as $G(x,y) = y - f(x)$ then it is seen as a mapping from $\mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$. We saw before that if $G(\bar{x},\bar{y}) = 0$ and $rank(G'(\bar{x},\bar{y})) = m$ then the solution set of $G(x,y) = 0$ near $(\bar{x},\bar{y})$ forms an (n)-dimensional manifold. This generalizes that in a sense because it allows for redundant conditions as indicated by $r < m$. On the other hand, this is less general than the previous implicit function theorem as it assumes a linear-dependence on the m-variables $y_1, \dots, y_m$ for the m equations defining the level set in $\mathbb{R}^m \times \mathbb{R}^n$. The formulation of this Theorem in Edward's text falls inline with the general pattern of that text to emphasize equations over mappings. It is both the strength and weakness of the text in my opinion. The next two examples attempt to work within the confines of the notation put forth in the theorem above, then we transition to examples where we informally apply the theorem.

Example 8.7.5. *Just to remind us of the counting: if* $y_1 = f(x_1, x_2) = x_1 + x_2 - 1$ *then*

$$G(x_1, x_2, y_1) = y_1 - (x_1 + x_2 - 1) = -x_1 - x_2 + y_1 + 1 = 0$$

is a plane in (x_1, x_2, y_1)*-space. Consider,* $(\bar{x}_1, \bar{x}_2) = (2, 3)$ *then* $\bar{y}_1 = 2 + 3 - 1 = 4$*. A point on this plane is* $(2, 3, 4)$ *we can use* (x_1, x_2) *as coordinates hence I expect* $r = 1$*. Observe, as* $G : \mathbb{R}^3 \to \mathbb{R}$ *we consider pull-back of the one-form* dt *in the range,*

$$G^* dt = -dx_1 - dx_2 + dy_1$$

apparently, in our context here:

$$\frac{\partial(y_1, \dots, y_r)}{\partial(x_1, \dots, x_r)} = \frac{\partial t}{\partial x_1} = -1 \neq 0$$

and clearly there is no 2*-form which pulls back under* G *since the only two-form in* t*-space is trivial. We can explicitly write* $x_1 = g_1(y_1, x_2) = -x_2 + y_1 + 1$ *and there is no* h_i *since* $r + 1 = 2 > m$*.*

Example 8.7.6. *Let us attempt another example to unravel the meaning of the Implicit Function Theorem. Suppose* $y_1 = x_1^2 - x_2^2$ *and* $y_2 = x_1^2 - x_2^2$ *clearly these are redundant. The theorem should deal with this. Let* $G(x_1, x_2, y_2, y_2) = (x_1^2 - x_2^2, x_1^2 - x_2^2)$ *so we might hope the solution set is* $4 - 2 = 2$*-dimensional however, the rank of* G' *is usually* 1 *so the solution set is likely* 3*-dimensional. No need for guessing, let's work it out.*

$$dy_1 = 2x_1 dx_1 - 2x_2 dx_2 = dy_2$$

Hence, if we attempt $r = 2$ *then consider:*

$$\frac{\partial(y_1, y_2)}{\partial(x_1, x_2)} = det \begin{bmatrix} 2x_1 & -2x_2 \\ 2x_1 & -2x_2 \end{bmatrix} = 0.$$

So, we are forced to look at $r = 1$ *for which* $\frac{\partial y_1}{\partial x_1} = 2x_1 \neq 0$ *for* $x_1 \neq 0$*. Consider then a point such that* $x_1 \neq 0$ *and see if we can derive the functions such that* $x_1 = g_1(y_1, x_2)$ *and* $y_2 = h_2(y_1)$*. Assume* $x_1, x_2 > 0$ *for convenience (we could replicate this calculation in other quadrants with appropriate signs adjoined),*

$$x_1 = \sqrt{y_1 + x_2^2} = g_1(y_1, x_2), \qquad \& \qquad y_2 = y_1 = h_1(y_1).$$

I'm mostly interested in this theorem to gain more geometric insight as to what the pull-back means. So, a better way to look at the last example is just to emphasize

$$dy_1 = 2x_1 dx_1 - 2x_2 dx_2 = dy_2$$

shows the pull-back of the basic one-forms can only give a one-form and where the coefficients are zero we cannot force the corresponding coordinate to be dependent. For example, $2x_1 dx_1$ is trivial when $x_1 = 0$ hence x_1 cannot be solved for as a function of the remaining variables. This is the computational essence of the theorem. Ideally, I want to get us to the point of calculating without reliance on the Jacobians. Towards that end, let's consider an example for which $m = r$ hence the h function is not needed and we can focus on the pull-back geometry.

Example 8.7.7. *For which variables can we solve the following system (possibly subject some condition). Let $F(x,y,z) = (s,t)$ defined as follows:*

$$\begin{aligned} s &= x + z - 1 \\ t &= y + z - 2 \end{aligned} \tag{8.7}$$

then clearly $F^*ds = dx + dz$ and $F^*dt = dy + dz$.

$$F^*(ds \wedge dt) = (dx + dz) \wedge (dy + dz) = \underbrace{dx \wedge dy}_{(1.)} + \underbrace{dx \wedge dz}_{(2.)} + \underbrace{dz \wedge dy}_{(3.)}$$

Therefore, by (1.), we can solve for x, y as functions of s, t, z. Or, by (2.), we could solve for x, z as functions of s, t, y. Or, by (3.), we could solve for z, y as functions of s, t, x. The fact that the rank is maximal is implicit within the fact that $ds \wedge dt$ is the top-form in the range. In contrast,

$$F^*ds = dx + dz$$

does not mean that I can solve Equation 8.7 by solving for x as a function of s, t, y, z. Of course, $x = s - z + 1$ solves the first equation, but the second equation is not contrained by the solution for x what so over. Conversely, we can solve for $z = t - y + 2$ but we cannot also solve for $z = s - x + 1$. The coefficients of 1 in $F^*ds = dx + dz$ are not applicable to the Implicit Function Theorem because this form is not the highest degree which pulls-back nontrivially. That role falls to $ds \wedge dt$ here.

Example 8.7.8. *For which variables can we solve the following system (possibly subject some condition). Let $F(u,v) = (x,y,z)$ defined as follows:*

$$\begin{aligned} x &= u^2 + v^2 \\ y &= u^2 - v^2 \\ z &= uv \end{aligned} \tag{8.8}$$

then

$$\begin{aligned} dx &= 2udu + 2vdv \\ dy &= 2udu - 2vdv \\ dz &= vdu + udv \end{aligned}$$

We can calculate,

$$dy \wedge dz = (2udu - 2vdv) \wedge (vdu + udv) = 2(u^2 + v^2)du \wedge dv$$

$$dz \wedge dx = (vdu + udv) \wedge (2udu + 2vdv) = 2(v^2 - u^2)du \wedge dv$$

$$dx \wedge dy = (2udu + 2vdv) \wedge (2udu - 2vdv) = -8uv\, du \wedge dv$$

What does this tell us geometrically? Well, notice that the top-form $dx \wedge dy \wedge dz$ pulls-back to zero since the two-form $du \wedge dv$ is the top-form in the domain. Therefore, $r = 2$, that is F has rank 2.

8.8 integration of forms

The general strategy is generally as follows:

(i) there is a natural way to calculate the integral of a k-form on a subset of $\mathbb{R}^k$

(ii) given a k-form on a manifold we can locally pull it back to a subset of $\mathbb{R}^k$ provided the manifold is an oriented[11] k-dimensional and thus by the previous idea we have an integral.

(iii) globally we should expect that we can add the results from various local charts and arrive at a total value for the manifold, assuming of course the integral in each chart is finite.

We will only investigate items (*i.*) and (*ii.*) in these notes. There are many other excellent texts which take great effort to carefully expand on point (iii.) and I do not wish to replicate that effort here. You can read Edwards and see about *pavings*, or read Munkres' where he has at least 100 pages devoted to the careful study of multivariate integration. I do not get into those topics in my notes because we simply do not have sufficient analytical power to do them justice. I would encourage the student interested in deeper ideas of integration to find time to talk to Dr. Skoumbourdis, he has thought a long time about these matters and he really understands integration in a way we dare not cover in the calculus sequence. You really should have that conversation after you've taken real analysis and have gained a better sense of what analysis' purpose is in mathematics. That said, what we do cover in this section and the next is fascinating whether or not we understand all the analytical underpinnings of the subject!

8.8.1 integration of k-form on $\mathbb{R}^k$

Note that on $U \subseteq \mathbb{R}^k$ a k-form α is the top form thus there exists some smooth function f on U such that $\alpha_x = f(x)dx^1 \wedge dx^2 \wedge \cdots \wedge dx^k$ for all $x \in U$. If D is a subset of U then we define the integral of α over D via the corresponding intgral of k-variables in $\mathbb{R}^k$. In particular,

$$\int_D \alpha = \int_D f(x) d^k x$$

where on the r.h.s. the symbol $d^k x$ is meant to denote the usual integral of k-variables on $\mathbb{R}^k$. It is sometimes convenient to write such an integral as:

$$\int_D f(x) d^k x = \int_D f(x) dx^1 dx^2 \cdots dx^k$$

but, to be more careful, the integration of f over D is a quantity which is independent of the particular order in which the variables on $\mathbb{R}^k$ are assigned. On the other hand, the order of the variables in the formula for α certainly can introuduce signs. Note

$$\alpha_x = -f(x)dx^2 \wedge dx^1 \wedge \cdots \wedge dx^k.$$

How can we reasonably maintain the integral proposed above? Well, the answer is to make a convention that we write the form to match the standard orientation of $\mathbb{R}^k$. The standard **orientation** of $\mathbb{R}^k$ is given by $Vol_k = dx^1 \wedge dx^2 \wedge \cdots \wedge dx^k$. If the given form is written $\alpha_x = f(x)Vol_k$ then we define $\int_D \alpha = \int_D f(x)d^k x$. Since it is always possible to write a k-form as a function multiplying

[11] we will discuss this as the section progresses

Vol_k on $\mathbb{R}^k$ this definition suffices to cover all possible k-forms. For example, if $\alpha_x = f(x)dx$ on some subset $D = [a,b]$ of $\mathbb{R}$,

$$\int_D \alpha = \int_D f(x)dx = \int_a^b f(x)dx.$$

Or, if $\alpha_{(x,y)} = f(x,y)dx \wedge dy$ then for D a aubset of $\mathbb{R}^2$,

$$\int_D \alpha = \int_D f(x,y)dxdy = \int_D fdA.$$

If $\alpha_{(x,y,z)} = f(x,y,z)dx \wedge dy \wedge dz$ then for D a aubset of $\mathbb{R}^3$,

$$\int_D \alpha = \int_D f(x,y,z)dxdydz = \int_D fdV.$$

In practice we tend to break the integrals above down into an interated integral thanks to Fubini's theorems. The integrals $\int_D fdA$ and $\int_D fdV$ are not in and of themselves dependent on orientation. However the set D may be oriented the value of those integrals are the same for a fixed function f. The orientation dependence of the form integral is completely wrapped up in our rule that the form must be written as a multiple of the volume form on the given space.

8.8.2 orientations and submanifolds

Given a k-manifold $\mathcal{M}$ we say it is an **oriented** manifold iff all coordinates on $\mathcal{M}$ are **consistently oriented**. If we make a choice and say $\phi_0 : U_0 \to V_0$ is **right-handed** then any overlapping patch $\phi_1 : U_1 \to V_1$ is said to be **right-handed** iff $det(d\theta_{01}) > 0$. Otherwise, if $det(d\theta_{01}) < 0$ then the patch $\phi_1 : U_1 \to V_1$ is said to be **left-handed**. If the manifold is **orientable** then as we continue to travel across the manifold we can choose coordinates such that on each overlap the transition functions satisfy $det(d\theta_{ij}) > 0$. In this way we find an atlas for an orientable $\mathcal{M}$ which is **right-handed**.

We can also say $\mathcal{M}$ is oriented is there exists a nonzero volume-form on $\mathcal{M}$. If $\mathcal{M}$ is k-dimensional then a volume form Vol is simply a nonzero k-form. At each point $p \in \mathcal{M}$ we can judge if a given coordinate system is left or right handed. We have to make a convention to be precise and I do so at this point. We assume Vol is positive and we say a coordinate system with chart $(x^1, x^2, \dots, x^k)$ is **positively oriented** iff $Vol(\partial_1|_p, \partial_2|_p, \dots, \partial_k|_p) > 0$. If a coordinate system is not positively oriented then it is said to be negatively oriented and we will find $Vol(\partial_1|_p, \partial_2|_p, \dots, \partial_k|_p) < 0$ in that case. It is important that we suppose $Vol_p \neq 0$ at each $p \in \mathcal{M}$ since that is what allows us to demarcate coordinate systems as positive or negatively oriented.

Naturally, you are probably wondering: is a positively oriented coordinate system is the same idea as a right-handed coordinate system as defined above? To answer that we should analyze how the Vol changes coordinates on an overlap. Suppose we are given a positive volume form Vol and a point $p \in \mathcal{M}$ where two coordinate systems x and y are both defined. There must exist some function f such that

$$Vol_x = f(x)dx^1 \wedge dx^2 \wedge \cdots \wedge dx^k$$

To change coordinates recall $dx^j = \sum_{j=1}^k \frac{\partial x^j}{\partial y^j} dy^j$ and subsitute,

$$\begin{aligned} Vol &= \sum_{j_1,\dots,j_k=1}^k (f \circ x \circ y^{-1})(y) \frac{\partial x^1}{\partial y^{j_1}} \frac{\partial x^2}{\partial y^{j_2}} \cdots \frac{\partial x^k}{\partial y^{j_k}} dy^{j_1} \wedge dy^{j_2} \wedge \cdots \wedge dy^{j_k} \\ &= (f \circ x \circ y^{-1})(y) det\left[\frac{\partial x}{\partial y}\right] dy^1 \wedge dy^2 \wedge \cdots \wedge dy^k \end{aligned} \tag{8.9}$$

If you calculate the value of Vol on $\partial_{x^I}|_p = \partial_{x_1}|_p, \partial_{x_2}|_p, \dots, \partial_{x_k}|_p$ you'll find $Vol(\partial_{x^I}|_p) = f(x(p)$. Whereas, if you evaluate Vol on $\partial_{y^I}|_p = \partial_{y_1}|_p, \partial_{y_2}|_p, \dots, \partial_{y_k}|_p$ then the value is $Vol(\partial_{y^I}|_p) = f(x(p))det\big[\frac{\partial x}{\partial y}(p)\big]$. But, we should recognize that $det\big[\frac{\partial x}{\partial y}\big] = det(d\theta_{ij})$ hence two coordinate systems which are positively oriented must also be consistently oriented. Why? Assume $Vol(\partial_{x^I}|_p) = f(x(p)) > 0$ then $Vol(\partial_{y^I}|_p) = f(x(p))det\big[\frac{\partial x}{\partial y}(p)\big] > 0$ iff $det\big[\frac{\partial x}{\partial y}(p)\big] > 0$ hence y is positively oriented if we are given that x is positively oriented **and** $det\big[\frac{\partial x}{\partial y}\big] > 0$.

Let $\mathcal{M}$ be an oriented k-manifold with orientation given by the volume form Vol and an associated atlas of positively oriented charts. Furthermore, let α be a p-form defined on $V \subseteq \mathcal{M}$. Suppose there exists a local parametrization $\phi : U \subseteq \mathbb{R}^k \to V \subseteq \mathcal{M}$ and $D \subset V$ then there is a smooth function h such that $\alpha_q = h(q)dx^1 \wedge dx^2 \wedge \cdots \wedge dx^k$ for each $q \in V$. We define the integral of α over D as follows:

$$\int_D \alpha = \int_{\phi^{-1}(D)} h(\phi(x))d^kx \qquad \leftarrow [\star_x]$$

Is this definition dependent on the coordinate system $\phi : U \subseteq \mathbb{R}^k \to V \subseteq \mathcal{M}$? If we instead used coordinate system $\psi : \bar{U} \subseteq \mathbb{R}^k \to \bar{V} \subseteq \mathcal{M}$ where coordinates $y^1, y^2, \dots, y^k$ on $\bar{V}$ then the given form α has a different coefficient of $dy^1 \wedge dy^2 \wedge \cdots \wedge dy^k$

$$\alpha = h(x)dx^1 \wedge dx^2 \wedge \cdots \wedge dx^k = (h \circ x \circ y^{-1})(y)det\left[\frac{\partial x}{\partial y}\right]dy^1 \wedge dy^2 \wedge \cdots \wedge dy^k$$

Thus, as we change over to y coordinates the function picks up a factor which is precisely the determinant of the derivative of the transition functions.

$$\begin{aligned}\int_D \alpha &= \int_D (h \circ x \circ y^{-1})(y)det\left[\frac{\partial x}{\partial y}\right]dy^1 \wedge dy^2 \wedge \cdots \wedge dy^k \\ &= \int_{\psi^{-1}(D)} (h \circ x \circ y^{-1})(y)det\left[\frac{\partial x}{\partial y}\right]d^ky \qquad \leftarrow [\star_y]\end{aligned}$$

We need $\star_x = \star_y$ in order for the integral $\int_D \alpha$ to be well-defined. Fortunately, the needed equality is *almost* provided by the change of variables theorem for multivariate integrals on $\mathbb{R}^k$. Recall,

$$\int_R f(x)d^kx = \int_{\bar{R}} \tilde{f}(y)\left|det\frac{\partial x}{\partial y}\right|d^ky$$

where $\tilde{f}$ is more pedantically written as $\tilde{f} = f \circ y^{-1}$, notation aside its just the function f written in terms of the new y-coordinates. Likewise, $\bar{R}$ limits y-coordinates so that the corresponding x-coordinates are found in R. Applying this theorem to our pull-back expression,

$$\int_{\phi^{-1}(D)} h(\phi(x))\, d^kx = \int_{\psi^{-1}(D)} (h \circ x \circ y^{-1})(y)\left|det\left[\frac{\partial x}{\partial y}\right]\right|d^ky.$$

Equality of $\star_x$ and $\star_y$ follows from the fact that $\mathcal{M}$ is oriented and has transition functions[12] θ_{ij} which satisfy $det(d\theta_{ij}) > 0$. We see that this integral to be well-defined only for oriented manifolds. To integrate over manifolds without an orientation additional ideas are needed, but it is possible.

Perhaps the most interesting case to consider is that of an embedded k-manifold in $\mathbb{R}^n$. In this context we must deal with both the coordinates of the ambient $\mathbb{R}^n$ and the local parametrizations

[12] once more recall the notation $\frac{\partial x}{\partial y}$ is just the matrix of the linear transformation $d\theta_{ij}$ and the determinant of a linear transformation is the determinant of the matrix of the transformation

of the k-manifold. In multivariate calculus we often consider vector fields which are defined on an open subset of $\mathbb{R}^3$ and then we calculate the flux over a surfaces or the work along a curve. What we have defined thus-far is in essence like definition how to integrate a vector field on a surface or a vector field along a curve, no mention of the vector field off the domain of integration was made. We supposed the forms were already defined on the oriented manifold, but, what if we are instead given a formula for a differential form on $\mathbb{R}^n$? How can we restrict that differential form to a surface or line or more generally a parametrized k-dimensional submanifold of $\mathbb{R}^n$? That is the problem we concern ourselvew with for the remainder of this section.

Let's begin with a simple object. Consider a one-form $\alpha = \sum_{i=1}^n \alpha_i dx^i$ where the function $p \to \alpha_i(p)$ is smooth on some subset of $\mathbb{R}^n$. Suppose C is a curve parametrized by $X : D \subseteq \mathbb{R} \to C \subseteq \mathbb{R}^n$ then the natural chart on C is provided by the parameter t in particular we have $T_pC = span\{\frac{\partial}{\partial t}\big|_{t_o}\}$ where $X(t_o) = p$ and $T_pC^* = span\{d_{t_o}t\}$ hence a vector field along C has the form $f(t)\frac{\partial}{\partial t}$ and a differential form has the form $g(t)dt$. How can we use the one-form α on $\mathbb{R}^n$ to naturally obtain a one-form defined along C? I propose:

$$\alpha\big|_C(t) = \sum_{i=1}^n \alpha_i(X(t))\frac{\partial X^i}{\partial t}dt$$

It can be shown that $\alpha\big|_C$ is a one-form on C. If we change coordinates on the curve by reparametrizing $t \to s$ it then the component relative to s vs. the component relative to t are related:

$$\sum_{i=1}^n \alpha_i(X(t(s)))\frac{\partial X^i}{ds} = \sum_{i=1}^n \alpha_i(X(t))\frac{dt}{ds}\frac{\partial X^i}{\partial t} = \frac{dt}{ds}\left(\sum_{i=1}^n \alpha_i(X(t))\frac{\partial X^i}{\partial t}\right)$$

This is precisely the transformation rule we want for the components of a one-form.

Example 8.8.1. *Suppose* $\alpha = dx + 3x^2dy + ydz$ *and* C *is the curve* $X : \mathbb{R} \to C \subseteq \mathbb{R}^3$ *defined by* $X(t) = (1, t, t^2)$ *we have* $x = 1$, $y = t$ *and* $z = t^2$ *hence* $dx = 0, dy = dt$ *and* $dz = 2tdt$ *on* C *hence* $\alpha\big|_C = 0 + 3dt + t(2tdt) = (3 + 2t^2)dt$.

Next, consider a two-form $\beta = \sum_{i,j=1}^n \frac{1}{2}\beta_{ij}dx^i \wedge dx^j$. Once more we consider a parametrized submanifold of $\mathbb{R}^n$. In particular use the notation $X : D \subseteq \mathbb{R}^2 \to S \subseteq \mathbb{R}^n$ where u, v serve as coordinates on the surface S. We can write an arbitrary two-form on S in the form $h(u, v)du \wedge dv$ where $h : S \to \mathbb{R}$ is a smooth function on S. How should we construct $h(u, v)$ given β? Again, I think the following formula is quite natural, honestly, what else would you do[13]?

$$\beta\big|_S(u, v) = \sum_{i,j=1}^n \beta_{ij}(X(u, v))\frac{\partial X^i}{\partial u}\frac{\partial X^j}{\partial v}du \wedge dv$$

The coefficient function of $du \wedge dv$ is smooth because we assume β_{ij} is smooth on $\mathbb{R}^n$ and the local parametrization is also assumed smooth so the functions $\frac{\partial X^i}{\partial u}$ and $\frac{\partial X^i}{\partial v}$ are smooth. Moreover, the component function has the desired coordinate change property with respect to a reparametrization of S. Suppose we reparametrize by s, t, then suppressing the point-dependence of β_{ij},

$$\beta\big|_S = \sum_{i,j=1}^n \beta_{ij}\frac{\partial Y^i}{\partial s}\frac{\partial Y^j}{\partial t}ds \wedge dt = \frac{du}{ds}\frac{dv}{dt}\sum_{i,j=1}^n \beta_{ij}\frac{\partial X^i}{\partial u}\frac{\partial X^j}{\partial v}ds \wedge dt = \sum_{i,j=1}^n \beta_{ij}\frac{\partial X^i}{\partial u}\frac{\partial X^j}{\partial v}du \wedge dv.$$

Therefore, the restriction of β to S is coordinate independent and we have thus constructed a two-form on a surface from the two-form in the ambient space.

[13] include the $\frac{1}{2}$ you say?, we'll see why not soon enough

Example 8.8.2. *Consider $\beta = y^2 dt \wedge dx + z dx \wedge dy + (x+y+z+t) dt \wedge dz$. Suppose $S \subseteq \mathbb{R}^4$ is parametrized by*

$$X(u,v) = (1,\ u^2v^2,\ 3u,\ v)$$

In other words, we are given that

$$t = 1,\ x = u^2v^2,\ y = 3u,\ z = v$$

Hence, $dt = 0$, $dx = 2uv^2du + 2u^2vdv$, $dy = 3du$ and $dz = dv$. Computing $\beta\big|_S$ is just a matter of substuting in all the formulas above, fortunately $dt = 0$ so only the $zdx \wedge dy$ term is nontrivial:

$$\beta\big|_S = v(2uv^2du + 2u^2vdv) \wedge (3du) = 6u^2v^2 dv \wedge du = -6u^2v^2 du \wedge dv.$$

It is fairly clear that we can restrict any p-form on $\mathbb{R}^n$ to a p-dimensional parametrized submanifold by the procedure we explained above for $p = 1, 2$. That is the underlying idea in the definitions which follow. Beyond that, once we have restricted the p-form β on $\mathbb{R}^n$ to $\beta|_{\mathcal{M}}$ then we pull-back the restricted form to an open subset of $\mathbb{R}^p$ and reduce the problem to an ordinary multivariate integral.

Remark 8.8.3. .

Just a warning, Einstein summation convention is used in what follows, by my count there are over a dozen places where we implicitly indicate a sum.

Our goal in the remainder of the section is to make contact with the[14] integrals we study in calculus. Note that an embedded manifold with a single patch is almost trivially oriented since there is no overlap to consider. In particular, if $\phi : U \subseteq \mathbb{R}^k \to \mathcal{M} \subseteq \mathbb{R}^n$ is a local parametrization with $\phi^{-1} = (u^1, u^2, \dots, u^k)$ then $du^1 \wedge du^2 \wedge \cdots \wedge du^k$ is a volume form for $\mathcal{M}$. This is the natural generalization of the normal-vector field construction for surfaces in $\mathbb{R}^3$.

Definition 8.8.4. integral of one-form along oriented curve:

Let $\alpha = \alpha_i dx^i$ be a one form and let C be an oriented curve with parametrization $X(t) : [a,b] \to C$ then we define the integral of the one-form α along the curve C as follows,

$$\int_C \alpha \equiv \int_a^b \alpha_i(X(t)) \frac{dX^i}{dt}(t)dt$$

where $X(t) = (X^1(t), X^2(t), \dots, X^n(t))$ so we mean X^i to be the i^{th} component of $X(t)$. Moreover, the indices are understood to range over the dimension of the ambient space, if we consider forms in $\mathbb{R}^2$ then $i = 1,2$ if in $\mathbb{R}^3$ then $i = 1,2,3$ if in Minkowski $\mathbb{R}^4$ then i should be replaced with $\mu = 0,1,2,3$ and so on.

Example 8.8.5. One form integrals vs. line integrals of vector fields: *We begin with a vector field $\vec{F}$ and construct the corresponding one-form $\omega_{\vec{F}} = F_i dx^i$. Next let C be an oriented curve with parametrization $X : [a,b] \subset \mathbb{R} \to C \subset \mathbb{R}$, observe*

$$\int_C \omega_{\vec{F}} = \int_a^b F_i(X(t)) \frac{dX^i}{dt}(t)dt = \int_C \vec{F} \cdot d\vec{l}$$

You may note that the definition of a line integral of a vector field is not special to three dimensions, we can clearly construct the line integral in n-dimensions, likewise the correspondance ω can be written between one-forms and vector fields in any dimension, provided we have a metric to lower the index of the vector field components. The same cannot be said of the flux-form correspondance, it is special to three dimensions for reasons we have explored previously.

[14] hopefully known to you already from multivariate calculus

Definition 8.8.6. integral of two-form over an oriented surface:

Let $\beta = \frac{1}{2}\beta_{ij}dx^i \wedge dx^j$ be a two-form and let S be an oriented piecewise smooth surface with parametrization $X(u,v) : D_2 \subset \mathbb{R}^2 \to S \subset \mathbb{R}^n$ then we define the integral of the two-form β over the surface S as follows,

$$\int_S \beta \equiv \int_{D_2} \beta_{ij}(X(u,v))\frac{\partial X^i}{\partial u}(u,v)\frac{\partial X^j}{\partial v}(u,v)dudv$$

where $X(u,v) = (X^1(u,v), X^2(u,v), \dots, X^n(u,v))$ so we mean X^i to be the i^{th} component of $X(u,v)$. Moreover, the indices are understood to range over the dimension of the ambient space, if we consider forms in $\mathbb{R}^2$ then $i,j = 1,2$ if in $\mathbb{R}^3$ then $i,j = 1,2,3$ if in Minkowski $\mathbb{R}^4$ then i,j should be replaced with $\mu,\nu = 0,1,2,3$ and so on.

Example 8.8.7. Two-form integrals vs. surface integrals of vector fields in $\mathbb{R}^3$: *We begin with a vector field $\vec{F}$ and construct the corresponding two-form $\Phi_{\vec{F}} = \frac{1}{2}\epsilon_{ijk}F_k dx^i \wedge dx^j$ which is to say $\Phi_{\vec{F}} = F_1 dy \wedge dz + F_2 dz \wedge dx + F_3 dx \wedge dy$. Next let S be an oriented piecewise smooth surface with parametrization $X : D \subset \mathbb{R}^2 \to S \subset \mathbb{R}^3$, then*

$$\int_S \Phi_{\vec{F}} = \int_S \vec{F} \cdot d\vec{A}$$

Proof: *Recall that the normal to the surface S has the form,*

$$N(u,v) = \frac{\partial X}{\partial u} \times \frac{\partial X}{\partial v} = \epsilon_{ijk}\frac{\partial X^i}{\partial u}\frac{\partial X^j}{\partial v}e_k$$

at the point $X(u,v)$. This gives us a vector which points along the outward normal to the surface and it is nonvanishing throughout the whole surface by our assumption that S is oriented. Moreover the vector surface integral of $\vec{F}$ over S was defined by the formula,

$$\int_S \vec{F} \cdot d\vec{A} \equiv \int\int_D \vec{F}(X(u,v)) \cdot \vec{N}(u,v)\ dudv.$$

now that the reader is reminded what's what, lets prove the proposition, dropping the (u,v) depence to reduce clutter we find,

$$\begin{aligned}
\int_S \vec{F} \cdot d\vec{A} &= \int\int_D \vec{F} \cdot \vec{N}\ dudv \\
&= \int\int_D F_k N_k\ dudv \\
&= \int\int_D F_k \epsilon_{ijk}\frac{\partial X^i}{\partial u}\frac{\partial X^j}{\partial v}\ dudv \\
&= \int\int_D (\Phi_{\vec{F}})_{ij}\frac{\partial X^i}{\partial u}\frac{\partial X^j}{\partial v}\ dudv \\
&= \int_S \Phi_{\vec{F}}
\end{aligned}$$

notice that we have again used our convention that $(\Phi_{\vec{F}})_{ij}$ refers to the tensor components of the 2-form $\Phi_{\vec{F}}$ meaning we have $\Phi_{\vec{F}} = (\Phi_{\vec{F}})_{ij}dx^i \otimes dx^j$ whereas with the wedge product $\Phi_{\vec{F}} = \frac{1}{2}(\Phi_{\vec{F}})_{ij}dx^i \wedge dx^j$, I mention this in case you are concerned there is a half in $\Phi_{\vec{F}}$ yet we never found a half in the integral. Well, we don't expect to because we defined the integral of the form with respect to the tensor components of the form, again they don't contain the half.

Example 8.8.8. *Consider the vector field $\vec{F} = (0,0,3)$ then the corresponding two-form is simply $\Phi_F = 3dx \wedge dy$. Lets calculate the surface integral and two-form integrals over the square $D = [0,1]\times[0,1]$ in the xy-plane, in this case the parameters can be taken to be x and y so $X(x,y) = (x,y)$ and,*

$$N(x,y) = \frac{\partial X}{\partial x} \times \frac{\partial X}{\partial y} = (1,0,0) \times (0,1,0) = (0,0,1)$$

which is nice. Now calculate,

$$\begin{aligned} \int_S \vec{F} \cdot d\vec{A} &= \int\int_D \vec{F} \cdot \vec{N}\ dxdy \\ &= \int\int_D (0,0,3) \cdot (0,0,1) dxdy \\ &= \int_0^1\int_0^1 3dxdy \\ &= 3. \end{aligned}$$

Consider that $\Phi_F = 3dx \wedge dy = 3dx \otimes dy - 3dy \otimes dx$ therefore we may read directly that $(\Phi_F)_{12} = -(\Phi_F)_{21} = 3$ and all other components are zero,

$$\begin{aligned} \int_S \Phi_F &= \int\int_D (\Phi_F)_{ij} \frac{\partial X^i}{\partial x}\frac{\partial X^j}{\partial y}\ dxdy \\ &= \int\int_D \left(3\frac{\partial X^1}{\partial x}\frac{\partial X^2}{\partial y} - 3\frac{\partial X^2}{\partial x}\frac{\partial X^1}{\partial y} \right) dxdy \\ &= \int_0^1\int_0^1 \left(3\frac{\partial x}{\partial x}\frac{\partial y}{\partial y} - 3\frac{\partial y}{\partial x}\frac{\partial x}{\partial y} \right) dxdy \\ &= 3. \end{aligned}$$

Definition 8.8.9. integral of a three-form over an oriented volume:

Let $\gamma = \frac{1}{6}\beta_{ijk} dx^i \wedge dx^j \wedge dx^k$ be a three-form and let V be an oriented piecewise smooth volume with parametrization $X(u,v,w) : D_3 \subset \mathbb{R}^3 \to V \subset \mathbb{R}^n$ then we define the integral of the three-form γ in the volume V as follows,

$$\int_V \gamma \equiv \int_{D_3} \gamma_{ijk}(X(u,v,w)) \frac{\partial X^i}{\partial u}\frac{\partial X^j}{\partial v}\frac{\partial X^k}{\partial w}\ dudvdw$$

where $X(u,v,w) = (X^1(u,v,w), X^2(u,v,w), \dots, X^n(u,v,w))$ so we mean X^i to be the i^{th} component of $X(u,v,w)$. Moreover, the indices are understood to range over the dimension of the ambient space, if we consider forms in $\mathbb{R}^3$ then $i,j,k = 1,2,3$ if in Minkowski $\mathbb{R}^4$ then i,j,k should be replaced with $\mu,\nu,\sigma = 0,1,2,3$ and so on.

Finally we define the integral of a p-form over an p-dimensional subspace of $\mathbb{R}$, we assume that $p \leq n$ so that it is possible to embed such a subspace in $\mathbb{R}$,

Definition 8.8.10. integral of a p-form over an oriented volume:

Let $\gamma = \frac{1}{p!}\beta_{i_1 \dots i_p} dx^{i_1} \wedge \cdots dx^{i_p}$ be a p-form and let S be an oriented piecewise smooth subspace with parametrization $X(u_1, \dots, u_p) : D_p \subset \mathbb{R}^p \to S \subset \mathbb{R}^n$ (for $n \geq p$) then we define the integral of the p-form γ in the subspace S as follows,

$$\int_S \gamma \equiv \int_{D_p} \beta_{i_1 \dots i_p}(X(u_1, \dots, u_p)) \frac{\partial X^{i_1}}{\partial u_1} \cdots \frac{\partial X^{i_p}}{\partial u_p}\, du_1 \cdots du_p$$

where $X(u_1, \dots, u_p) = (X^1(u_1, \dots, u_p), X^2(u_1, \dots, u_p), \dots, X^n(u_1, \dots, u_p))$ so we mean X^i to be the i^{th} component of $X(u_1, \dots, u_p)$. Moreover, the indices are understood to range over the dimension of the ambient space.

Integrals of forms play an important role in modern physics. I hope you can begin to appreciate that forms recover all the formulas we learned in multivariate calculus and give us a way to extend calculation into higher dimensions with ease. I include a toy example at the conclusion of this chapter just to show you how electromagnetism is easily translated into higher dimensions.

8.9 Generalized Stokes Theorem

The generalized Stokes theorem contains within it most of the main theorems of integral calculus, namely the fundamental theorem of calculus, the fundamental theorem of line integrals (a.k.a the FTC in three dimensions), Greene's Theorem in the plane, Gauss' Theorem and also Stokes Theorem, not to mention a myriad of higher dimensional not so commonly named theorems. The breadth of its application is hard to overstate, yet the statement of the theorem is simple,

Theorem 8.9.1. *Generalized Stokes Theorem:*

Let S be an oriented, piecewise smooth (p+1)-dimensional subspace of $\mathbb{R}^n$ where $n \geq p+1$ and let ∂S be it boundary which is consistently oriented then for a p-form α which behaves reasonably on S we have that

$$\boxed{\int_S d\alpha = \int_{\partial S} \alpha}$$

The proof of this theorem (and a more careful statement of it) can be found in a number of places, Susan Colley's *Vector Calculus* or Steven H. Weintraub's *Differential Forms: A Complement to Vector Calculus* or Spivak's *Calculus on Manifolds* just to name a few. I believe the argument in Edward's text is quite complete. In any event, you should already be familar with the idea from the usual Stokes Theorem where we must insist the boundary curve to the surface is related to the surface's normal field according to the right-hand-rule. Explaining how to orient the boundary $\partial\mathcal{M}$ given an oriented $\mathcal{M}$ is the problem of generalizing the right-hand-rule to many dimensions. I leave it to your homework for the time being.

Lets work out how this theorem reproduces the main integral theorems of calculus.

Example 8.9.2. Fundamental Theorem of Calculus in $\mathbb{R}$: *Let $f : \mathbb{R} \to \mathbb{R}$ be a zero-form then consider the interval $[a, b]$ in $\mathbb{R}$. If we let $S = [a, b]$ then $\partial S = \{a, b\}$. Further observe that $df = f'(x)dx$. Notice by the definition of one-form integration*

$$\int_S df = \int_a^b f'(x)dx$$

However on the other hand we find (the integral over a zero-form is taken to be the evaluation map, perhaps we should have defined this earlier, oops., but its only going to come up here so I'm leaving it.)

$$\int_{\partial S} f = f(b) - f(a)$$

Hence in view of the definition above we find that

$$\int_a^b f'(x)dx = f(b) - f(a) \quad \Longleftrightarrow \quad \int_S df = \int_{\partial S} f$$

Example 8.9.3. Fundamental Theorem of Calculus in $\mathbb{R}^3$: *Let $f : \mathbb{R}^3 \to \mathbb{R}$ be a zero-form then consider a curve C from $p \in \mathbb{R}^3$ to $q \in \mathbb{R}^3$ parametrized by $\phi : [a,b] \to \mathbb{R}^3$. Note that $\partial C = \{\phi(a) = p, \phi(b) = q\}$. Next note that*

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

Then consider that the exterior derivative of a function corresponds to the gradient of the function thus we are not to surprised to find that

$$\int_C df = \int_a^b \frac{\partial f}{\partial x^i}\frac{dx^i}{dt} dt = \int_C (\nabla f) \cdot d\vec{l}$$

On the other hand, we use the definition of the integral over a a two point set again to find

$$\int_{\partial C} f = f(q) - f(p)$$

Hence if the Generalized Stokes Theorem is true then so is the FTC in three dimensions,

$$\int_C (\nabla f) \cdot d\vec{l} = f(q) - f(p) \quad \Longleftrightarrow \quad \int_C df = \int_{\partial C} f$$

another popular title for this theorem is the "fundamental theorem for line integrals". As a final thought here we notice that this calculation easily generalizes to 2,4,5,6,... dimensions.

Example 8.9.4. Greene's Theorem: *Let us recall the statement of Greene's Theorem as I have not replicated it yet in the notes, let D be a region in the xy-plane and let ∂D be its consistently oriented boundary then if $\vec{F} = (M(x,y), N(x,y), 0)$ is well behaved on D*

$$\int_{\partial D} Mdx + Ndy = \int\int_D \left(\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y}\right) dxdy$$

We begin by finding the one-form corresponding to $\vec{F}$ namely $\omega_F = Mdx + Ndy$ consider then that

$$d\omega_F = d(Mdx + Ndy) = dM \wedge dx + dN \wedge dy = \frac{\partial M}{\partial y} dy \wedge dx + \frac{\partial N}{\partial x} dx \wedge dy$$

which simplifies to,

$$d\omega_F = \left(\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y}\right) dx \wedge dy = \Phi_{(\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y})\hat{k}}$$

Thus, using our discussion in the last section we recall

$$\int_{\partial D} \omega_F = \int_{\partial D} \vec{F} \cdot d\vec{l} = \int_{\partial D} Mdx + Ndy$$

where we have reminded the reader that the notation in the rightmost expression is just another way of denoting the line integral in question. Next observe,

$$\int_D d\omega_F = \int_D (\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y})\hat{k} \cdot d\vec{A}$$

And clearly, since $d\vec{A} = \hat{k}dxdy$ we have

$$\int_D (\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y})\hat{k} \cdot d\vec{A} = \int_D (\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y})dxdy$$

Therefore,

$$\int_{\partial D} Mdx + Ndy = \int\int_D \left(\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y}\right) dxdy \quad \Longleftrightarrow \quad \int_D d\omega_F = \int_{\partial D} \omega_F$$

Example 8.9.5. Gauss Theorem: *Let us recall Gauss Theorem to begin, for suitably defined $\vec{F}$ and V,*

$$\int_{\partial V} \vec{F} \cdot d\vec{A} = \int_V \nabla \cdot \vec{F} \; d\tau$$

First we recall our earlier result that

$$d(\Phi_F) = (\nabla \cdot \vec{F})dx \wedge dy \wedge dz$$

Now note that we may integrate the three form over a volume,

$$\int_V d(\Phi_F) = \int_V (\nabla \cdot \vec{F})dxdydz$$

whereas,

$$\int_{\partial V} \Phi_F = \int_{\partial V} \vec{F} \cdot d\vec{A}$$

so there it is,

$$\int_V (\nabla \cdot \vec{F})d\tau = \int_{\partial V} \vec{F} \cdot d\vec{A} \quad \Longleftrightarrow \quad \int_V d(\Phi_F) = \int_{\partial V} \Phi_F$$

I have left a little detail out here, I may assign it for homework.

Example 8.9.6. Stokes Theorem: *Let us recall Stokes Theorem to begin, for suitably defined $\vec{F}$ and S,*

$$\int_S (\nabla \times \vec{F}) \cdot d\vec{A} = \int_{\partial S} \vec{F} \cdot d\vec{l}$$

Next recall we have shown in the last chapter that,

$$d(\omega_F) = \Phi_{\nabla \times \vec{F}}$$

Hence,

$$\int_S d(\omega_F) = \int_S (\nabla \times \vec{F}) \cdot d\vec{A}$$

whereas,

$$\int_{\partial S} \omega_F = \int_{\partial S} \vec{F} \cdot d\vec{l}$$

which tells us that,

$$\int_S (\nabla \times \vec{F}) \cdot d\vec{A} = \int_{\partial S} \vec{F} \cdot d\vec{l} \quad \Longleftrightarrow \quad \int_S d(\omega_F) = \int_{\partial S} \omega_F$$

The Generalized Stokes Theorem is perhaps the most persausive argument for mathematicians to be aware of differential forms, it is clear they allow for more deep and sweeping statements of the calculus. The generality of differential forms is what drives modern physicists to work with them, string theorists for example examine higher dimensional theories so they are forced to use a language more general than that of the conventional vector calculus.

8.10 Poincare lemma

This section is in large part inspired by M. Gockeler and T. Schucker's *Differential geometry, gauge theories, and gravity* page 20-22. The converse calculation is modelled on the argument found in H. Flanders *Differential Forms with Applications to the Physical Sciences*. The original work was done around the dawn of the twentieth century and can be found in many texts besides the two I mentioned here.

8.10.1 exact forms are closed

Proposition 8.10.1.

The exterior derivative of the exterior derivative is zero. $d^2 = 0$

Proof: Let α be an arbitrary p-form then

$$d\alpha = \frac{1}{p!}(\partial_m \alpha_{i_1 i_2 \dots i_p}) dx^m \wedge dx^{i_1} \wedge dx^{i_2} \wedge \cdots \wedge dx^{i_p} \tag{8.10}$$

then differentiate again,

$$\begin{aligned} d(d\alpha) &= d\left[\tfrac{1}{p!}(\partial_m \alpha_{i_1 i_2 \dots i_p}) dx^m \wedge dx^{i_1} \wedge dx^{i_2} \wedge \cdots \wedge dx^{i_p}\right] \\ &= \tfrac{1}{p!}(\partial_k \partial_m \alpha_{i_1 i_2 \dots i_p}) dx^k \wedge dx^m \wedge dx^{i_1} \wedge dx^{i_2} \wedge \cdots \wedge dx^{i_p} \\ &= 0 \end{aligned} \tag{8.11}$$

since the partial derivatives commute whereas the wedge product anticommutes so we note that the pair of indices (k,m) is symmetric for the derivatives but antisymmetric for the wedge, as we know the sum of symmetric against antisymmetric vanishes (see equation **??** part *iv* if you forgot.)

Definition 8.10.2.

A differential form α is **closed** iff $d\alpha = 0$. A differential form β is **exact** iff there exists γ such that $\beta = d\gamma$.

Proposition 8.10.3.

All exact forms are closed. However, there exist closed forms which are not exact.

Proof: Exact implies closed is easy, let β be exact such that$\beta = d\gamma$ then

$$d\beta = d(d\gamma) = 0$$

using the theorem $d^2 = 0$. To prove that there exists a closed form which is not exact it suffices to give an example. A popular example (due to its physical significance to magnetic monopoles, Dirac Strings and the like..) is the following differential form in $\mathbb{R}^2$

$$\phi = \frac{1}{x^2 + y^2}(xdy - ydx) \tag{8.12}$$

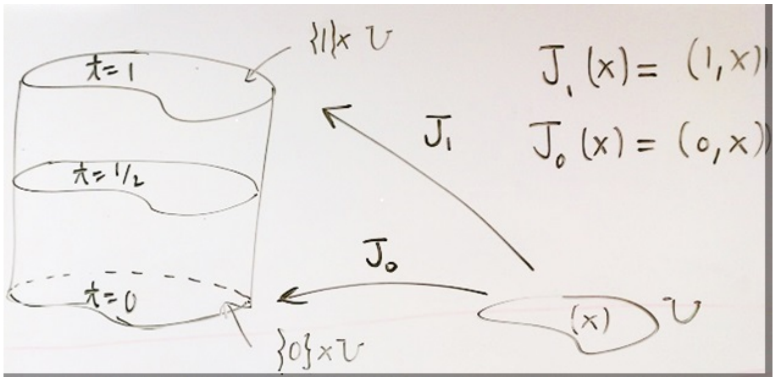
You may verify that $d\phi = 0$ in homework. Observe that if ϕ were exact then there would exist f such that $\phi = df$ meaning that

$$\frac{\partial f}{\partial x} = -\frac{y}{x^2+y^2}, \qquad \frac{\partial f}{\partial y} = \frac{x}{x^2+y^2}$$

which are solved by $f = arctan(y/x) + c$ where c is arbitrary. Observe that f is ill-defined along the y-axis $x = 0$ (this is the Dirac String if we put things in context), however the natural domain of ϕ is $\mathbb{R}^{\,n\times n} - \{(0,0)\}$. $\square$

8.10.2 potentials for closed forms

Poincare' suggested the following partial converse, he said closed implies exact provided we place a topological restriction on the domain of the form. In particular, if the domain of a closed form is smoothly deformable to a point then each closed form is exact. We'll work out a proof of that result for a subset of $\mathbb{R}^n$. Be patient, we have to build some toys before we play.



Suppose $U \subseteq \mathbb{R}^n$ and $I = [0,1]$ we denote a typical point in $I \times U$ as (t,x) where $t \in I$ and $x \in \mathbb{R}^n$. Define maps,

$$J_1 : U \to I \times U, \qquad J_0 : U \to I \times U$$

by $J_1(x) = (1,x)$ and $J_0(x) = (0,x)$ for each $x \in U$. Flanders encourages us to view $I \times U$ as a cylinder and where the map J_1 maps U to the top and J_0 maps U to the base. We can pull-back forms on the cylinder to the U on the top ($t = 1$) or to the base ($t = 0$). For instance, if we consider $\omega = (x+t)dx + dt$ for the case $n = 1$ then

$$J_0^*\omega = xdx \qquad J_1^*\omega = (x+1)dx.$$

Define a smooth mapping K of $(p+1)$ forms on $I \times U$ to p-forms on U as follows:

$$(1.)\ K(a(t,x)dx^I) = 0, \qquad (2.)\ K(a(t,x)dt \wedge dx^J) = \left(\int_0^1 a(t,x)dt\right)dx^J$$

for multi-indices I of length $(p+1)$ and J of length p. The cases (1.) and (2.) simply divide the possible monomial[15] inputs from $\Lambda^{p+1}(I \times U)$ into forms which have dt and those which don't. Then K is defined for a general $(p+1)$-form on $I \times U$ by linearly extending the formulas above to multinomials of the basic monomials.

It turns out that the following identity holds for K:

[15] $dx \wedge dy$ is a monomial whereas $dx + dy$ is a binomial in this context

Lemma 8.10.4. *the K-lemma.*

If ω is a differential form on $I \times U$ then

$$K(d\omega) + d(K(\omega)) = J_1^*\omega - J_0^*\omega.$$

Proof: Since the equation is given for linear operations it suffices to check the formula for monomials since we can extend the result linearly once those are affirmed. As in the definition of K there are two basic categories of forms on $I \times U$:

Case 1: If $\omega = a(t,x)dx^I$ then clearly $K(\omega) = 0$ hence $d(K(\omega)) = 0$. Observe,

$$d\omega = da \wedge dx^I = \sum_j \frac{\partial a}{\partial x^j} dx^j \wedge dx^I + \frac{\partial a}{\partial t} dt \wedge dx^I$$

Hence $K(d\omega)$ is calculated as follows:

$$\begin{aligned} K(d\omega) &= K\left(\sum_j \frac{\partial a}{\partial x^j} dx^j \wedge dx^I\right) + K\left(\frac{\partial a}{\partial t} dt \wedge dx^I\right) \\ &= \left(\int_0^1 \frac{\partial a}{\partial t} dt\right) dx^I \\ &= \big[a(x,1) - a(x,0)\big] dx^I \\ &= J_1^*\omega - J_0^*\omega \end{aligned} \tag{8.13}$$

where we used the FTC in the next to last step. The pull-backs in this case just amount to evaluation at $t = 0$ or $t = 1$ as there is no dt-type term to squash in ω. The identity follows.

Case 2: Suppose $\omega = a(t,x)dt \wedge dx^J$. Calculate,

$$d\omega = \sum_j \frac{\partial a}{\partial x^k} dx^k \wedge dt \wedge dx^J + \frac{\partial a}{\partial t} \underbrace{dt \wedge dt}_{zero\ !} \wedge dx^J$$

Thus, using $dx^k \wedge dt = -dt \wedge dx^k$, we calculate:

$$\begin{aligned} K(d\omega) &= K\left(-\sum_k \frac{\partial a}{\partial x^k} dt \wedge dx^k \wedge dx^I\right) \\ &= -\sum_k \left(\int_0^1 \frac{\partial a}{\partial x^k} dt\right) dx^k \wedge dx^I \end{aligned}$$

at which point we cannot procede further since a is an arbitrary function which can include a nontrivial time-dependence. We turn to the calculation of $d(K(\omega))$. Recall we defined

$$K(\omega) = \left(\int_0^1 a(t,x)dt\right) dx^J.$$

We calculate the exterior derivative of $K(\omega)$:

$$\begin{aligned} d(K(\omega)) &= d\Big(\int_0^1 a(t,x)dt\Big) \wedge dx^J \\ &= \Big(\frac{\partial}{\partial t}\underbrace{\Big[\int_0^1 a(\tau,x)d\tau\Big]}_{constant\ in\ t} dt + \sum_k \frac{\partial}{\partial x^k}\Big[\int_0^1 a(t,x)dt\Big]dx^k\Big) \wedge dx^J \\ &= \sum_k \Big(\int_0^1 \frac{\partial a}{\partial x^k}dt\Big)dx^k \wedge dx^J. \qquad (8.14) \end{aligned}$$

Therefore, $K(d\omega)+d(K(\omega)) = 0$ and clearly $J_0^*\omega = J_1^*\omega = 0$ in this case since the pull-backs squash the dt to zero. The lemma follows. $\square$.

Definition 8.10.5.

A subset $U \subseteq \mathbb{R}^n$ is deformable to a point P if there exists a smooth mapping $G : I \times U \to U$ such that $G(1,x) = x$ and $G(0,x) = P$ for all $x \in U$.

The map G deforms U smoothly into the point P. Recall that $J_1(x) = (1,x)$ and $J_0(x) = (0,x)$ hence the conditions on the deformation can be expressed as:

$$G(J_1(x)) = x \qquad G(J_0(x)) = P$$

Denoting Id for the identity on U and P as the constant mapping with value P on U we have

$$G \circ J_1 = Id \qquad G \circ J_0 = P$$

Therefore, if γ is a $(p+1)$-form on U we calculate,

$$(G \circ J_1)^*\gamma = Id^*\gamma \quad \Rightarrow \quad J_1^*[G^*\gamma] = \gamma$$

whereas,

$$(G \circ J_0)^*\gamma = P^*\gamma = 0 \quad \Rightarrow \quad J_0^*[G^*\gamma] = 0$$

Apply the K-lemma to the form $\omega = G^*\gamma$ on $I \times U$ and we find:

$$K(d(G^*\gamma)) + d(K(G^*\gamma)) = \gamma.$$

However, recall that we proved that pull-backs and exterior derivatives commute thus

$$d(G^*\gamma) = G^*(d\gamma)$$

and we find an extremely interesting identity,

$$\boxed{K(G^*(d\gamma)) + d(K(G^*\gamma)) = \gamma.}$$

Proposition 8.10.6.

If $U \subseteq \mathbb{R}$ is deformable to a point P then a p-form γ on U is closed iff ϕ is exact.

Proof: Suppose γ is exact then $\gamma = d\beta$ for some $(p-1)$-form β on U hence $d\gamma = d(d\beta) = 0$ by Proposition 8.10.1 hence γ is closed. Conversely, suppose γ is closed. Apply the boxed consequence of the K-lemma, note that $K(G^*(0)) = 0$ since we assume $d\gamma = 0$. We find,

$$d(K(G^*\gamma)) = \gamma$$

identify that $G^*\gamma$ is a p-form on $I \times U$ whereas $K(G^*\gamma)$ is a $(p-1)$-form on U by the very construction of K. It follows that γ is exact since we have shown how it is obtained as the exterior derivative of another differential form of one degree less. □

Where was deformability to a point P used in the proof above? The key is the existence of the mapping G. In other words, if you have a space which is not deformable to a point then no deformation map G is available and the construction via K breaks down. Basically, if the space has a hole which you get stuck on as you deform loops to a point then it is not deformable to a point. Often we call such spaces *simply connected.* Careful definition of these terms is too difficult for calculus, deformation of loops and higher dimensional objects is properly covered in algebraic topology. In any event, the connection of the deformation and exactness of closed forms allows topologists to use differential forms detect holes in spaces. In particular:

Definition 8.10.7. de Rham cohomology:

We define several real vector spaces of differential forms over some subset U of $\mathbb{R}$,

$$Z^p(U) \equiv \{\phi \in \Lambda^p U \mid \phi \text{ closed}\}$$

the space of closed p-forms on U. Then,

$$B^p(U) \equiv \{\phi \in \Lambda^p U \mid \phi \text{ exact}\}$$

the space of exact p-forms where by convention $B^0(U) = \{0\}$ The de Rham cohomology groups are defined by the quotient of closed/exact,

$$H^p(U) \equiv Z^p(U)/B^p(U).$$

the $dim(H^pU) = p^{th}$ Betti number of U.

We observe that simply connected regions have all the Betti numbers zero since $Z^p(U) = B^p(U)$ implies that $H^p(U) = \{0\}$. Of course there is much more to say about de Rahm Cohomology, I just wanted to give you a taste and alert you to the fact that differential forms can be used to reveal aspects of topology. Not all algebraic topology uses differential forms though, there are several other calculational schemes based on triangulation of the space, or studying singular simplexes. One important event in 20-th century mathematics was the discovery that all these schemes described the same homology groups. The Steenrod reduced the problem to a few central axioms and it was shown that all the calculational schemes adhere to that same set of axioms.

One interesting aspect of the proof we (copied from Flanders [16]) is that it is not a mere existence proof. It actually lays out how to calculate the form which provides exactness. Let's call β the potential form of γ if $\gamma = d\beta$. Notice this is totally reasonable langauge since in the case of classical mechanics we consider conservative forces $\vec{F}$ which as derivable from a scalar potential

[16] I don't know the complete history of this calculation at the present. It would be nice to find it since I doubt Flanders is the originator.

V by $\vec{F} = -\nabla V$. Translated into differential forms we have $\omega_{\vec{F}} = -dV$. Let's explore how the K-mapping and proof indicate the potential of a vector field ought be calculated.

Suppose U is deformable to a point and F is a smooth conservative vector field on U. Perhaps you recall that for conservative F are irrotational hence $\nabla \times F = 0$. Recall that $d\omega_F = \Phi_{\nabla \times F} = \Phi_0 = 0$ thus the one-form corresponding to a conservative vector field is a closed form. Apply the identity: let $G : I \times U \to U \subseteq \mathbb{R}^3$ be the deformation of U to a point P,

$$d(K(G^*\omega_F)) = \omega_F$$

Hence, including the minus to make energy conservation natural,

$$V = -K(G^*\omega_F)$$

For convenience, lets suppose the space considered is the unit-ball B and lets use a deformation to the origin. Explicitly, $G(t, r) = tr$ for all $r \in \mathbb{R}^3$ such that $||r|| \leq 1$. Note that clearly $G(0, r) = 0$ whereas $G(1, r) = r$ and G has a nice formula so it's smooth[17]. We wish to calculate the pull-back of $\omega_F = Pdx + Qdy + Rdz$ under G, from the definition of pull-back we have

$$(G^*\omega_F)(X) = \omega_F(dG(X))$$

for each smooth vector field X on $I \times B$. Differential forms on $I \times B$ are written as linear combinations of dt, dx, dy, dz with smooth functions as coefficients. We can calculate the coefficents by evalutaion on the corresponding vector fields $\partial_t, \partial_x, \partial_y, \partial_z$. Observe, since $G(t, x, y, z) = (tx, ty, tz)$ we have

$$dG(\partial_t) = \frac{\partial G^1}{\partial t}\frac{\partial}{\partial x} + \frac{\partial G^2}{\partial t}\frac{\partial}{\partial y} + \frac{\partial G^3}{\partial t}\frac{\partial}{\partial z} = x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y} + z\frac{\partial}{\partial z}$$

wheras,

$$dG(\partial_x) = \frac{\partial G^1}{\partial x}\frac{\partial}{\partial x} + \frac{\partial G^2}{\partial x}\frac{\partial}{\partial y} + \frac{\partial G^3}{\partial x}\frac{\partial}{\partial z} = t\frac{\partial}{\partial x}$$

and similarly,

$$dG(\partial_y) = t\frac{\partial}{\partial y} \qquad dG(\partial_x) = t\frac{\partial}{\partial z}$$

Furthermore,

$$\omega_F(dG(\partial_t)) = \omega_F(x\partial_x + y\partial_y + z\partial_z) = xP + yQ + zR$$

$$\omega_F(dG(\partial_x)) = \omega_F(t\partial_x) = tP, \qquad \omega_F(dG(\partial_y)) = \omega_F(t\partial_y) = tQ, \qquad \omega_F(dG(\partial_z)) = \omega_F(t\partial_z) = tR$$

Therefore,

$$G^*\omega_F = (xP + yQ + zR)dt + tPdx + tQdy + tRdz = (xP + yQ + zR)dt + t\omega_F$$

Now we can calculate $K(G^*\omega_F)$ and hence V [18]

$$K(G^*\omega_F)(t, x, y, z) = K\bigg(\big(xP(tx, ty, tz) + yQ(tx, ty, tz) + zR(tx, ty, tz)\big)dt\bigg)$$

[17] there is of course a deeper meaning to the word, but, for brevity I gloss over this.

[18] Note that only the coefficient of dt gives a nontrivial contribution so in retrospect we did a bit more calculation than necessary. That said, I'll just keep it as a celebration of extreme youth for calculation. Also, I've been a bit careless in omiting the point up to this point, let's include the point dependence since it will be critical to properly understand the formula.

Therefore,

$$V(x,y,z) = -K(G^*\omega_F) = -\int_0^1 \big(xP(tx,ty,tz) + yQ(tx,ty,tz) + zR(tx,ty,tz)\big)dt$$

Notice this is precisely the line-integral of $F =< P,Q,R >$ along the line C with direction $< x,y,z >$ from the origin to (x,y,z). In particular, if $\vec{r}(t) =< tx,ty,tz >$ then $\frac{d\vec{r}}{dt} =< x,y,z >$ hence we identify

$$V(x,y,z) = -\int_0^1 \vec{F}\big(\vec{r}(t)\big)\cdot\frac{d\vec{r}}{dt}\,dt = -\int_C \vec{F}\cdot d\vec{r}$$

Perhaps you recall this is precisely how we calculate the potential function for a conservative vector field provided we take the origin as the zero for the potential.

Actually, this calculation is quite interesting. Suppose we used a different deformation $\tilde{G} : I \times U \to U$. For fixed point Q we travel to from the origin to the point by the path $t \mapsto \tilde{G}(t,Q)$. Of course this path need not be a line. The space considered might look like a snake where a line cannot reach from the base point P to the point Q. But, the same potential is derived. Why? Path independence of the vector field is one answer. The criteria $\nabla \times F = 0$ suffices for a simply connected region. However, we see something deeper. The criteria of a closed form paired with a simply connected (deformable) domain suffices to construct a potential for the given form. This result reproduces the familar case of conservative vector fields derived from scalar potentials and *much more*. In Flanders he calculates the potential for a closed two-form. This ought to be the mathematics underlying the construction of the so-called **vector potential** of magnetism. In junior-level electromagnetism[19] the magnetic field B satisfies $\nabla \cdot B = 0$ and thus the two-form Φ_B has exterior derivative $d\Phi_B = \nabla \cdot B dx \wedge dy \wedge dz = 0$. The magnetic field corresponds to a closed form. Poincare's lemma shows that there exists a one-form ω_A such that $d\omega_A = \Phi_B$. But this means $\Phi_{\nabla\times A} = \Phi_B$ hence in the langauge of vector fields we expect the vector potential A generated the magnetic field B throught the curl $B = \nabla \times A$. Indeed, this is precisely what a typical junior level physics student learns in the magnetostatic case. Appreciate that it goes deeper still, the Poincare lemma holds for p-forms which correspond to objects which don't match up with the quaint vector fields of 19-th century physics. We can be confident to find potential for 3-form fluxes in a 10-dimensional space, or wherever our imagination takes us. I explain at the end of this chapter how to translate electromagnetics into the langauge of differential forms, it may well be that in the future we think about forms the way we currently think about vectors. This is one of the reasons I like Flanders text, he really sticks with the langauge of differential forms throughout. In contrast to these notes, he just does what is most interesting. I think undergraduates need to see more detail and not just the most clever calculations, but, I can hardly blame Flanders! He makes no claim to be an undergraduate work.

Finally, I should at least mention that though we can derive a potential β for a given closed form α on a simply connected domain it need not be unique. In fact, it will not be unique unless we add further criteria for the potential. This ambuity is called **gauge freedom** in physics. Mathematically it's really simple give form language. If $\alpha = d\beta$ where β is a $(p-1)$-form then we can take any smooth $(p-2)$ form and calculate that

$$d(\alpha + d\lambda) = d\beta + d^2\lambda = d\beta = \alpha$$

Therefore, if β is a potential-form for α then $\beta + d\lambda$ is also a potential-form for α.

[19] just discussing magnetostatic case here to keep it simple

8.11 introduction to geometric differential equations

Differential forms, pull-backs and submanifolds provide a language in which the general theory of partial differential equations is naturally expressed. That said, let us begin with an application to the usual differential equations course.

8.11.1 exact differential equations

Throughout what follows assume that M, N, I, F are continuously differentiable functions of x, y, perhaps just defined for some open subset of $\mathbb{R}^2$. Recall (or learn) that a differential equation $Mdx + Ndy = 0$ is said to be **exact** iff there is a function F such that $dF = Mdx + Ndy$. This is a very nice kind of differential equation since the solution is simply $F(x, y) = c$. A convenient test for exactness was provided from the fact that partial derivatives commute. This exchange of partial derivatives implies that $\partial_y M = \partial_x N$. Observe, we can recover this condition via exterior differentiation:

Proposition 8.11.1.

If a differential equation $Mdx + Ndy = 0$ is exact then $d(Mdx + Ndy) = 0$ (this zero is derived from M, N alone, independent of the given differential equation).

Proof: if $Mdx + Ndy = 0$ is exact then by definition $dF = Mdx + Ndy$ hence $d(Mdx + Ndy) = d(dF) = 0$. □

Pfaff did pioneering work in the theory of differential forms. One Theorem due to Pfaff states that any first order differential equation can be made exact by the method of integrating factors. In particular, if $Mdx + Ndy = 0$ is not exact then there exists I such that $IMdx + INdy = 0$ is an exact differential equation. The catch, it is as hard or harder to find I as it is to solve the given differential equation. That said, the integrating factor method is an example of this method. Although, we don't usually think of linear ordinary differential equations as an exact equation, it can be viewed as such[20].

A differential equation in $x_1, x_2, \dots, x_n$ of the form:

$$M_1dx_1 + M_2dx_2 + \cdots + M_ndx_n = 0$$

can be written as $dF = 0$ locally iff

$$d(M_1dx_1 + M_2dx_2 + \cdots + M_ndx_n) = 0.$$

The fact that exact implies closed is just $d^2 = 0$. The converse direction, assuming closed near a point, only gives the existence of a potential form F close to the point. Globally, there could be a topological obstruction as we saw in the Poincare Lemma section.

Example 8.11.2. Problem: *Suppose $xdy + ydx - xdz = 0$. Is there a point(s) in $\mathbb{R}^3$ near which there exists F such that $dF = xdy + ydx - xdz$?*

Solution: *If there was the the differential form of the DEqn would vanish identically. However:*

$$d(xdy + ydx - xdz) = dx \wedge dy + dy \wedge dx - dx \wedge dz = dz \wedge dx. \qquad \textit{Therefore, no.}$$

[20] see my differential equations notes, it's in there

We can try to find an integrating factor. Let's give it a shot, this problem is simple enough it may be possible to work it out. We want I such that $I(xdy + ydx - xdz)$ is a closed one-form. Use the Leibniz rule and our previous calculation:

$$\begin{aligned} d[I(xdy + ydx - xdz)] &= dI \wedge (xdy + ydx - xdz) + Idz \wedge dx \\ &= (I_x dx + I_y dy + I_z dz) \wedge (xdy + ydx - xdz) + Idz \wedge dx \\ &= (xI_x - yI_y)dx \wedge dy + (xI_x + yI_z + I)dz \wedge dx + (-xI_y - xI_z)dy \wedge dz \end{aligned}$$

Therefore, our integrating factor must satisfy the following partial differential equations,

$$(I.)\ xI_x - yI_y = 0, \qquad (II.)\ xI_x + yI_z + I = 0, \qquad (III.)\ -xI_y - xI_z = 0.$$

I'll leave this to the reader. I'm not sure if it has a solution. It seems possible that the differential consquences of this system are nonsensical. I just wanted to show how differential forms allow us to extend to higher dimensional problems. Notice, we could just as well have not solved a problem with 4 or 5 variables.

8.11.2 differential equations via forms

I follow Example 1.2.3. of *Cartan for Beginners: Differential Geometry via Moving Frames and Exterior Differential Systems* by Thomas A. Ivey and J.M Landsberg. I suspect understanding this text makes you quite a bit more than a *beginner*. Consider,

$$u_x = A(x, y, u) \tag{8.15}$$

$$u_y = B(x, y, u) \tag{8.16}$$

This section is not finished.

Chapter 9

Electromagnetism in differential form

Warning: I will use Einstein's implicit summation convention throughout this section. I have made a point of abstaining from Einstein's convention in these notes up to this point. However, I just can't bear the summations in this section. They're just too ugly.

9.1 differential forms in Minkowski space

The logic here follows fairly close to the last section, however the wrinkle is that the metric here demands more attention. We must take care to raise the indices on the forms when we Hodge dual them. First we list the basis differential forms, we have to add time to the mix (again $c = 1$ so $x^0 = ct = t$ if you worried about it)

Greek indices are defined to range over $0, 1, 2, 3$. Here the top form is degree four since in four dimensions we can have four differentials without a repeat. Wedge products work the same as they have before, just now we have dt to play with. Hodge duality may offer some surprises though.

Definition 9.1.1. *The antisymmetric symbol in* **flat** $\mathbb{R}^4$ *is denoted* $\epsilon_{\mu\nu\alpha\beta}$ *and it is defined by the value*

$$\epsilon_{0123} = 1$$

plus the demand that it be completely antisymmetric.

We must not assume that this symbol is invariant under a cyclic exhange of indices. Consider,

$$\begin{aligned} \epsilon_{0123} &= -\epsilon_{1023} && \text{flipped (01)} \\ &= +\epsilon_{1203} && \text{flipped (02)} \\ &= -\epsilon_{1230} && \text{flipped (03).} \end{aligned} \tag{9.1}$$

Example 9.1.2. *We now compute the Hodge dual of $\gamma = dx$ with respect to the Minkowski metric $\eta_{\mu\nu}$. First notice that dx has components $\gamma_\mu = \delta^1_\mu$ as is readily verified by the equation $dx = \delta^1_\mu dx^\mu$. We raise the index using η, as follows*

$$\gamma^\mu = \eta^{\mu\nu}\gamma_\nu = \eta^{\mu\nu}\delta^1_\nu = \eta^{1\mu} = \delta^{1\mu}.$$

Beginning with the definition of the Hodge dual we calculate

$$\begin{aligned} {}^*(dx) &= \tfrac{1}{(4-1)!}\delta^{1\mu}\epsilon_{\mu\nu\alpha\beta}dx^\nu \wedge dx^\alpha \wedge dx^\beta \\ &= (1/6)\epsilon_{1\nu\alpha\beta}dx^\nu \wedge dx^\alpha \wedge dx^\beta \\ \\ &= (1/6)[\epsilon_{1023}dt \wedge dy \wedge dz + \epsilon_{1230}dy \wedge dz \wedge dt + \epsilon_{1302}dz \wedge dt \wedge dy \\ &\qquad +\epsilon_{1320}dz \wedge dy \wedge dt + \epsilon_{1203}dy \wedge dt \wedge dz + \epsilon_{1032}dt \wedge dz \wedge dy] \\ \\ &= (1/6)[-dt \wedge dy \wedge dz - dy \wedge dz \wedge dt - dz \wedge dt \wedge dy \\ &\qquad +dz \wedge dy \wedge dt + dy \wedge dt \wedge dz + dt \wedge dz \wedge dy] \\ \\ &= -dy \wedge dz \wedge dt. \end{aligned} \tag{9.2}$$

The difference between the three and four dimensional Hodge dual arises from two sources, for one we are using the Minkowski metric so indices up or down makes a difference, and second the antisymmetric symbol has more possibilities than before because the Greek indices take four values.

Example 9.1.3. *We find the Hodge dual of $\gamma = dt$ with respect to the Minkowski metric $\eta_{\mu\nu}$. Notice that dt has components $\gamma_\mu = \delta^0_\mu$ as is easily seen using the equation $dt = \delta^0_\mu dx^\mu$. Raising the index using η as usual, we have*

$$\gamma^\mu = \eta^{\mu\nu}\gamma_\nu = \eta^{\mu\nu}\delta^0_\nu = -\eta^{0\mu} = -\delta^{0\mu}$$

where the minus sign is due to the Minkowski metric. Starting with the definition of Hodge duality we calculate

$$\begin{aligned} {}^*(dt) &= -(1/6)\delta^{0\mu}\epsilon_{\mu\nu\alpha\beta}dx^\nu \wedge dx^\alpha \wedge dx^\beta \\ &= -(1/6)\epsilon_{0\nu\alpha\beta}dx^\nu \wedge dx^\alpha \wedge dx^\beta \\ \\ &= -(1/6)\epsilon_{0ijk}dx^i \wedge dx^j \wedge dx^k \\ &= -(1/6)\epsilon_{ijk}dx^i \wedge dx^j \wedge dx^k \\ &= -dx \wedge dy \wedge dz. \end{aligned} \tag{9.3}$$

for the case here we are able to use some of our old three dimensional ideas. The Hodge dual of dt cannot have a dt in it which means our answer will only have dx, dy, dz in it and that is why we were able to shortcut some of the work, (compared to the previous example).

Example 9.1.4. *Finally, we find the Hodge dual of $\gamma = dt \wedge dx$ with respect to the Minkowski metric $\eta_{\mu\nu}$. Recall that ${}^*(dt \wedge dx) = \frac{1}{(4-2)!}\epsilon_{01\mu\nu}\gamma^{01}(dx^\mu \wedge dx^\nu)$ and that $\gamma^{01} = \eta^{0\lambda}\eta^{1\rho}\gamma_{\lambda\rho} = (-1)(1)\gamma_{01} = -1$. Thus*

$$\begin{aligned} {}^*(dt \wedge dx) &= -(1/2)\epsilon_{01\mu\nu}dx^\mu \wedge dx^\nu \\ &= -(1/2)[\epsilon_{0123}dy \wedge dz + \epsilon_{0132}dz \wedge dy] \\ \\ &= -dy \wedge dz. \end{aligned} \tag{9.4}$$

Notice also that since $dt \wedge dx = -dx \wedge dt$ we find $*(dx \wedge dt) = dy \wedge dz$

The other Hodge duals of the basic two-forms follow from similar calculations. Here is a table of all the basic Hodge dualities in Minkowski space, In the table the terms are grouped as they are to emphasize the isomorphisms between the one-dimensional $\Lambda^0(M)$ and $\Lambda^4(M)$, the four-dimensional $\Lambda^1(M)$ and $\Lambda^3(M)$, the six-dimensional $\Lambda^2(M)$ and itself. Notice that the dimension of $\Lambda(M)$ is 16 which just happens to be 2^4.

Now that we've established how the Hodge dual works on the differentials we can easily take the Hodge dual of arbitrary differential forms on Minkowski space. We begin with the example of the 4-current $\mathcal{J}$

Example 9.1.5. Four Current: *often in relativistic physics we would even just call the four current simply the current, however it actually includes the charge density ρ and current density $\vec{J}$. Consequently, we define,*

$$(\mathcal{J}^\mu) \equiv (\rho, \vec{J}),$$

moreover if we lower the index we obtain,

$$(\mathcal{J}_\mu) = (-\rho, \vec{J})$$

which are the components of the current one-form,

$$\mathcal{J} = \mathcal{J}_\mu dx^\mu = -\rho dt + J_x dx + J_y dy + J_z dz$$

This equation could be taken as the definition of the current as it is equivalent to the vector definition. Now we can rewrite the last equation using the vectors $\mapsto$ forms mapping as,

$$\mathcal{J} = -\rho dt + \omega_{\vec{J}}.$$

Consider the Hodge dual of $\mathcal{J}$,

$$\begin{aligned} {}^*\mathcal{J} &= {}^*(-\rho dt + J_x dx + J_y dy + J_z dz) \\ &= -\rho^* dt + J_x{}^* dx + J_y{}^* dy + J_z{}^* dz \\ &= \rho dx \wedge dy \wedge dz - J_x dy \wedge dz \wedge dt - J_y dz \wedge dx \wedge dt - J_z dx \wedge dy \wedge dt \\ &= \rho dx \wedge dy \wedge dz - \Phi_{\vec{J}} \wedge dt. \end{aligned} \tag{9.5}$$

we will find it useful to appeal to this calculation in a later section.

Example 9.1.6. Four Potential: *often in relativistic physics we would call the four potential simply the potential, however it actually includes the scalar potential V and the vector potential $\vec{A}$ (discussed at the end of chapter 3). To be precise we define,*

$$(A^\mu) \equiv (V, \vec{A})$$

we can lower the index to obtain,

$$(A_\mu) = (-V, \vec{A})$$

which are the components of the current one-form,

$$A = A_\mu dx^\mu = -V dt + A_x dx + A_y dy + A_z dz$$

Sometimes this equation is taken as the definition of the four potential. We can rewrite the four potential vector field using the vectors $\mapsto$ forms mapping as,

$$A = -V dt + \omega_{\vec{A}}.$$

The Hodge dual of A is

$${}^*A = V dx \wedge dy \wedge dz - \Phi_{\vec{A}} \wedge dt. \tag{9.6}$$

Several steps were omitted because they are identical to the calculation of the dual of the 4-current above.

Definition 9.1.7. *Faraday tensor.*

Given an electric field $\vec{E} = (E_1, E_2, E_3)$ and a magnetic field $\vec{B} = (B_1, B_2, B_3)$ we define a 2-form F by

$$F = \omega_E \wedge dt + \Phi_B.$$

This 2-form is often called the **electromagnetic field tensor or the Faraday tensor.** If we write it in tensor components as $F = \frac{1}{2} F_{\mu\nu} dx^\mu \wedge dx^\nu$ and then consider its matrix $(F_{\mu\nu})$ of components then it is easy to see that

$$(F_{\mu\nu}) = \begin{pmatrix} 0 & -E_1 & -E_2 & -E_3 \\ E_1 & 0 & B_3 & -B_2 \\ E_2 & -B_3 & 0 & B_1 \\ E_3 & B_2 & -B_1 & 0 \end{pmatrix} \tag{9.7}$$

Convention: Notice that when we write the matrix version of the tensor components we take the first index to be the row index and the second index to be the column index, that means $F_{01} = -E_1$ whereas $F_{10} = E_1$.

Example 9.1.8. *In this example we demonstrate various conventions which show how one can transform the field tensor to other type tensors. Define a type $(1,1)$ tensor by raising the first index by the inverse metric $\eta^{\alpha\mu}$ as follows,*

$$F^{\alpha}{}_{\nu} = \eta^{\alpha\mu} F_{\mu\nu}$$

The zeroth row,

$$(F^0{}_\nu) = (\eta^{0\mu} F_{\mu\nu}) = (0, E_1, E_2, E_3)$$

Then row one is unchanged since $\eta^{1\mu} = \delta^{1\mu}$,

$$(F^1{}_\nu) = (\eta^{1\mu} F_{\mu\nu}) = (E_1, 0, B_3, -B_2)$$

and likewise for rows two and three. In summary the (1,1) tensor $F' = F^\alpha_\nu (\frac{\partial}{\partial x^\alpha} \otimes dx^\nu)$ has the components below

$$(F^\alpha{}_\nu) = \begin{pmatrix} 0 & E_1 & E_2 & E_3 \\ E_1 & 0 & B_3 & -B_2 \\ E_2 & -B_3 & 0 & B_1 \\ E_3 & B_2 & -B_1 & 0 \end{pmatrix}. \tag{9.8}$$

At this point we raise the other index to create a $(2,0)$ tensor,

$$\boxed{F^{\alpha\beta} = \eta^{\alpha\mu}\eta^{\beta\nu} F_{\mu\nu}} \tag{9.9}$$

and we see that it takes one copy of the inverse metric to raise each index and $F^{\alpha\beta} = \eta^{\beta\nu} F^\alpha{}_\nu$ so we can pick up where we left off in the $(1,1)$ case. We could proceed case by case like we did with the $(1,1)$ case but it is better to use matrix multiplication. Notice that $\eta^{\beta\nu} F^\alpha{}_\nu = F^\alpha{}_\nu \eta^{\nu\beta}$ is just the (α,β) component of the following matrix product,

$$(F^{\alpha\beta}) = \begin{pmatrix} 0 & E_1 & E_2 & E_3 \\ E_1 & 0 & B_3 & -B_2 \\ E_2 & -B_3 & 0 & B_1 \\ E_3 & B_2 & -B_1 & 0 \end{pmatrix} \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & E_1 & E_2 & E_3 \\ -E_1 & 0 & B_3 & -B_2 \\ -E_2 & -B_3 & 0 & B_1 \\ -E_3 & B_2 & -B_1 & 0 \end{pmatrix}. \tag{9.10}$$

So we find a $(2,0)$ tensor $F'' = F^{\alpha\beta}(\frac{\partial}{\partial x^\alpha} \otimes \frac{\partial}{\partial x^\beta})$. Other books might even use the same symbol F for F' and F'', it is in fact typically clear from the context which version of F one is thinking about. Pragmatically physicists just write the components so its not even an issue for them.

Example 9.1.9. Field tensor's dual: *We now calculate the Hodge dual of the field tensor,*

$$\begin{aligned} {}^*F &= {}^*(\omega_E \wedge dt + \Phi_B) \\ &= E_x{}^*(dx \wedge dt) + E_y{}^*(dy \wedge dt) + E_z{}^*(dz \wedge dt) \\ &\quad + B_x{}^*(dy \wedge dz) + B_y{}^*(dz \wedge dx) + B_z{}^*(dx \wedge dy) \\ &= E_x dy \wedge dz + E_y dz \wedge dx + E_z dx \wedge dy \\ &\quad - B_x dx \wedge dt - B_y dy \wedge dt - B_z dz \wedge dt \\ &= \Phi_E - \omega_B \wedge dt. \end{aligned}$$

we can also write the components of ${}^*F$ in matrix form:

$$({}^*F_{\mu\nu}) = \begin{pmatrix} 0 & B_1 & B_2 & B_3 \\ -B_1 & 0 & E_3 & -E_2 \\ -B_2 & -E_3 & 0 & E_1 \\ -B_3 & E_2 & -E_1 & 0. \end{pmatrix} \tag{9.11}$$

Notice that the net-effect of Hodge duality on the field tensor was to make the exchanges $\vec{E} \mapsto -\vec{B}$ and $\vec{B} \mapsto \vec{E}$.

9.2 exterior derivatives of charge forms, field tensors, and their duals

In the last chapter we found that the single operation of the exterior differentiation reproduces the gradiant, curl and divergence of vector calculus provided we make the appropriate identifications under the "work" and "flux" form mappings. We now move on to some four dimensional examples.

Example 9.2.1. Charge conservation: *Consider the 4-current we introduced in example* 9.1.5. *Take the exterior derivative of the dual of the current to get,*

$$\begin{aligned} d({}^*\mathcal{J}) &= d(\rho dx \wedge dy \wedge dz - \Phi_{\vec{J}} \wedge dt) \\ &= (\partial_t \rho) dt \wedge dx \wedge dy \wedge dz - d[(J_x dy \wedge dz + J_y dz \wedge dx + J_z dx \wedge dy) \wedge dt] \\ &= d\rho \wedge dx \wedge dy \wedge dz \\ &\quad -\partial_x J_x dx \wedge dy \wedge dz \wedge dt - \partial_y J_y dy \wedge dz \wedge dx \wedge dt - \partial_z J_z dz \wedge dx \wedge dy \wedge dt \\ &= (\partial_t \rho + \nabla \cdot \vec{J}) dt \wedge dx \wedge dy \wedge dz. \end{aligned}$$

We work through the same calculation using index techniques,

$$\begin{aligned} d({}^*\mathcal{J}) &= d(\rho dx \wedge dy \wedge dz - \Phi_{\vec{J}} \wedge dt) \\ &= d(\rho) \wedge dx \wedge dy \wedge dz - d[\tfrac{1}{2}\epsilon_{ijk} J_i dx^j \wedge dx^k \wedge dt) \\ &= (\partial_t \rho) dt \wedge dx \wedge dy \wedge dz - \tfrac{1}{2}\epsilon_{ijk} \partial_\mu J_i dx^\mu \wedge dx^j \wedge dx^k \wedge dt \\ &= (\partial_t \rho) dt \wedge dx \wedge dy \wedge dz - \tfrac{1}{2}\epsilon_{ijk} \partial_m J_i dx^m \wedge dx^j \wedge dx^k \wedge dt \\ &= (\partial_t \rho) dt \wedge dx \wedge dy \wedge dz - \tfrac{1}{2}\epsilon_{ijk} \epsilon_{mjk} \partial_m J_i dx \wedge dy \wedge dz \wedge dt \\ &= (\partial_t \rho) dt \wedge dx \wedge dy \wedge dz - \tfrac{1}{2} 2\delta_{im} \partial_m J_i dx \wedge dy \wedge dz \wedge dt \\ &= (\partial_t \rho + \nabla \cdot \vec{J}) dt \wedge dx \wedge dy \wedge dz. \end{aligned}$$

Observe that we can now phrase charge conservation by the following equation

$$d({}^*\mathcal{J}) = 0 \qquad \Longleftrightarrow \qquad \partial_t \rho + \nabla \cdot \vec{J} = 0.$$

In the classical scheme of things this was a derived consequence of the equations of electromagnetism, however it is possible to build the theory regarding this equation as fundamental. Rindler describes that formal approach in a late chapter of "Introduction to Special Relativity".

Proposition 9.2.2.

If $(A_\mu) = (-V, \vec{A})$ is the vector potential (which gives the magnetic field) and $A = -V dt + \omega_{\vec{A}}$, then $dA = \omega_{\vec{E}} + \Phi_{\vec{B}} = F$ where F is the electromagnetic field tensor. Moreover, $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$.

Proof: The proof uses the definitions $\vec{E} = -\nabla V - \partial_t A$ and $\vec{B} = \nabla \times \vec{A}$ and some vector identities:

$$\begin{aligned}
dA &= d(-V dt + \omega_{\vec{A}}) \\
&= -dV \wedge dt + d(\omega_{\vec{A}}) \\
&= -dV \wedge dt + (\partial_t A_i) dt \wedge dx^i + (\partial_j A_i) dx^j \wedge dx^i \\
&= \omega_{(-\nabla V)} \wedge dt - \omega_{\partial_t \vec{A}} \wedge dt + \Phi_{\nabla \times \vec{A}} \\
&= (\omega_{(-\nabla V)} - \omega_{\partial_t \vec{A}}) \wedge dt + \Phi_{\nabla \times \vec{A}} \\
&= \omega_{(-\nabla V - \partial_t \vec{A})} \wedge dt + \Phi_{\nabla \times \vec{A}} \\
&= \omega_{\vec{E}} \wedge dt + \Phi_{\vec{B}} \\
&= F = \frac{1}{2} F_{\mu\nu} dx^\mu \wedge dx^\nu.
\end{aligned}$$

Moreover we also have:

$$\begin{aligned}
dA &= d(A_\nu) \wedge dx^\nu \\
&= \partial_\mu A_\nu dx^\mu \wedge dx^\nu \\
&= \tfrac{1}{2}(\partial_\mu A_\nu - \partial_\nu A_\mu) dx^\mu \wedge dx^\nu + \tfrac{1}{2}(\partial_\mu A_\nu + \partial_\nu A_\mu) dx^\mu \wedge dx^\nu \\
&= \tfrac{1}{2}(\partial_\mu A_\nu - \partial_\nu A_\mu) dx^\mu \wedge dx^\nu.
\end{aligned}$$

Comparing the two identities we see that $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ and the proposition follows.

Example 9.2.3. Exterior derivative of the field tensor: *We have just seen that the field tensor is the exterior derivative of the potential one-form. We now compute the exterior derivative of the field tensor expecting to find Maxwell's equations since the derivative of the fields are governed by Maxwell's equations,*

$$\begin{aligned}
dF &= d(E_i dx^i \wedge dt) + d(\Phi_{\vec{B}}) \\
&= \partial_m E_i (dx^m \wedge dx^i \wedge dt) + (\nabla \cdot \vec{B}) dx \wedge dy \wedge dz + \tfrac{1}{2}\epsilon_{ijk}(\partial_t B_i)(dt \wedge dx^j \wedge dx^k).
\end{aligned} \tag{9.12}$$

W pause here to explain our logic. In the above we dropped the $\partial_t E_i dt \wedge \cdots$ term because it was wedged with another dt in the term so it vanished. Also we broke up the exterior derivative on the flux form of $\vec{B}$ into the space and then time derivative terms and used our work in example 8.6.7. *Continuing the calculation,*

$$\begin{aligned}
dF &= [\partial_j E_k + \tfrac{1}{2}\epsilon_{ijk}(\partial_t B_i)] dx^j \wedge dx^k \wedge dt + (\nabla \cdot \vec{B}) dx \wedge dy \wedge dz \\
&= [\partial_x E_y - \partial_y E_x + \epsilon_{i12}(\partial_t B_i)] dx \wedge dy \wedge dt \\
&\quad + [\partial_z E_x - \partial_x E_z + \epsilon_{i31}(\partial_t B_i)] dz \wedge dx \wedge dt \\
&\quad + [\partial_y E_z - \partial_z E_y + \epsilon_{i23}(\partial_t B_i)] dy \wedge dz \wedge dt \\
&\quad + (\nabla \cdot \vec{B}) dx \wedge dy \wedge dz \\
&= (\nabla \times \vec{E} + \partial_t \vec{B})_i \Phi_{e_i} \wedge dt + (\nabla \cdot \vec{B}) dx \wedge dy \wedge dz \\
&= \Phi_{\nabla \times \vec{E} + \partial_t \vec{B}} \wedge dt + (\nabla \cdot \vec{B}) dx \wedge dy \wedge dz
\end{aligned} \tag{9.13}$$

where we used the fact that Φ is an isomorphism of vector spaces (at a point) and $\Phi_{e_1} = dy \wedge dz$, $\Phi_{e_2} = dz \wedge dx$, and $\Phi_{e_3} = dx \wedge dy$. Behold, we can state two of Maxwell's equations as

$$\boxed{dF = 0 \quad \Longleftrightarrow \quad \nabla \times \vec{E} + \partial_t \vec{B} = 0, \quad \nabla \cdot \vec{B} = 0} \tag{9.14}$$

Example 9.2.4. We now compute the exterior derivative of the dual to the field tensor:

$$\begin{aligned} d^*F &= d(-B_i dx^i \wedge dt) + d(\Phi_{\vec{E}}) \\ &= \Phi_{-\nabla\times\vec{B}+\partial_t\vec{E}} \wedge dt + (\nabla\cdot\vec{E})dx\wedge dy\wedge dz \end{aligned} \tag{9.15}$$

This follows directly from the last example by replacing $\vec{E} \mapsto -\vec{B}$ and $\vec{B} \mapsto \vec{E}$. We obtain the two inhomogeneous Maxwell's equations by setting d^*F equal to the Hodge dual of the 4-current,

$$\boxed{d^*F = \mu_o {}^*\mathcal{J} \qquad \iff \qquad -\nabla\times\vec{B} + \partial_t\vec{E} = -\mu_o\vec{J}, \qquad \nabla\cdot\vec{E} = \rho} \tag{9.16}$$

Here we have used example 9.1.5 *to find the RHS of the Maxwell equations.*

We now know how to write Maxwell's equations via differential forms. The stage is set to prove that Maxwell's equations are Lorentz covariant, that is they have the same form in all inertial frames.

9.3 coderivatives and comparing to Griffith's relativitic E & M

Optional section, for those who wish to compare our tensorial E & M with that of Griffith's, you may skip ahead to the next section if not interested

I should mention that this is not the only way to phrase Maxwell's equations in terms of differential forms. If you try to see how what we have done here compares with the equations presented in Griffith's text it is not immediately obvious. He works with $F^{\mu\nu}$ and $G^{\mu\nu}$ and J^{μ} none of which are the components of differential forms. Nevertheless he recovers Maxwell's equations as $\partial_\mu F^{\mu\nu} = J^\nu$ and $\partial_\mu G^{\mu\nu} = 0$. If we compare the components of ${}^*F$ with equation 12.119 (the matrix form of $G^{\mu\nu}$) in Griffith's text,

$$(G^{\mu\nu}(c=1)) = \begin{pmatrix} 0 & B_1 & B_2 & B_3 \\ -B_1 & 0 & -E_3 & E_2 \\ -B_2 & -E_3 & 0 & -E_1 \\ -B_3 & E_2 & -E_1 & 0 \end{pmatrix} = -({}^*F^{\mu\nu}). \tag{9.17}$$

we find that we obtain the negative of Griffith's "dual tensor" (recall that raising the indices has the net-effect of multiplying the zeroth row and column by -1). The equation $\partial_\mu F^{\mu\nu} = J^\nu$ does not follow directly from an exterior derivative, rather it is the component form of a "coderivative". The coderivative is defined $\delta = {}^*d^*$, it takes a p-form to an $(n-p)$-form then d makes it a $(n-p+1)$-form then finally the second Hodge dual takes it to an $(n-(n-p+1))$-form. That is δ takes a p-form to a $p-1$-form. We stated Maxwell's equations as

$$dF = 0 \qquad\qquad d^*F = {}^*\mathcal{J}$$

Now we can take the Hodge dual of the inhomogeneous equation to obtain,

$${}^*d^*F = \delta F = {}^{**}\mathcal{J} = \pm\mathcal{J}$$

where I leave the sign for you to figure out. Then the other equation

$$\partial_\mu G^{\mu\nu} = 0$$

can be understood as the component form of $\delta^*F = 0$ but this is really $dF = 0$ in disguise,

$$0 = \delta^*F = {}^*d^{**}F = \pm{}^*dF \iff dF = 0$$

so even though it looks like Griffith's is using the dual field tensor for the homogeneous Maxwell's equations and the field tensor for the inhomogeneous Maxwell's equations it is in fact not the case. The key point is that there are **coderivatives** implicit within Griffith's equations, so you have to read between the lines a little to see how it matched up with what we've done here. I have not entirely proved it here, to be complete we should look at the component form of $\delta F = \mathcal{J}$ and explicitly show that this gives us $\partial_\mu F^{\mu\nu} = J^\nu$, I don't think it is terribly difficult but I'll leave it to the reader.

Comparing with Griffith's is fairly straightforward because he uses the same metric as we have. Other texts use the mostly negative metric, its just a convention. If you try to compare to such a book you'll find that our equations are almost the same up to a sign. One good careful book is Reinhold A. Bertlmann's *Anomalies in Quantum Field Theory* you will find much of what we have done here done there with respect to the other metric. Another good book which shares our conventions is Sean M. Carroll's *An Introduction to General Relativity: Spacetime and Geometry*, that text has a no-nonsense introduction to tensors forms and much more over a curved space (in contrast to our approach which has been over a vector space which is flat). By now there are probably thousands of texts on tensors; these are a few we have found useful here.

9.4 Maxwell's equations are relativistically covariant

Let us begin with the definition of the field tensor once more. We define the components of the field tensor in terms of the 4-potentials as we take the view-point those are the basic objects (not the fields). If

$$F_{\mu\nu} \equiv \partial_\mu A_\nu - \partial_\nu A_\mu,$$

then the field tensor $F = F_{\mu\nu} dx^\mu \otimes dx^\nu$ is a tensor, or is it ? We should check that the components transform as they ought according to the discussion in section **??**. Let $\bar{x}^\mu = \Lambda^\mu_\nu x^\nu$ then we observe,

$$\begin{array}{l} (1.)\ \bar{A}_\mu = (\Lambda^{-1})^\alpha_\mu A_\alpha \\ (2.)\ \frac{\partial}{\partial \bar{x}^\nu} = \frac{\partial x^\beta}{\partial \bar{x}^\nu}\frac{\partial}{\partial x^\beta} = (\Lambda^{-1})^\beta_\nu \frac{\partial}{\partial x^\beta} \end{array} \tag{9.18}$$

where (2.) is simply the chain rule of multivariate calculus and (1.) is not at all obvious. We will assume that (1.) holds, that is we assume that the 4-potential transforms in the appropriate way for a one-form. In principle one could prove that from more base assumptions. After all electromagnetism is the study of the interaction of charged objects, we should hope that the potentials are derivable from the source charge distribution. Indeed, there exist formulas to calculate the potentials for moving distributions of charge. We could take those as definitions for the potentials, then it would be possible to actually calculate if (1.) is true. We'd just change coordinates via a Lorentz transformation and verify (1.). For the sake of brevity we will just assume that (1.) holds. We should mention that alternatively one can show the electric and magnetic fields transform as to make $F_{\mu\nu}$ a tensor. Those derivations assume that charge is an invariant quantity and just apply Lorentz transformations to special physical situations to deduce the field transformation rules. See Griffith's chapter on special relativity or look in Resnick for example.

Let us find how the field tensor transforms assuming that (1.) and (2.) hold, again we consider $\bar{x}^\mu = \Lambda^\mu_\nu x^\nu$,

$$\begin{aligned} \bar{F}_{\mu\nu} &= \bar{\partial}_\mu \bar{A}_\nu - \bar{\partial}_\nu \bar{A}_\mu \\ &= (\Lambda^{-1})^\alpha_\mu \partial_\alpha((\Lambda^{-1})^\beta_\nu A_\beta) - (\Lambda^{-1})^\beta_\nu \partial_\beta((\Lambda^{-1})^\alpha_\mu A_\alpha) \\ &= (\Lambda^{-1})^\alpha_\mu (\Lambda^{-1})^\beta_\nu (\partial_\alpha A_\beta - \partial_\beta A_\alpha) \\ &= (\Lambda^{-1})^\alpha_\mu (\Lambda^{-1})^\beta_\nu F_{\alpha\beta}. \end{aligned} \tag{9.19}$$

therefore the field tensor really is a tensor over Minkowski space.

Proposition 9.4.1.

> The dual to the field tensor is a tensor over Minkowski space. For a given Lorentz transformation $\bar{x}^\mu = \Lambda^\mu_\nu x^\nu$ it follows that
>
> $$^*\bar{F}_{\mu\nu} = (\Lambda^{-1})^\alpha_\mu (\Lambda^{-1})^\beta_\nu {}^*F_{\alpha\beta}$$

Proof: homework (just kidding in 2010), it follows quickly from the definition and the fact we already know that the field tensor is a tensor.

Proposition 9.4.2.

> The four-current is a four-vector. That is under the Lorentz transformation $\bar{x}^\mu = \Lambda^\mu_\nu x^\nu$ we can show,
>
> $$\bar{\mathcal{J}}_\mu = (\Lambda^{-1})^\alpha_\mu \mathcal{J}_\alpha$$

Proof: follows from arguments involving the invariance of charge, time dilation and length contraction. See Griffith's for details, sorry we have no time.

Corollary 9.4.3.

> The dual to the four current transforms as a 3-form. That is under the Lorentz transformation $\bar{x}^\mu = \Lambda^\mu_\nu x^\nu$ we can show,
>
> $$^*\bar{\mathcal{J}}_{\mu\nu\sigma} = (\Lambda^{-1})^\alpha_\mu (\Lambda^{-1})^\beta_\nu (\Lambda^{-1})^\gamma_\sigma \mathcal{J}_{\alpha\beta\gamma}$$

Up to now the content of this section is simply an admission that we have been a little careless in defining things upto this point. The main point is that if we say that something is a tensor then we need to make sure that is in fact the case. With the knowledge that our tensors are indeed tensors the proof of the covariance of Maxwell's equations is trivial.

$$dF = 0 \qquad\qquad d^*F = {}^*\mathcal{J}$$

are coordinate invariant expressions which we have already proved give Maxwell's equations in one frame of reference, thus they must give Maxwell's equations in all frames of reference.
The essential point is simply that

$$F = \frac{1}{2}F_{\mu\nu}dx^\mu \wedge dx^\nu = \frac{1}{2}\bar{F}_{\mu\nu}d\bar{x}^\mu \wedge d\bar{x}^\nu$$

Again, we have no hope for the equation above to be true unless we know that
$\bar{F}_{\mu\nu} = (\Lambda^{-1})^\alpha_\mu (\Lambda^{-1})^\beta_\nu F_{\alpha\beta}$. That transformation follows from the fact that the four-potential is a four-vector. It should be mentioned that others prefer to "prove" the field tensor is a tensor by studying how the electric and magnetic fields transform under a Lorentz transformation. We in contrast have derived the field transforms based ultimately on the seemingly innocuous assumption that the four-potential transforms according to $\bar{A}_\mu = (\Lambda^{-1})^\alpha_\mu A_\alpha$. OK enough about that.

So the fact that Maxwell's equations have the same form in all **relativistically inertial** frames of reference simply stems from the fact that we found Maxwell's equation were given by an arbitrary frame, and the field tensor looks the same in the new barred frame so we can again go through all the same arguments with barred coordinates. Thus we find that Maxwell's equations are the same in all relativistic frames of reference, that is if they hold in one inertial frame then they will hold in any other frame which is related by a Lorentz transformation.

9.5 Electrostatics in Five dimensions

We will endeavor to determine the electric field of a point charge in 5 dimensions where we are thinking of adding an extra spatial dimension. Lets call the fourth spatial dimension the w-direction so that a typical point in space time will be (t, x, y, z, w). First we note that the electromagnetic field tensor can still be derived from a one-form potential,

$$A = -\rho dt + A_1 dx + A_2 dy + A_3 dz + A_4 dw$$

we will find it convenient to make our convention for this section that $\mu, \nu, ... = 0, 1, 2, 3, 4$ whereas $m, n, ... = 1, 2, 3, 4$ so we can rewrite the potential one-form as,

$$A = -\rho dt + A_m dx^m$$

This is derived from the vector potential $A^\mu = (\rho, A^m)$ under the assumption we use the natural generalization of the Minkowski metric, namely the 5 by 5 matrix,

$$(\eta_{\mu\nu}) = \begin{pmatrix} -1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} = (\eta^{\mu\nu}) \tag{9.20}$$

we could study the linear isometries of this metric, they would form the group $O(1, 4)$. Now we form the field tensor by taking the exterior derivative of the one-form potential,

$$F = dA = \frac{1}{2}(\partial_\mu \partial_\nu - \partial_\nu \partial_\mu) dx^\mu \wedge dx^\nu$$

now we would like to find the electric and magnetic "fields" in 4 dimensions. Perhaps we should say 4+1 dimensions, just understand that I take there to be 4 spatial directions throughout this discussion if in doubt. Note that we are faced with a dilemma of interpretation. There are 10 independent components of a 5 by 5 antisymmetric tensor, naively we wold expect that the electric and magnetic fields each would have 4 components, but that is not possible, we'd be missing two components. The solution is this, the time components of the field tensor are understood to correspond to the electric part of the fields whereas the remaining 6 components are said to be magnetic. This aligns with what we found in 3 dimensions, its just in 3 dimensions we had the fortunate quirk that the number of linearly independent one and two forms were equal at any point. This definition means that the magnetic field will in general not be a vector field but rather a "flux" encoded by a 2-form.

$$(F_{\mu\nu}) = \begin{pmatrix} 0 & -E_x & -E_y & -E_z & -E_w \\ E_x & 0 & B_z & -B_y & H_1 \\ E_y & -B_z & 0 & B_x & H_2 \\ E_z & B_y & -B_x & 0 & H_3 \\ E_w & -H_1 & -H_2 & -H_3 & 0 \end{pmatrix} \tag{9.21}$$

Now we can write this compactly via the following equation,

$$F = E \wedge dt + B$$

I admit there are subtle points about how exactly we should interpret the magnetic field, however I'm going to leave that to your imagination and instead focus on the electric sector. What is the generalized Maxwell's equation that E must satisfy?

$$d^*F = \mu_o {}^*\mathcal{J} \implies d^*(E \wedge dt + B) = \mu_o {}^*\mathcal{J}$$

where $\mathcal{J} = -\rho dt + J_m dx^m$ so the 5 dimensional Hodge dual will give us a $5 - 1 = 4$ form, in particular we will be interested in just the term stemming from the dual of dt,

$$^*(-\rho dt) = \rho dx \wedge dy \wedge dz \wedge dw$$

the corresponding term in d^*F is $d^*(E \wedge dt)$ thus, using $\mu_o = \frac{1}{\epsilon_o}$,

$$d^*(E \wedge dt) = \frac{1}{\epsilon_o}\rho dx \wedge dy \wedge dz \wedge dw \tag{9.22}$$

is the 4-dimensional Gauss's equation. Now consider the case we have an isolated point charge which has somehow always existed at the origin. Moreover consider a 3-sphere that surrounds the charge. We wish to determine the generalized Coulomb field due to the point charge. First we note that the solid 3-sphere is a 4-dimensional object, it the set of all $(x, y, z, w) \in \mathbb{R}^4$ such that

$$x^2 + y^2 + z^2 + w^2 \leq r^2$$

We may parametrize a three-sphere of radius r via generalized spherical coordinates,

$$\begin{aligned} x &= rsin(\theta)cos(\phi)sin(\psi) \\ y &= rsin(\theta)sin(\phi)sin(\psi) \\ z &= rcos(\theta)sin(\psi) \\ w &= rcos(\psi) \end{aligned} \tag{9.23}$$

Now it can be shown that the volume and surface area of the radius r three-sphere are as follows,

$$vol(S^3) = \frac{\pi^2}{2}r^4 \qquad\qquad area(S^3) = 2\pi^2 r^3$$

We may write the charge density of a smeared out point charge q as,

$$\rho = \begin{cases} 2q/\pi^2 a^4, & 0 \leq r \leq a \\ 0, & r > a \end{cases}. \tag{9.24}$$

Notice that if we integrate ρ over any four-dimensional region which contains the solid three sphere of radius a will give the enclosed charge to be q. Then integrate over the Gaussian 3-sphere S^3 with radius r call it M,

$$\int_M d^*(E \wedge dt) = \frac{1}{\epsilon_o}\int_M \rho dx \wedge dy \wedge dz \wedge dw$$

now use the Generalized Stokes Theorem to deduce,

$$\int_{\partial M} {}^*(E \wedge dt) = \frac{q}{\epsilon_o}$$

but by the "spherical" symmetry of the problem we find that E must be independent of the direction it points, this means that it can only have a radial component. Thus we may calculate the integral with respect to generalized spherical coordinates and we will find that it is the product of $E_r \equiv E$ and the surface volume of the four dimensional solid three sphere. That is,

$$\int_{\partial M} {}^*(E \wedge dt) = 2\pi^2 r^3 E = \frac{q}{\epsilon_o}$$

Thus,

$$\boxed{E = \frac{q}{2\pi^2 \epsilon_o r^3}}$$

the Coulomb field is weaker if it were to propogate in 4 spatial dimensions. Qualitatively what has happened is that the have taken the same net flux and spread it out over an additional dimension, this means it thins out quicker. A very similar idea is used in some *brane world* scenarios. String theorists posit that the gravitational field spreads out in more than four dimensions while in contrast the standard model fields of electromagnetism, and the strong and weak forces are confined to a four-dimensional brane. That sort of model attempts an explanation as to why gravity is so weak in comparison to the other forces. Also it gives large scale corrections to gravity that some hope will match observations which at present don't seem to fit the standard gravitational models.

This example is but a taste of the theoretical discussion that differential forms allow. As a final comment I remind the reader that we have done things for flat space for the most part in this course, when considering a curved space there are a few extra considerations that must enter. Coordinate vector fields e_i must be thought of as derivations $\partial/\partial x^\mu$ for one. Also the metric is not a constant tensor like δ_{ij} or $\eta_{\mu\nu}$ rather is depends on position, this means Hodge duality aquires a coordinate dependence as well. Doubtless I have forgotten something else in this brief warning. One more advanced treatment of many of our discussions is Dr. Fulp's Fiber Bundles 2001 notes which I have posted on my webpage. He uses the other metric but it is rather elegantly argued, all his arguments are coordinate independent. He also deals with the issue of the magnetic induction and the dielectric, issues which we have entirely ignored since we always have worked in free space.

References and Acknowledgements:

I have drawn from many sources to assemble the content of the last couple chapters, the references are listed approximately in the order of their use to the course, additionally we are indebted to Dr. Fulp for his course notes from many courses (ma 430, ma 518, ma 555, ma 756, ...). Also Manuela Kulaxizi helped me towards the correct (I hope) interpretation of 5-dimensional E&M in the last example.

Vector Calculus, Susan Jane Colley

Introduction to Special Relativity, Robert Resnick

Differential Forms and Connections, R.W.R. Darling

Differential geometry, gauge theories, and gravity, M. Göckerler & T. Schücker

Anomalies in Quantum Field Theory, Reinhold A. Bertlmann

”The Differential Geometry and Physical Basis for the Applications of Feynman Diagrams”, S.L. Marateck, Notices of the AMS, Vol. 53, Number 7, pp. 744-752

Abstract Linear Algebra, Morton L. Curtis

Gravitation, Misner, Thorne and Wheeler

Introduction to Special Relativity, Wolfgang Rindler

Differential Forms A Complement to Vector Calculus, Steven H. Weintraub

Differential Forms with Applications to the Physical Sciences, Harley Flanders

Introduction to Electrodynamics, (3rd ed.) David J. Griffiths

The Geometry of Physics: An Introduction, Theodore Frankel

An Introduction to General Relativity: Spacetime and Geometry, Sean M. Carroll

Gauge Theory and Variational Principles, David Bleeker

Group Theory in Physics, Wu-Ki Tung

Chapter 10

Banach's Fixed Point Theorem and its Applications

In this chapter we'll prove Banach's Fixed Point Theorem which is also known as the contraction mapping theorem. Then we will study its application in several contexts. First, Newton's Method shows us how to find roots of the equation $f(x) = 0$. Then, we see how a very similar construction allows us to prove an inverse function theorem for functions on $\mathbb{R}$. We then turn to developing some background to understand the inverse mapping theorem. We introduce the derivative for maps from $\mathbb{R}^n \to \mathbb{R}^m$ and show how the mean value theorem generalizes to such a context. Once this background is settled, we state and prove the inverse mapping theorem. The remainder of the chapter is devoted showing how the contraction mapping theorem allows us to establish the major existence and uniqueness theorems for ordinary differential equations.

10.1 Contraction Mapping

Let me begin by review a well-known bit of calculus; the *geometric series.* If $r \in (-1, 1)$ then define $S_n = 1 + r + \cdots + r^{n-1} + r^n$ and notice $rS_n = r + r^2 + \cdots + r^n + r^{n+1}$ hence

$$S_n - rS_n = (1 + r + \cdots + r^{n-1} + r^n) - (r + r^2 + \cdots + r^n + r^{n+1}) = 1 - r^{n+1}.$$

Thus,

$$S_n = \frac{1 - r^{n+1}}{1 - r}$$

The sequence $\{S_n\}$ converges since $|r| < 1$ implies $r^{n+1} \to 0$ as $n \to \infty$ and so

$$\lim_{n\to\infty} S_n = \lim_{n\to\infty} \frac{1 - r^{n+1}}{1 - r} = \frac{1}{1 - r}.$$

Since $\{S_n\}$ is a convergent sequence it is also a Cauchy sequence. Suppose $0 < r < 1$. If $\varepsilon > 0$ then there exists $N \in \mathbb{N}$ for which $N < m < n$ implies $|S_n - S_m| < \varepsilon$. Thus, for such $m < n$ algebra yields:

$$|S_n - S_m| = r^n + \cdots + r^{m+1} < \varepsilon$$

Let $S = 1 + r + r^2 + \cdots$ then for $|r| < 1$ we have shown $S = \frac{1}{1-r}$. Moreover, this bounds the magnitude of the error in the n-th partial sum S_n:

$$S_n = \frac{1}{1 - r} - \frac{r^{n+1}}{1 - r} \quad \Rightarrow \quad |S - S_n| = \frac{|r|^{n+1}}{1 - r}$$

The error in the n-th partial is alternatively viewed as the tail of the series since:

$$S - S_n = \sum_{k=1}^{\infty} r^k - \sum_{k=1}^{n} r^k = \sum_{k=n+1}^{\infty} r^k \quad \Rightarrow \quad |r^{n+1} + r^{n+2} + \cdots| = \frac{|r|^{n+1}}{1-r}$$

Reindexing a bit for future convenience, for $|r| < 1$,

$$\boxed{|r^m + r^{m+1} + \cdots| = \frac{|r|^m}{1-r}}$$

Definition 10.1.1. *Let V be a normed linear space with norm $\|\cdot\|$. Suppose $M \subseteq V$ and let $\varphi : M \to M$ be a mapping. If there exists a constant $\alpha \in [0,1)$ for which $\|\varphi(x) - \varphi(y)\| \leq \alpha\|x-y\|$ for all $x, y \in M$ then φ is called a* **contraction mapping** *on M with* **contraction constant** *α.*

While our primary interest is the context of the definition above[1], the concept of fixed point is more general so I will intentionally be vague in what follows:

Definition 10.1.2. *Let φ be a map. If $\varphi(x_o) = x_o$ then we say x_o is a* **fixed point** *of φ.*

The words *map* and *function* are used interchangeably in this document [2].

Theorem 10.1.3. Banach's Fixed Point Theorem:
Let V be a Banach space with norm $\|\cdot\|$ and let M be a closed subset of V. If $\varphi : M \to M$ is a contraction mapping then φ has a unique fixed point $x_\star \in M$. Moreover, if $x \in M$ and we define $x_0 = x$ and $x_n = \varphi^n(x)$ for $n = 1, 2, \dots$ then $\{x_n\}_{n=1}^{\infty}$ is a Cauchy sequence which converges to $x_\star$. Furthermore, the distance between the n-th iterate and the fixed point is bounded according to

$$\|x_\star - x_n\| \leq \frac{\alpha^n \|x_1 - x_0\|}{1-\alpha}.$$

Proof: since V is complete we know that a Cauchy sequence in V necessarily converges. Moreover, since M is closed, a sequence in M must converge to a point in M. Therefore, if we can show a sequence in M is Cauchy then it follows that the limit of the sequence is a point in M.

Let $x \in M$ and recursively define $x_0 = x$ and $x_n = \varphi(x_{n-1})$ for each $n \in \mathbb{N}$. Notice[3] $\varphi : M \to M$ thus $\{x_n\}$ is a sequence in M. Let $\alpha \in (0,1)$ be the contraction constant of φ and observe

$$\|\varphi(x_1) - \varphi(x_0)\| \leq \alpha\|x_1 - x_0\| \quad \Rightarrow \quad \|x_2 - x_1\| \leq \alpha\|x_1 - x_0\|$$

Furthermore[4],

$$\|\varphi(x_2) - \varphi(x_1)\| \leq \alpha\|x_2 - x_1\| \leq \alpha^2\|x_1 - x_0\| \quad \Rightarrow \quad \|x_3 - x_2\| \leq \alpha^2\|x_1 - x_0\|.$$

[1] Notice the definition of contraction map can be reformulated in terms of distance alone. Banach's fixed point theorem can also be formulated in the context of a complete metric space. Probably I will assign the formulation and proof in that context as homework. It can be found in *Introduction to Analysis* by Maxwell Rosenlicht, page 170-172, Dover Edition.

[2] Some texts use function only in the one-dimensional context but reserve map for higher dimensional applications. I make no such distinction. Which word I used is just a matter of which one sounds better as I write the sentence

[3] in practice, this must be shown for a proposed contraction map and it may require some work.

[4] yes, this line of the proof is not logically necessary, I'm leaving it here so the induction argument below is easier to follow

We claim $\|x_{n+1} - x_n\| \leq \alpha^n \|x_1 - x_0\|$ for all $n \in \mathbb{N}$. We already proved the claim holds for $n = 1$ and $n = 2$. Suppose the claim holds for some $n \in \mathbb{N}$. Use the definition of the sequence and the definition of contraction mapping to derive:

$$\|x_{n+2} - x_{n+1}\| = \|\varphi(x_{n+1}) - \varphi(x_n)\| \leq \alpha \|x_{n+1} - x_n\| \leq \alpha^{n+1} \|x_1 - x_0\|$$

where we used the induction hypothesis in the last step. Thus the claim is true for $n + 1$ and we have shown $\|x_{n+1} - x_n\| \leq \alpha^n \|x_1 - x_0\|$ for all $n \in \mathbb{N}$ by induction on n. Suppose $n > m$,

$$\begin{aligned}
\|x_n - x_m\| &= \|x_n - x_{n-1} + x_{n-1} - x_{n-2} + \cdots + x_{m+2} - x_{m+1} + x_{m+1} - x_m\| \\
&\leq \|x_n - x_{n-1}\| + \|x_{n-1} - x_{n-2}\| + \cdots + \|x_{m+2} - x_{m+1}\| + \|x_{m+1} - x_m\| \\
&\leq \left(\alpha^{n-1} + \alpha^{n-2} + \cdots + \alpha^{m+1} + \alpha^m\right) \|x_1 - x_0\| \\
&\leq \left(\alpha^m + \alpha^{m+1} + \cdots\right) \|x_1 - x_0\| \\
&= \frac{\alpha^m \|x_1 - x_0\|}{1 - \alpha} \qquad (10.1)
\end{aligned}$$

Let $\varepsilon > 0$. Notice $0 < \alpha < 1$ implies $\lim_{m \to \infty} \alpha^m = 0$ thus there exists $N \in \mathbb{N}$ for which $m > N$ implies $\alpha^m < \frac{(1-\alpha)\varepsilon}{\|x_1 - x_0\|}$. Suppose $m, n > N$ with $n > m$ then

$$\|x_n - x_m\| \leq \frac{\alpha^m \|x_1 - x_0\|}{1 - \alpha} < \frac{\|x_1 - x_0\|}{1 - \alpha} \cdot \frac{(1 - \alpha)\varepsilon}{\|x_1 - x_0\|} = \varepsilon.$$

Therefore $\{x_n\}$ is a Cauchy sequence in M and we find there exists $x_\star \in M$ for which $x_n \to x_\star$ as $n \to \infty$. Notice this is a fixed point for φ since you can show φ is continuous and so

$$\lim_{n \to \infty} \varphi(x_n) = \lim_{n \to \infty} x_{n+1} \quad \Rightarrow \quad \varphi\left(\lim_{n \to \infty} x_n\right) = x_\star \quad \Rightarrow \quad \varphi(x_\star) = x_\star.$$

If $y_\star$ is another fixed point of φ; $\varphi(y_\star) = y_\star$ then we have

$$\|y_\star - x_\star\| = \|\varphi(y_\star) - \varphi(x_\star)\| \leq \alpha \|y_\star - x_\star\|$$

Thus $(1 - \alpha)\|y_\star - x_\star\| \leq 0$ hence $\|y_\star - x_\star\| = 0$ and we find $y_\star = x_\star$. The fixed point is unique.

It remains to prove the estimate on the error of the n-th iterate. Notice $\|\cdot\|$ is a continuous function from V to $[0, \infty)$ and as such we can pass the limit as $n \to \infty$ inside the norm and obtain for fixed m from Equation 10.1

$$\lim_{n \to \infty} \|x_n - x_m\| = \lim_{n \to \infty} \left[\frac{\alpha^m \|x_1 - x_0\|}{1 - \alpha}\right] \quad \Rightarrow \quad \|x_\star - x_m\| \leq \frac{\alpha^m \|x_1 - x_0\|}{1 - \alpha}. \qquad \square$$

10.2 Newton's Method

The genesis of Banach's theorem is almost certainly due to Newton. Newton's Method is just about as old as calculus itself in the modern formulation. The basic problem is to solve $f(x) = 0$ near some given point x_0. Assume $f(x)$ changes sign near x_0 and $f'(x) \neq 0$ near x_0. The method goes like this:

- (Step 0) Check $f(x_0)$, if $f(x_0)$ is sufficiently close to zero then congratulations, you're done. Otherwise proceed,

- (Step 1) Find the x-intercept of $y = f(x_0) + f'(x_0)(x - x_0)$ and call it x_1; that is, define $x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}$. Once more calculate $f(x_1)$ and if it is sufficiently close to zero then you're done. Otherwise, proceed,

- (Step 2) Find the x-intercept of $y = f(x_1) + f'(x_1)(x - x_1)$ and call it x_2; that is, define $x_2 = x_1 - \frac{f(x_1)}{f'(x_1)}$. Once more calculate $f(x_2)$ and if it is sufficiently close to zero then you're done. Otherwise, proceed, ...

- (Step n) Find the x-intercept of $y = f(x_{n-1}) + f'(x_{n-1})(x - x_{n-1})$ and call it x_n; that is, define $x_n = x_{n-1} - \frac{f(x_{n-1})}{f'(x_{n-1})}$. Once more calculate $f(x_n)$ and if it is sufficiently close to zero then you're done. Otherwise, proceed.

I do not intend to prove this converges, indeed, it does not for all examples. Even when $f'(x) \neq 0$ near x_0 it is possible that the method gets stuck in a loop which never zooms into the actual zero for $f(x)$ near x_0. A simplification of Newton's Method which always works is given by adjusting the slope of the linear iterates to be maximal. If we choose the maximum slope possible for $y = f(x)$ near x_0 then we can show that modified Newton's Method will find the root of $f(x) = 0$ near x_0, and thanks to the contraction mapping theorem we'll even be able to give an estimate of the maximum error in each of our estimates. I found this in C.H. Edward's *Advanced Calculus of Several Variables*, Dover Edition, page 164-165.

Theorem 10.2.1. Modified Newton's Method:
Let $f : [a, b] \to \mathbb{R}$ be a differentiable function with $f(a) < 0 < f(b)$ and $0 < m < f'(x) < M$ for each $x \in [a, b]$. Given $x_0 \in [a, b]$ the sequence $\{x_n\}_{n=0}^{\infty}$ defined by

$$x_{n+1} = x_n - \frac{f(x_n)}{M}$$

converges to the unique solution $x_\star \in [a, b]$ of $f(x) = 0$. Moreover,

$$|x_n - x_\star| \leq \frac{|f(x_0)|}{m}\left(1 - \frac{m}{M}\right)^n.$$

Proof: define $\varphi : [a, b] \to \mathbb{R}$ by

$$\varphi(x) = x - \frac{f(x)}{M}$$

Observe φ is differentiable and

$$\varphi'(x) = 1 - \frac{f'(x)}{M}.$$

Observe $0 < m < f'(x) < M$ implies $0 < \frac{m}{M} < \frac{f'(x)}{M} < 1$ hence $0 > -\frac{m}{M} > -\frac{f'(x)}{M} > -1$ which implies $0 < 1 - \frac{f'(x)}{M} < 1 - \frac{m}{M} < 1$. Therefore,

$$0 < \varphi'(x) < 1 - \frac{m}{M} < 1$$

Let $\alpha = 1 - \frac{m}{M}$ and suppose $x, y \in [a, b]$. By the mean value theorem there exists c between x and y for which $\varphi'(c) = \frac{\varphi(x) - \varphi(y)}{x - y}$. Thus,

$$|\varphi(x) - \varphi(y)| = \varphi'(c)|x - y| < \alpha|x - y|.$$

Moreover, $\varphi[a, b] \subseteq [a, b]$ since φ is a strictly increasing function and we can calculate

$$\varphi(a) = a - \frac{f(a)}{M} > a \qquad \& \qquad \varphi(b) = b - \frac{f(b)}{M} < b.$$

We can argue from calculus $\varphi[a,b] = [\varphi(a), \varphi(b)]$. Therefore, we can regard $\varphi : [a,b] \to [a,b]$ as a contraction map with contraction constant $\alpha = 1 - \frac{m}{M}$. Banach's Fixed Point Theorem gives the existence of a unique $x_\star \in [a,b]$ for which $\varphi(x_\star) = x_\star$ which means

$$x_\star = x_\star - \frac{f(x_\star)}{M} \quad \Rightarrow \quad f(x_\star) = 0.$$

Moreover, the estimate of Banach's Theorem gives

$$|x_n - x_\star| \leq \frac{\alpha^n |x_1 - x_0|}{1-\alpha} = \frac{\alpha^n |\varphi(x_0) - x_0|}{1-\alpha} = \frac{\alpha^n}{1-\alpha}\frac{|f(x_0)|}{M} = \frac{|f(x_0)|}{m}\left(1 - \frac{m}{M}\right)^n. \quad \square$$

10.3 inverse function theorem for functions on ℝ

In this section we study the inverse function theorem for functions on ℝ. If the derivative is continuous and nonzero at a given point then we can use the contraction mapping theorem to show there exists and inverse function for an appropriate restriction of the given function to some neighborhood of the given point.

Suppose $f(a) = b$ and $f'(x) \neq 0$ for all x near a then

$$y - b = f(x) - f(a) \approx f'(a)(x-a) \quad \Rightarrow \quad x \approx a - \frac{f(a) - y}{f'(a)}.$$

Define $x_0 = a$ and $x_1 = a - \frac{f(a)-y}{f'(a)}$ and $y_1 = f(x_1)$. Then recursively continue to define:

$$x_{n+1} = x_n - \frac{f(x_n) - y_n}{f'(x_n)}$$

where $y_n = f(x_n)$ for $n \in \mathbb{N}$. Essentially the point is this: to find the inverse function is to solve $y = f(x)$ for x. When the derivative is nonzero we can linearize the function and solve the linearization for x. That solution gives an approximation of the inverse function. Then continue in this fashion and the limit of this iteration constructs the inverse function locally.

The theorem below, due to C.H. Edward's *Advanced Calculus of Several Variables*, Dover Edition, page 165-167, modifies the iteration above slightly much as we modified the standard Newton's Method to a slightly less complicated iteration.

Theorem 10.3.1. Inverse Function Theorem on ℝ:
Let $f : \mathbb{R} \to \mathbb{R}$ be a continuously differentiable function with $f(a) = b$ and $f'(a) \neq 0$. Then there exist neighborhoods $U = [a-\delta, a+\delta]$ and $V = [b-\varepsilon, b+\varepsilon]$ such that given $y_\star \in V$, the sequence defined inductively by

$$x_0 = a, \qquad x_{n+1} = x_n - \frac{f(x_n) - y_\star}{f'(a)}$$

converges to a unique point $x_\star \in U$ such that $f(x_\star) = y_\star$.

Proof: By continuity of the derivative function at $x = a$ we may choose $\delta > 0$ for which $x \in U = [a-\delta, a+\delta]$ implies

$$|f'(x) - f'(a)| < \frac{1}{2}|f'(a)|.$$

Let $\varepsilon = \frac{1}{2}\delta|f'(a)|$. Suppose $y_\star \in V = [b-\varepsilon, b+\varepsilon]$ then we claim that

$$\varphi(x) = x - \frac{f(x) - y_\star}{f'(a)}$$

gives a contraction mapping of U with contraction constant $\alpha = 1/2$. To verify this claim, we begin by examining the derivative of φ,

$$|\varphi'(x)| = \left|1 - \frac{f'(x)}{f'(a)}\right| = \frac{|f'(a) - f'(x)|}{|f'(a)|} \leq \frac{1}{2}.$$

Also, consider by the Mean Value Theorem and our choice of ε,

$$\begin{aligned}
|\varphi(x) - a| &\leq |\varphi(x) - \varphi(a)| + |\varphi(a) - a| \\
&\leq \frac{1}{2}|x-a| + \frac{|f(a) - y_\star|}{|f'(a)|} \\
&\leq \frac{1}{2}\delta + \frac{\varepsilon}{|f'(a)|} \\
&\leq \frac{1}{2}\delta + \frac{\frac{1}{2}\delta|f'(a)|}{|f'(a)|} \\
&= \delta.
\end{aligned}$$

Therefore, $\varphi(x) \in [a-\delta, a+\delta]$. We have shown $\varphi : [a-\delta, a+\delta] \to [a-\delta, a+\delta]$. Let $x, y \in [a-\delta, a+\delta]$ and note by the MVT we have c between x and y for which

$$\frac{\varphi(y) - \varphi(x)}{x-y} = \varphi'(c) \quad \Rightarrow \quad |\varphi(y) - \varphi(x)| = |\varphi'(c)||x-y| \leq \frac{1}{2}|x-y|.$$

Thus φ is a contraction map on $[a-\delta, a+\delta]$ with contraction constant $\alpha = 1/2$. Banach's fixed point theorem gives there exists a unique $x_\star \in [a-\delta, a+\delta]$ for which $\varphi(x_\star) = x_\star$. Thus

$$x_\star = x_\star - \frac{f(x_\star) - y_\star}{f'(a)} \quad \Rightarrow \quad f(x_\star) = y_\star. \quad \square$$

Suppose U and V are as were constructed in the proof above. Given $y \in V$, we define $g(y)$ to be the point $x \in U$ given by the theorem to solve $f(x) = y$. Thus g serves as a local inverse function for f with respect to the restriction to the set U. Moreover, defining

$$g_0(y) = a, \qquad g_{n+1}(y) = g_n(y) - \frac{f(g_n(y)) - y}{f'(a)}$$

gives a sequence of functions which converges to g.

Example 10.3.2. *I'll ignore the details of U and V and simply calculate the sequence suggested above for an example function. Consider $f(x) = 1/x$ and study $f'(1) = -1 \neq 0$. Identify $a = 1$ and $b = 1$ then*

$$g_0(y) = 1$$

$$g_1(y) = 1 - \frac{f(1) - y}{-1} = 1 + 1 - y = -y$$

$$g_2(y) = -y - \left(\frac{1}{-y} - y\right) = \frac{1}{y}$$

$$g_3(y) = \frac{1}{y} - \left(\frac{1}{\frac{1}{y}} - y \right) = \frac{1}{y}$$

Beautiful, $g_n(y) = 1/y$ for all $n \geq 3$. Indeed, the inverse function of $f(x) = 1/x$ is $g(y) = 1/y$. It just so happens our local inverse here serves as a global inverse.

That last example was neat. What happens if we start at a different point ?

Example 10.3.3. *Let $f(x) = 1/x$ and study $a = 1/2$ and $b = 2$ then $f'(x) = -1/x^2$ yieds $f'(1/2) = -4$. Then*

$$g_{n+1}(y) = g_n(y) + 4(f(g_n(y)) - y) = g_n(y) - 4y + \frac{4}{g_n(y)}$$

and $g_0(y) = 1/2$ thus

$$g_1(y) = 1/2 - 4y + \frac{4}{1/2} = \frac{17}{2} - 4y$$

$$g_2(y) = \frac{17}{2} - 4y - 4y + \frac{4}{\frac{17}{2} - 4y} = \frac{128y^2 - 408y + 305}{2(17 - 8y)}$$

Well, I think that is far enough for the last example. It just gets uglier past $n = 2$. My apologies if I made a mistake in there. Let's try another.

Example 10.3.4. *Consider $f(x) = e^x$ and study $a = 0$ and $b = 1$ since $e^0 = 1$. Notice $f'(x) = e^x$ hence $f'(0) = 1$. Begin with $g_0(y) = 0$ then*

$$g_{n+1}(y) = g_n(y) - \frac{f(g_n(y)) - y}{1} = g_n(y) + y - exp((g_n(y))$$

Thus,

$$g_1(y) = 0 + y - exp(0) = y - 1$$

$$g_2(y) = y - 1 + y - exp(y - 1) = 2y - 1 - e^{y-1}$$

$$g_3(y) = 2y - 1 - e^{y-1} + y - exp(2y - 1 - e^{y-1})$$

Yep. It's ugly. The following example is from C.H. Edward's *Advanced Calculus of Several Variables*, Dover Edition, page 168.

Example 10.3.5. *Let $f(x) = x^2 - 1$ and study $a = 1$ with $b = 0$. Notice $f'(x) = 2x$ thus $f'(1) = 2$ hence*

$$g_0(y) = 1, \qquad g_{n+1}(y) = g_n(y) - \frac{1}{2}\left((g_n(y))^2 - 1 - y\right) = g_n(y) + \frac{1}{2}\left(1 + y - (g_n(y))^2\right).$$

Hence,

$$g_1(y) = 1 + \frac{1}{2}\left(1 + y - (1)^2\right) = 1 + \frac{y}{2}$$

$$g_2(y) = 1 + \frac{y}{2} + \frac{1}{2}\left(1 + y - \left(1 + \frac{y}{2}\right)^2\right) = 1 + \frac{y}{2} - \frac{y^2}{8}$$

$$g_3(y) = 1 + \frac{y}{2} - \frac{y^2}{8} + \frac{1}{2}\left(1 + y - \left(1 + \frac{y}{2} - \frac{y^2}{8}\right)^2\right) = 1 + \frac{y}{2} - \frac{y^2}{8} + \frac{y^3}{16} - \frac{y^4}{128}.$$

Notice $y = x^2 - 1$ implies $x^2 = y + 1$ hence $x = \pm\sqrt{1+y}$. Recall $f(x) = (1+x)^\alpha$ has $f(0) = 1$, $f'(0) = \alpha$ and $f''(0) = \alpha(\alpha-1)$ and $f'''(0) = \alpha(\alpha-1)(\alpha-2)$ etc. Thus, by Taylor's Theorem at zero,

$$(1+x)^\alpha = 1 + \alpha x + \frac{1}{2}\alpha(\alpha-1)x^2 + + \frac{1}{3!}\alpha(\alpha-1)(\alpha-2)x^3 + \frac{1}{4!}\alpha(\alpha-1)(\alpha-2)(\alpha-3)x^4 + \cdots$$

Set $\alpha = 1$ and see that:

$$\frac{1}{2}\alpha(\alpha-1) = \frac{-1}{8} \qquad \& \qquad \frac{1}{3!}\alpha(\alpha-1)(\alpha-2) = \frac{1}{6}\cdot\frac{1}{2}\cdot\frac{-1}{2}\cdot\frac{-3}{2} = \frac{1}{16}$$

and

$$\frac{1}{4!}\alpha(\alpha-1)(\alpha-2)(\alpha-3) = \frac{1}{24}\cdot\frac{1}{2}\cdot\frac{-1}{2}\cdot\frac{-3}{2}\cdot\frac{-5}{2} = \frac{-5}{128}.$$

So, $g_3(y)$ almost matches the binomial expansion of $\sqrt{1+y}$. But, as the iteration continues, I believe the order 4 term continues to be modified. Edwards is careful to only include up to order three in his example. This why that is.

Remark 10.3.6. I had originally hoped to add the inverse function theorem for functions of a complex variable. However, I was reminded as I was about to write the section that there is no simple version of the Mean Value Theorem for complex analysis. There is a modification of the MVT for functions of several variables which I intend to prove it in an upcoming section. Once we have that settled, then we'll see how to prove the inverse mapping theorem. The inverse function theorem of complex analysis then appears as a special case of that theorem. All of this said, it might be interesting to attempt a proof of the inverse function theorem for $f : \mathbb{C} \to \mathbb{C}$ directly.

10.4 multivariate mean value theorem

The MVT of single variable calculus says for f which is differentiable on $[a, b]$ there exists $c \in [a, b]$ for which

$$f'(c) = \frac{f(b) - f(a)}{b - a}.$$

In words, the MVT says there is a value of x for which the average rate of change over $[a, b]$ is matched by the instantaneous rate of change $f'(c)$. How should this theorem generalize to functions where the domain and range have more dimensions ? Can it be generalized ?

In order to proceed we need to settle some notation and definitions about differentiation of mappings from $\mathbb{R}^n$ to $\mathbb{R}^m$. I'll roughly follow the notation of C.H. Edward's *Advanced Calculus of Several Variables*, Dover Edition.

Definition 10.4.1. *Suppose $U \subseteq \mathbb{R}^n$ and $F : U \to \mathbb{R}^m$ is a mapping with $p \in U$ then F is* **real differentiable** *at p if there exists a linear map $L : \mathbb{R}^n \to \mathbb{R}^m$ for which*

$$\lim_{h \to 0} \frac{F(p+h) - F(p) - L(h)}{||h||} = 0.$$

In this case we write $d_pF(h) = L(h)$ and call d_pF the **differential of** *F* **at** *p. If d_pF exists for each $p \in U$ then we say F is differentiable on U with differential $dF : U \to \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ defined by $p \mapsto dF(p) = d_pF$ for each $p \in U$.*

If $F : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ then we can write $F = (F_1, \dots, F_m)$ where $F_1, \dots, F_m$ define the **component functions** of F. Each component function is a real-valued function whereas F is generally a vector-valued function. The differential itself is a mapping from a subset of $\mathbb{R}^n$ to a space of linear transformations! The analysis of this will require some effort on our part here, but you'll have to wait a few pages before I tackle the issue.

Theorem 10.4.2. Differentiability of Component Functions:
If $F : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ and $F = (F_1, \dots, F_m)$ then F is differentiable at $p \in U$ if and only if F_j is differentiable at $p \in U$ for each $j = 1, \dots, m$.

Proof: this follows immediately from the vector limit theorem; a limit of a vector valued function exists if and only if each component function limit exists. □

The derivatives discussed thus far are all defined in terms of a multivariate limit which requires multivariate analysis. We can reduce such derivatives to a cloak of partial differentiation which is knit together with appropriate continuity conditions. Partial differentiation reduces multivariate functions to single variable functions constructed by freezing all but one variable.

Definition 10.4.3. *Suppose $U \subseteq \mathbb{R}^n$ and $F : U \to \mathbb{R}^m$ is a mapping with $p \in U$ the* **i-th partial derivative of** *F at p, if it exists, is defined to be*

$$D_iF(p) = \lim_{h \to 0} \frac{F(p + he_i) - F(p)}{h} = \frac{d}{dt}F(p + te_i)\big|_{t=0}$$

If $D_iF(p)$ exists at each $p \in U$ then $D_iF : U \to \mathbb{R}^n$ defines the i-th partial derivative of F.

Essentially the i-th partial derivative captures the change in the map in the x_i-direction by calculating the tangent vector(s) to the path(s)[5] $t \mapsto F(p + te_i)$ at p. Once more, by the vector limit theorem:

Theorem 10.4.4. Partial Differentiability of Component Functions: *If $F : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ and $F = (F_1, \dots, F_m)$ then $D_iF(p)$ exists if and only if $D_iF_j(p)$ exists for each $j = 1, \dots, m$.*

Furthermore, if the partial derivatives of $F = (F_1, \dots, F_m)$ at p exists,

$$D_iF(p) = (D_iF_1(p), D_iF_2(p), \dots, D_iF_m(p)).$$

Example 10.4.5. *Let $\vec{r} : D \subseteq \mathbb{R}^2 \to \mathbb{R}^3$ then using $\vec{r} = (x, y, z)$,*

$$\partial_u\vec{r} = (\partial_u x, \partial_u y, \partial_u z) \quad \& \quad \partial_v\vec{r} = (\partial_v x, \partial_v y, \partial_v z)$$

give the partial velocities of the parametrization $\vec{r}$. In multivariate calculus we find the normal vector field to the surface by calculating $\partial_u\vec{r} \times \partial_v\vec{r}$. Translating into the notation of this work, $\partial_u\vec{r} = D_1\vec{r}$ and $\partial_v\vec{r} = D_2\vec{r}$ since we assume $(u, v) \in D$ which makes u the first coordinate and v the second coordinate.

Theorem 10.4.6. Differentiable implies Partial Differentiable at point: *If $F : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at p then $D_iF(p)$ exists and $dF_p(e_i) = D_iF(p)$ for each $i = 1, \dots, n$.*

[5] the (s) is to deal with $m \geq 2$.

Often the notation $D_iF = \frac{\partial F}{\partial x_i} = \partial_i F$ is used. Notice the theorem above shows that

$$[dF_p] = [dF_p(e_1)|\cdots|dF_p(e_n)] = [D_1F(p)|\cdots|D_nF(p)] = \left[\begin{array}{cccc} \partial_1 F_1 & \partial_2 F_1 & \cdots & \partial_n F_1 \\ \partial_1 F_2 & \partial_2 F_2 & \cdots & \partial_n F_2 \\ \vdots & \vdots & \vdots & \vdots \\ \partial_1 F_m & \partial_2 F_m & \cdots & \partial_n F_m \end{array}\right]$$

Definition 10.4.7. *If $F : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ is a mapping then the* **Jacobian matrix** *of F is*

$$J_F = \left[\frac{\partial F}{\partial x_1}\bigg|\frac{\partial F}{\partial x_2}\bigg|\cdots\bigg|\frac{\partial F}{\partial x_n}\right] = \left[\begin{array}{cccc} \partial_1 F_1 & \partial_2 F_1 & \cdots & \partial_n F_1 \\ \partial_1 F_2 & \partial_2 F_2 & \cdots & \partial_n F_2 \\ \vdots & \vdots & \vdots & \vdots \\ \partial_1 F_m & \partial_2 F_m & \cdots & \partial_n F_m \end{array}\right] = \left[\begin{array}{c} \nabla F_1^T \\ \nabla F_2^T \\ \vdots \\ \nabla F_m^T \end{array}\right].$$

Where $\nabla F_j = (row_j(J_F))^T$. We also call ∇F_j the **gradient** *of F_j.*

We can either look at the Jacobian matrix as partial derivatives of the map concatenated column-by-column, or as a stack of gradients, one for each component function of the map. Notice $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$ has a Jacobian matrix which is a $1 \times n$ matrix, in particular

$$J_f = [\partial_1 f, \dots, \partial_n f] = (\nabla f)^T.$$

Now that we have set all the notation we are ready to look at a case where a nice generalization of the Mean Value Theorem (MVT) can be made.

Theorem 10.4.8. MVT for real-valued function of n-variables:
If $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$ is differentiable on U and $[a, b] = \{a + t(b - a) \mid 0 \leq t \leq 1\} \subseteq U$ then there exists $c \in [a, b]$ for which

$$f(b) - f(a) = (\nabla f)(c) \bullet (b - a)$$

Proof: Let $\gamma(t) = a + t(b - a)$ define the path from $\gamma(0) = a$ to $\gamma(1) = b$ for $0 \leq t \leq 1$. Define $h : [0, 1] \to \mathbb{R}$ by $h(t) = f(\gamma(t))$ and calculate by the chain-rule that:

$$\frac{dh}{dt} = \frac{d}{dt} f \circ \gamma = \nabla f(\gamma(t) \bullet \frac{d\gamma}{dt}$$

Since f is differentiable it is continuous[6] on U and thus $h = f \circ \gamma$ is the composition of continuous functions which is once more continuous. Likewise, h is differentiable on $[0, 1]$ thus we may apply the Mean Value Theorem from first semester Calculus and find there exists $\mathfrak{c} \in [0, 1]$ for which $h'(\mathfrak{c}) = \frac{h(1)-h(0)}{1-0} = h(1) - h(0)$. But, $h(0) = f(\gamma(0)) = f(a)$ and $h(1) = f(\gamma(1)) = f(b)$ hence

$$h'(\mathfrak{c}) = f(b) - f(a) = \nabla f(\gamma(\mathfrak{c}) \bullet \frac{d\gamma}{dt}.$$

We calculate $\frac{d\gamma}{dt} = b - a$ and set $c = \gamma(\mathfrak{c}) \in [a, b]$ to conclude our proof:

$$f(b) - f(a) = \nabla f(c) \bullet (b - a). \qquad \square$$

This result does not generalize directly to mappings from $\mathbb{R}^n$ to $\mathbb{R}^m$ with $m \geq 2$. It is not generally the case that $F(b) - F(a) = J_F(c)(b - a)$ for some $c \in [a, b]$. It is perhaps helpful to study how this breaks down in the context of complex analysis.

[6] perhaps this would be a good homework problem

Example 10.4.9. *Let $f = u + iv$ define a function on $\mathbb{C}$. Then $u, v : \mathbb{R}^2 \to \mathbb{R}$ are the* **component functions** *of f. Let $[a, b] \subseteq \mathbb{C}$ denote the line-segment from a to b then we can apply Theorem 10.4.8 and select $c_1, c_2 \in [a, b]$ for which*

$$u(b) - u(a) = (\nabla u)(c_1) \bullet (b - a) \;\; \& \;\; v(b) - v(a) = (\nabla v)(c_2) \bullet (b - a)$$

If $c_1 = c_2 = c$ then since $J_f = \begin{bmatrix} u_x & u_y \\ v_x & v_y \end{bmatrix}$ we would have

$$J_f(c)(b - a) = \begin{bmatrix} (\nabla u)(c) \bullet (b - a) \\ (\nabla v)(c) \bullet (b - a) \end{bmatrix} = \begin{bmatrix} u(b) - u(a) \\ v(b) - v(a) \end{bmatrix} = f(b) - f(a).$$

However, logically, there is no reason in general that c_1, c_2 ought to be identical, or even close.

Remark 10.4.10. In the case $[a, b] = [\gamma(t_j), \gamma(t_{j+1})]$ where $t_{j+1} - t_j = \triangle t$ if we pick $c_1, c_2 \in [a, b]$ in this circumstance, and then $\triangle t \to 0$ then in that limiting circumstance $c_1 \to c_2$. This calculation arises in Complex Analysis as we study the basic properties of the complex integral and relate a Riemann-type formulation to a parametric formulation for the complex integral.

Corollary 10.4.11. MVT-type estimate for real-valued function of n-variables: *If $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$ is differentiable on U and $[a, b] = \{a + t(b - a) \mid 0 \leq t \leq 1\} \subseteq U$ then*

$$|f(b) - f(a)| \leq \max\{ \; \|\nabla f(c)\| \mid c \in [a, b] \; \}\|b - a\| = \max_{x \in [a,b]} (\|\nabla f(x)\|)\|b - a\|.$$

Proof: we know $f(b) - f(a) = (\nabla f)(c) \bullet (b - a)$ for some $c \in [a, b]$ by Theorem 10.4.8. The Cauchy Schwarz inequality gives

$$|(\nabla f)(c) \bullet (b - a)| \leq \|(\nabla f)(c)\|\|b - a\|$$

and the Corollary follows immediately. □

The result above generalizes to the context of mappings. However, we first need to develop several results concerning the norm of transformation.

Definition 10.4.12. *Suppose V, W are normed linear spaces and $T : V \to W$ is a linear transformation then the* **norm** *of T is defined by*

$$\|T\| = sup\{\|T(x)\| \mid x \in V, \|x\| = 1\}$$

If the supremum above does not exist in $\mathbb{R}$ then we write $\|T\| = \infty$ and call T an **unbounded linear operator***.*

If $\|T\| < \infty$ then we have the following estimate:

Proposition 10.4.13. *Let $T \in \mathcal{L}(V, W)$ where V, W are normed linear spaces over $\mathbb{R}$. If $\|T\| < \infty$ then $\|T(x)\| \leq \|T\|\|x\|$ for each $x \in V$.*

Proof: clearly if $x = 0$ the claim holds. Suppose $x \neq 0$, notice

$$\|T(x)\| = \left\| T\left(\frac{\|x\|}{\|x\|}x\right) \right\| = \left\| \|x\| T\left(\frac{x}{\|x\|}\right) \right\| = \|x\| \left\| T\left(\frac{x}{\|x\|}\right) \right\| \leq \|x\|\|T\|$$

since $\|x/\|x\|\| = 1$ so $\|T\|$ provides the indicated bound. □

At this point I will defer to Edward's text since he has detailed calculations in a particular direction. If I had more time, it would be nice to do things differently, but I must not let perfect be the enemy of good. Edward's makes strategic use of the *sup*-norm and the *one*-norm. In fact, his uses a mixture of both to set up some rather nice identities. Let us go through these since I intend to follow his arguments here and also in the technical sections which follow on the inverse mapping and implicit mapping theorems.

Definition 10.4.14. *Let $x \in \mathbb{R}^n$ then define $|x|_0 = \max\{|x_1|, \dots, |x_n|\}$ as the* **max norm** *of $\mathbb{R}^n$. The n-***cube of radius** *$r > 0$ is defined by $C_r^n = \{x \in \mathbb{R}^n \mid |x|_0 \leq 1\}$.*

It might be better to use $|x|_\infty$ since if we use $|x|_p = \sqrt[p]{\sum_{i=1}^p |x_i|^p}$ then essentially as $p \to \infty$ we obtain the max norm. Notice, $|x|_2$ is the usual *Euclidean norm* whereas $|x|_1$ is the so-called *taxi-cab* or *one* norm. Notice the analog of the n-cube of radius r is the usual n-ball of radius r for the Euclidean norm. Cubes can be expressed as Cartesian products of the interval $[-r, r]$:

$$C_r^2 = [-r,r] \times [-r,r], \quad C_r^3 = [-r,r] \times [-r,r] \times [-r,r], \quad C_r^n = [-r,r]^n$$

This makes the max norm especially easy to think about as we examine questions of analysis that involve coordinate-by-coordinate estimation. The boundaries of these cubes are also especially nice:

$$\partial C_r^2 = (\{-r\} \times [-r,r]) \cup (\{r\} \times [-r,r]) \cup ([-r,r] \times \{-r\}) \cup ([-r,r] \times \{r\})$$

which is otherwise known as the square with vertices $(-r,-r), (-r,r), (r,-r), (r,r)$. The larger take away from these musings is that for $x \in \partial C_r^n$ there exists at least one coordinate x_j for which $|x_j| = r$. Consequently,

Proposition 10.4.15. *If $x \in \partial C_r^n$ then $|x|_0 = r$.*

Next we define the norm of a linear transformation with respect to the max norm:

Definition 10.4.16. *Let $L : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation then $\|L\|_0 = \max_{x \in \partial C_1^n} (|L(x)|_0)$.*

Since the mapping $x \mapsto |L(x)|_0$ is formed by the composition of continuous maps[7] it follows from the Extreme Value Theorem that the maximum value $|L(x)|_0$ exists since ∂C_1^n is a compact set. Thus $\|L\|$ is defined for any $L \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$. Following the argument for Proposition 10.4.13 we see $|L(x)|_0 \leq \|L\|_0 |x|_0$ for any $x \in \mathbb{R}^n$. In this context, we can say more:

Proposition 10.4.17. *If M is a constant such that $|L(x)|_0 \leq M|x|_0$ for all $x \in \mathbb{R}^n$ then $\|L\| \leq M$.*

Proof: if $x \in \partial C_1^n$ then $|x|_0 = 1$ thus $|L(x)|_0 \leq M$. However, by definition of $\|L\|_0$ we also have $\|L\|_0 \leq |L(x)|_0$ hence $\|L\|_0 \leq |L(x)|_0 \leq M$ and we conclude $\|L\| \leq M$. □

C.H. Edward now gives Lemma 2.2 on page 174 which show the norm of a component function of a linear transformation is less than the norm of the whole transformation.

[7] my apologies, there is some gap here since this is not a complete course with the whole treatment of continuous maps, sometimes I go over the rudiments of such theory in Math 332.

Proposition 10.4.18. *If $L : \mathbb{R}^n \to \mathbb{R}^m$ has $L = (L_1, \dots, L_m)$ then $\|L_j\|_0 \leq \|L\|_0$ for $1 \leq j \leq m$.*

Proof: suppose $1 \leq j \leq m$ and let $x_0 \in \partial C_1^n$ be the point for which $\|L_j\|_0 = |L_j(x_0)|_0$. Then,

$$\|L_j\|_0 = |L_j(x_0)|_0 \leq \max(|L_1(x_0)|_0, \dots, |L_m(x_0)|_0) = |L(x_0)|_0 \leq \max_{x \in \partial C_1^n}(|L(x)|_0) = \|L\|_0. \ \square$$

Suppose $L : \mathbb{R}^n \to \mathbb{R}^m$ has standard matrix A; $L(x) = Ax$ for each $x \in \mathbb{R}^n$. Notice

$$L(x) = \left[\begin{array}{c} row_1(A) \bullet x \\ row_2(A) \bullet x \\ \vdots \\ row_m(A) \bullet x \end{array} \right]$$

If we take $x \in \partial C_1^n$ then $|row_i(A) \bullet x|$ will be largest if we take $x = e_k$ where A_{ik} is the largest magnitude entry in $row_i(A)$. Then, if we repeat this for each row then we will maximize the possible value of $|L(x)|_0$ as x ranges over $x \in \partial C_1^n$. For this reason we give the following[8] definition for the norm of a matrix:

Definition 10.4.19. *Let $A \in \mathbb{R}^{m \times n}$ then define*

$$\|A\|_0 = \max\{|row_1(A)|_1, \dots, |row_m(A)|_1\}$$

where $|[A_{i1}, \dots, A_{in}]|_1 = |A_{i1}| + \cdots + |A_{in}|$ for each $row_i(A) = [A_{i1}, \dots, A_{in}]$ for $1 \leq i \leq m$. Equivalently,

$$\|A\|_0 = \max_{1 \leq i \leq m} \sum_{j=1}^{n} |A_{ij}|.$$

Proposition 10.4.20. *If $L : \mathbb{R}^n \to \mathbb{R}^m$ has $L(x) = Ax$ for all $x \in x \in \mathbb{R}^n$ then $\|L\|_0 = \|A\|_0$.*

Proof: let $x \in \mathbb{R}^n$ and calculate,

$$\begin{aligned} |L(x)|_0 &= \max_{1 \leq i \leq m} \left| \sum_{j=1}^{n} A_{ij} x_j \right| \\ &\leq \max_{1 \leq i \leq m} \sum_{j=1}^{n} |A_{ij} x_j| \\ &\leq |x|_0 \max_{1 \leq i \leq m} \sum_{j=1}^{n} |A_{ij}| \\ &= |x|_0 \|A\|_0. \end{aligned}$$

Thus $\|L\|_0 = sup\{|L(x)|_0 \mid |x|_0 = 1\} \leq \|A\|_0$.

Note, by Proposition 10.4.18, $\|L_i\|_0 \leq \|L\|_0$ for $1 \leq i \leq m$. Thus $\|L_i\|_0 \leq \|A\|_0$ for each i. If we can find $z \in \partial C_1^n$ for which $|L_i(z)|_0 = \|A\|_0$ then it follows $\|L_i\|_0 = \|A\|_0$. We argue this is possible for the special choice of k for which $|row_k(A)|_1 \geq |row_i(A)|$ for $1 \leq i \leq m$. Observe

[8] in my view, rather odd, but actually quite clever if you see how it works together with $\|L\|_0$ already defined.

$\|A\|_0 = |row_k(A)|_1$. Define $sign(r) = 1$ if $r \geq 0$ and $sign(r) = -1$ if $r < 0$. Let $z_j = sign(A_{kj})$ for $j = 1, \dots, n$ then z is a column vector for which every entry is ± 1. Thus $z \in \partial C_1^n$ and

$$(L(z))_k = (Az)_k = row_k(A) \bullet z = \sum_{j=1}^{n} sign(A_{kj})A_{kj} = \sum_{j=1}^{n} |A_{jk}| = |row_k(A)|_1 = \|A\|_0.$$

Therefore, $\|L_k\|_0 = \|A\|_0$. However, $\|L_k\|_0 \leq \|L\|_0$ thus $\|A\|_0 \leq \|L\|_0$. Hence $\|L\|_0 = \|A\|_0$. □

With the results above settled we can easily prove continuity of a linear transformation.

Proposition 10.4.21. *If $L : \mathbb{R}^n \to \mathbb{R}^m$ has $L(x) = Ax$ for all $x \in x \in \mathbb{R}^n$ then L is continuous.*

Proof: suppose $x \in \mathbb{R}^n$ then if $x \neq 0$ we can calculate $x = |x|_0 \hat{x}$ where $\hat{x} = \frac{x}{|x|_0}$ has $|\hat{x}|_0 = 1$ thus for

$$|L(x)|_0 = L\left(|x|_0 \hat{x}\right) = |x|_0 L\left(\hat{x}\right) \leq |x|_0 \|A\|_0.$$

Therefore, for $x \neq a$,

$$\|L(x) - L(a)\|_0 = \|L(x-a)\| \leq |x-a|_0 \|A\|_0$$

Let $a \in \mathbb{R}^n$ and assume $\|A\|_0 \neq$. Let $\varepsilon > 0$ and choose $\delta = \frac{\varepsilon}{\|A\|_0}$ if $0 < |x-a|_0 < \delta$ then

$$\|L(x) - L(a)\|_0 \leq |x-a|_0 \|A\|_0 \leq \frac{\varepsilon}{\|A\|_0} \|A\|_0 = \varepsilon.$$

Furthermore, if $L(x) = 0$ for all $x \in \mathbb{R}^n$ then it is easy to argue continuity of L at $a \in \mathbb{R}^n$. □

Remark 10.4.22. There is a different proof of continuity of a linear transformation I typically offer in Math 332 which does not require the special definitions given in this article. Basically, if you let $M = \max |A_{ij}|$ then it is easy to bound the values of Ax and as such continuity can be easily argued. Alternatively, the component functions of $L(x) = Ax$ are linear combinations of the coordinate functions thus are sums and scalar multiples of continuous functions.

Definition 10.4.23. *Let $F : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ be differentiable on U then we say F is* **continuously differentiable on** *U if each partial derivative function of F is continuous on U. In this case we say $F \in C^1 U$.*

Since J_F has component functions which are partial derivatives of components of F it is also true to say $F \in C^1 U$ if and only if $p \mapsto J_F(p)$ is continuous for $p \in U$.

The differential for $F : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ is the assignment of a linear transformation d_pF for each $p \in U$. To say the differential is continuous on U means for each $\varepsilon > 0$ we need to be able to supply $\delta > 0$ for which $0 < |x-p|_0 < \delta$ implies $\|d_xF - d_pF\|_0 < \varepsilon$. Notice $\|d_pF\|_0 = \|J_F\|_0$ since the Jacobian matrix is the standard matrix of the differential. Consequently,

$$\|d_xF - d_pF\|_0 = \|J_F(x) - J_F(p)\|_0.$$

Therefore, if F is differentiable on U then we can think of the map being **continuously differentiable** in three equivalent manners:

(i.) dF is continuous on U

(ii.) J_F is continuous on U

(iii.) $\frac{\partial F_i}{\partial x_j}$ is continuous on U for each $1 \leq i \leq m$, $1 \leq j \leq n$.

Often in undergraduate texts we define continuous differentiability of a map in terms of the continuity of the partial derivatives. Thus, as a default, I take (iii.) to define continuous differentiability of a map. I will not prove the following theorem, it is non-trivial:

Theorem 10.4.24. *If $F : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ is continuously differentiable, in the sense that all the partial derivative functions of F are continuous on U then F is differentiable on U.*

Proof: see Theorem 2.5 on page 72-73 of C.H. Edward's *Advanced Calculus of Several Variables*, Dover Edition. I usually prove this in Math 332. □

Finally, we reach the main goal, the generalization of the mean value theorem for maps:

Theorem 10.4.25. Mean Value Estimation Theorem for Maps:
If $F : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ is continuously differentiable on the open set U and $[a, b] \subseteq U$ then

$$|F(b) - F(a)|_0 \leq |b - a|_0 \max_{x \in [a,b]} \left(\|J_F(x)\|_0 \right).$$

Proof: suppose $\gamma : [0, 1] \to \mathbb{R}^m$ is defined by $\gamma(t) = F(a + t(b - a))$. Let $h = b - a$ for convenience of exposition; $\gamma(t) = F(a + th)$. Notice $\gamma(0) = F(a)$ and $\gamma(1) = F(b)$. The chain-rule gives that

$$\frac{d\gamma}{dt} = dF_{a+th}(h) \quad \Rightarrow \quad \frac{d\gamma_k}{dt} = d(F_k)_{a+th}(h).$$

Following Edward's proof on page 177 roughly, assuming the k-th coordinate is maximal,

$$\begin{aligned}
|F(b) - F(a)|_0 &= |F_k(b) - F_k(a)| \\
&= |\gamma_k(1) - \gamma_k(0)| \\
&= \left| \int_0^1 \frac{d\gamma_k}{dt} dt \right| \\
&\leq \int_0^1 \left| \frac{d\gamma_k}{dt} \right| dt \\
&\leq \int_0^1 |d(F_k)_{a+th}(h)|\, dt \\
&\leq \max_{t \in [0,1]} |d(F_k)_{a+th}(h)| \\
&\leq \max_{t \in [0,1]} \left(|h|_0 \|d(F_k)_{a+th}\|_0 \right) \\
&\leq |h|_0 \max_{t \in [0,1]} \left(\|d(F_k)_{a+th}\|_0 \right) \\
&\leq |h|_0 \|d(F_k)_{a+\tau h}\|_0 \qquad \text{(extreme value theorem gives such } \tau \in [0,1]) \\
&\leq |h|_0 \|dF_{a+\tau h}\|_0 \qquad \text{(applying Proposition 10.4.18)} \\
&\leq |h|_0 \max_{t \in [0,1]} \|dF_{a+th}\|_0 \\
&\leq |h|_0 \max_{t \in [0,1]} \|J_F(a + th)\|_0 \\
&\leq |h|_0 \max_{x \in [a,b]} \|J_F(x)\|_0. \quad \square
\end{aligned}$$

The theorem above is central to the more sophisticated estimates which underly both the inverse mapping theorem and the implicit mapping theorem.

Corollary 10.4.26. *If $F : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ is continuously differentiable on the open set U and $[a, a+h] \subseteq U$ and $L : \mathbb{R}^n \to \mathbb{R}^m$ is a linear mapping then*

$$|F(a+h) - F(a) - L(h)|_0 \leq |h|_0 \max_{x \in [a,a+h]} (\|dF_x - L\|_0) .$$

Proof: Let $G(x) = F(x) - L(x)$ and note $dG = dF - dL$ and a homework exercise shows $dL = L$ thus $dG = dF - L$. Setting $b = a + h$ we have $h = b - a$ and

$$G(b) - G(a) = F(b) - L(b) - F(a) + L(a) = F(a+h) - F(a) - L(h).$$

Moreover, $\|J_G(x)\|_0 = \|dG_x\|_0 = \|dF_x - L\|_0$ hence the Corollary follows from the theorem. □

Corollary 10.4.27. Motion of Cube by Map with Invertible Differential: *Suppose $F : U \subseteq \mathbb{R}^n \to \mathbb{R}^n$ is continuously differentiable mapping on the open set U which contains the cube C_r and $F(0) = 0$ and $dF_0 = I$ where I denotes the identity map on $\mathbb{R}^n$. If $\|dF_x - I\|_0 < \varepsilon$ for each $x \in C_r$ then $f(C_r) \subseteq C_{(1+\varepsilon)r}$.*

Proof: apply the previous corollary with $L = I$ and $a = 0$ and $h = x \in C_r$ hence

$$|F(x) - x|_0 \leq |h|_0 \max_{z \in [0,x]} (\|dF_x - I\|_0) \leq \varepsilon |x|_0.$$

However, using the triangle inequality,

$$|\ |F(x)|_0 - |x|_0\ | \leq |F(x) - x|_0 \leq \varepsilon |x|_0.$$

Thus,

$$-\varepsilon|x|_0 \leq |F(x)|_0 - |x|_0 \leq \varepsilon|x|_0 \quad \Rightarrow \quad (1-\varepsilon)|x|_0 \leq |F(x)|_0 \leq (1+\varepsilon)|x|_0 \leq (1+\varepsilon)r.$$

Therefore, $F(x) \in C_{(1+\varepsilon)r}$. □

To say a mapping is **continuously differentiable at a point** means there exists some open set containing the point on which the mapping is continuously differentiable.

Corollary 10.4.28. Local Injectivity at point for Map gives injectivity nearby: *Suppose $F : U \subseteq \mathbb{R}^n \to \mathbb{R}^n$ is continuously differentiable at a. If dF_a is injective then there exists a neighborhood of a on which F is injective.*

Proof: suppose dF_a is injective then $dF_a \neq 0$ hence there exists a minimum value $m > 0$ which is attained by $|dF_a(x)|_0$ for $x \in \partial C_1^n$. Let $0 < \varepsilon < m$. As F is continuously differentiable at a, there exists $\delta > 0$ for which $|x - a|_0 < \delta$ implies $\|dF_x - dF_a\|_0 < \varepsilon$. Let $U = \{x \in \mathbb{R}^n \mid |x - a|_0 < \delta\}$ and choose a distinct pair $x, y \in U$. Then, apply Corollary 10.4.26 with $L = dF_a$ and $[x, y] = [a, a+h]$ so $h = y - x$ and

$$|F(y) - F(x) - dF_a(y-x)|_0 \leq |y - x|_0 \max_{z \in [x,y]} (\|dF_z - dF_a\|_0) < \varepsilon |y - x|_0.$$

Then, by the triangle inequality,

$$||dF_a(y-x)|_0 - |F(y) - F(x)|_0| < \varepsilon |y - x|_0$$

unwraping the abolute value inequality gives:

$$-\varepsilon|y - x|_0 < |dF_a(y-x)|_0 - |F(y) - F(x)|_0 < \varepsilon|y - x|_0$$

Hence

$$|F(y) - F(x)|_0 > |dF_a(y-x)|_0 - \varepsilon|y - x|_0 \geq (m - \varepsilon)|y - x|_0 > 0$$

Consequently, $|F(y) - F(x)|_0 \neq 0$ thus $F(y) \neq F(x)$. Therefore, F is injective on the neighborhood U which contains a. □

10.5 inverse mapping theorem

Given a mapping $F : \mathbb{R}^n \to \mathbb{R}^n$ when is it possible to find a *local inverse* to F. That is, given $p \in \mathbb{R}^n$ when is it possible to find a set $S \subseteq \mathbb{R}^n$ containing p for which $F_S : S \to F(S)$ is invertible ? Since F_S invertible means that F_S must be injective. If both F_S and its inverse F_S^{-1} are differentiable then the chain-rule for mappings applied to $F_S^{-1} \circ F_S = Id_S$ yields

$$dF_S^{-1} \circ dF_S = d(Id_S) = Id_{\mathbb{R}^n}$$

Or, in terms of Jacobian matrices, and I'll put in the point dependence, $J_{F_S^{-1}}(F(p))J_{F_S}(p) = I$ for each $p \in S$. Therefore, if we wish to invert a differentiable map near p then it is desireable to have an invertible differential near p.

Example 10.5.1. *Suppose $f : \mathbb{R} \to \mathbb{R}$ is defined by $f(x) = x^3$ and we wish to find a local inverse at $p = 0$. We can solve $y = x^3$ for $x = \sqrt[3]{y}$ and deduce $f^{-1}(y) = \sqrt[3]{y}$ for $y \in \mathbb{R}$. Apparently, f has a* **global inverse***. However, $f'(x) = 3x^2$ hence $J_f(x) = 3x^2$ and $J_f(0) = 0$ is certainly not invertible. Your homework problem is to explain why this example does not contradict the previous paragraph. What gives ?*

Example 10.5.2. *Another important example is given by complex analysis. Define $F : \mathbb{R}^2 \to \mathbb{R}^2$ by $F(x, y) = (x^2 - y^2, 2xy)$. In complex notation, $F(z) = z^2$ since $z = x + iy$ and $(x^2 - y^2, 2xy) = x^2 - y^2 + 2xyi = (x + iy)^2 = z^2$. Here we use the notation $z = (x, y) = x + iy$. If we define $S = \{(x, y) \in \mathbb{R}^2 \mid (x, y) \neq (0, 0)$ then note*

$$J_F = \left[\begin{array}{cc} 2x & -2y \\ 2y & 2x \end{array}\right] \quad \Rightarrow \quad \det(J_F) = 4x^2 + 4y^2 = 4|z|^2$$

Thus dF_p is invertible for each $p \in S$. However, we cannot find an inverse for F_S since F is not one-to-one on S. Let's see what happens when we try to solve $w = z^2$ for z. Let $w = \rho e^{i\beta}$ and $z = re^{i\theta}$ then

$$\rho e^{i\beta} = \left(re^{i\theta}\right)^2 = r^2 e^{2i\theta} \quad \Rightarrow \quad \rho = r^2 \;\;\&\;\; \beta + 2\pi k = 2\theta$$

for some $k \in \mathbb{Z}$. Thus,

$$r = \sqrt{\rho} \qquad \& \qquad \theta = \frac{\beta}{2} + k\pi$$

where $k \in \mathbb{Z}$. Notice, $k = 0, 1$ are sufficient to cover all geometrically distinct cases. Observe,

$$F(\sqrt{\rho}e^{i\beta/2}) = \rho e^{i\beta} \qquad \& \qquad F(\sqrt{\rho}e^{i(\beta/2+\pi)}) = \rho e^{i\beta}$$

Yet, $\sqrt{\rho}e^{i\beta/2} \neq \sqrt{\rho}e^{i(\beta/2+\pi)}$. At this point, I have achieved my goal, namely to confuse the students. I could have made this way easier. Notice $(\pm z)^2 = z^2$ thus F_S is not one-to-one so long as we include antipodal points. We can construct an inverse function on any half-plane with the origin removed. For instance, if $S = \{(x, y) \mid x > 0\}$ then $F(S) = \mathbb{C}^- = \mathbb{R}^2 - \{(0, y) \mid y \leq 0\}$ and for $(x, y) \in \mathbb{C}^-$,

$$F_S^{-1}(x, y) = \sqrt{\sqrt{x^2 + y^2}exp\,(i\theta/2)}$$

where $\theta \in (-\pi, \pi]$ solves $x = \sqrt{x^2 + y^2}\cos\theta$ and $y = \sqrt{x^2 + y^2}\sin\theta$. That, is

$$F_S^{-1}(z) = \sqrt{|z|}\exp\left(\frac{i}{2}Arg(z)\right)$$

where $Arg(z) = \theta$ and $|z| = \sqrt{x^2 + y^2}$. I will likely give a homework based on modifying this example.

Very well, with these cautionary examples in mind, we now present the main theorem. This theorem explains when we can find a *differentiable* local inverse at p for a continuously differentiable mapping on $\mathbb{R}^n$. This is based on Theorem 3.3 in Edwards, page 185.

Theorem 10.5.3. inverse mapping theorem for maps on $\mathbb{R}^n$:

Suppose the mapping $F : \mathbb{R}^n \to \mathbb{R}^n$ is continuously differentiable in a neighborhood W of a point p and suppose $det(J_F(p)) \neq 0$. Then F is locally invertible at p. In particular, there exist open sets $U \subset W$ with $p \in U$ and $V = F(U)$ with $F(p) = q \in V$ and an injective mapping $G : V \to U$ such that

$$G(F(x)) = x \qquad F(G(y)) = y$$

for each $x \in U$ and each $y \in V$. Moreover, the local inverse G is the limit of the sequence $\{G_k\}_0^\infty$ of successive approximations defined inductively by

$$G_0(y) = p, \quad G_{k+1}(y) = G_k(y) - J_F(p)^{-1}\left[F(G_k(y)) - y\right]$$

for $y \in V$.

Proof: we begin by stating and proving a lemma.

Lemma 10.5.4. inverse mapping theorem at zero:

Suppose the mapping $F : \mathbb{R}^n \to \mathbb{R}^n$ is continuously differentiable in a neighborhood W of the origin and suppose $dF_0 = I$. Further, suppose $\|dF_x - I\|_0 \leq \varepsilon < 1$ for all $x \in C_r$. Then

$$C_{(1-\varepsilon)r} \subseteq F(C_r) \subseteq C_{(1+\varepsilon)r}.$$

If $V = int\left(C_{(1-\varepsilon)r}\right)$ and $U = int\left(C_r \cap F^{-1}(V)\right)$, then $F : U \to V$ is an bijection and the inverse mapping $G : V \to U$ is differentiable at zero. Moreover, G is the limit of the sequence $\{G_m\}_0^\infty$ of successive approximations defined inductively by

$$G_0(y) = p, \quad G_{m+1}(y) = G_m(y) - F(G_m(y)) + y$$

for $y \in V$.

Proof: apply Corollary 10.4.26 with $L = Id$ to see for $x, y \in C_r$ we have

$$|F(x) - F(y) - (x - y)|_0 \leq |x - y|_0 \max_{z \in [x,y]} \|dF_x - I\|_0 \leq \varepsilon |x - y|_0 \tag{10.2}$$

Therefore,

$$(1 - \varepsilon)|x - y|_0 \leq |F(x) - F(y)|_0 \leq (1 + \varepsilon)|x - y|_0. \tag{10.3}$$

Notice the left inequality implies F is injective on C_r. Also, notice the right inequality shows $F(C_r) \subseteq C_{(1+\varepsilon)r}$. It remains to show $C_{(1-\varepsilon)r} \subseteq F(C_r)$.

Let $y \in C_{(1-\varepsilon)r}$. Define $\varphi : \mathbb{R}^n \to \mathbb{R}^n$ by

$$\varphi(x) = x - F(x) - y.$$

We wish to show φ defines a contraction mapping of C_r with fixedpoint x for which $F(x) = y$. Let $x \in C_r$ and consider

$$\begin{aligned}
|\varphi(x)|_0 &= |x - F(x) - y|_0 \\
&\leq |F(x) - x|_0 + |y|_0 \\
&\leq |F(x) - F(0) - dF_0(x - 0)|_0 + |y|_0 \ \ \text{(using } F(0) = 0 \text{ and } dF_0 = I) \\
&\leq |y|_0 + |x|_o \max_{x \in C_r} \|dF_x - dF_0\|_0 \ \ \text{(using Corollary 10.4.26)} \\
&\leq (1 - \varepsilon)r + r\varepsilon \\
&= r.
\end{aligned} \tag{10.4}$$

The equation above shows that if $x \in C_r$ then $|\varphi(x)|_0 < r$ thus $\varphi(C_r) \subseteq int(C_r)$ provided we construct φ using $y \in V = int\big(C_{(1-\varepsilon)r}\big)$. Therefore, $\varphi(C_r) \subseteq C_r$ hence we can study φ as a potential contraction mapping on C_r. If $x, y \in C_r$ then

$$|\varphi(x) - \varphi(y)|_0 = |F(x) - F(y) - (x-y)|_0 \leq \varepsilon|x-y|_0$$

thus φ is a contraction mapping on C_r with contraction constant ε. By Banach's fixed point theorem, there exists $x \in C_r$ for which $\varphi(x) = x$ hence $x - F(x) - y = x$ and so $F(x) = y$.

Define $U = F^{-1}(V) \cap int(C_r)$ and note U and $V = int\big(C_{(1-\varepsilon)r}\big)$ are open neighborhoods of zero such that F is a bijection from U to V. For each $y \in V$ we define x uniquely by $F(x) = y$. Denote the inverse mapping by $G(y) = x$ and note we obtain such x as the limit of a sequence defined by:

$$x_0 = 0, \quad x_{m+1} = \varphi(x_m) = x_m - F(x_m) + y$$

by the contraction mapping theorem.

Finally, we show G is differentiable at 0. Recall Equation 10.2

$$|F(x) - F(y) - (x-y)|_0 \leq |x-y|_0 \leq \varepsilon|x-y|_0$$

Let $y = 0$ and $x = G(h)$ and $h = F(x)$ to obtain

$$|G(h) - h|_0 = |x - F(x)|_0 = |F(x) - x|_0 \leq \varepsilon|x|_0$$

Next, apply the left inequality of Equation 10.3 to see $|x|_0 \leq \frac{1}{1-\varepsilon}|F(x)|_0$ thus

$$|G(h) - h|_0 \leq \varepsilon|x|_0 \leq \frac{\varepsilon}{1-\varepsilon}|F(x)|_0 = \frac{\varepsilon}{1-\varepsilon}|h|_0.$$

Consequently,

$$\frac{|G(h)-h|_0}{|h|_0} \leq \frac{\varepsilon}{1-\varepsilon}$$

Notice that F is assumed to be continuously differentiable at zero with $dF_0 = I$ thus we may force $\|dF_x - I\|_0 < \varepsilon$ for ε as small as we wish. The whole argument likewise transfers for such ε with $\varepsilon < 1$. In short, we can reasonably argue

$$\frac{|G(h)-h|_0}{|h|_0} \leq \frac{\varepsilon}{1-\varepsilon}$$

for arbitrarily small ε and as such

$$\lim_{h \to 0} \frac{|G(h)-h|_0}{|h|_0} = 0$$

Thus G is differentiable at zero with $dG_0 = I$. ∇

Now that our little lemma is done, we can return to the proof of the main result. The strategy is to recast the data in the main theorem so that it resembles that of the lemma. Toward this end, we introduce translations on $\mathbb{R}^n$ defined by:

$$T_p(x) = x + p \qquad \& \qquad T_q(x) = x + q.$$

Remember, $F(p) = q$ is the relation of p and q. Let $L = dF_p$ and recall $[dF_p] = J_F(p)$ is assumed nonzero, thus L^{-1} exists and we may define[9]:

$$\widetilde{F} = T_q^{-1} \circ F \circ T_p \circ L^{-1}$$

Observe,

$$\widetilde{F}(0) = (T_q^{-1} \circ F \circ T_p \circ L^{-1})(0) = T_q^{-1}(F((T_p(0)))) = T_q^{-1}(F(p)) = T_q^{-1}(q) = 0$$

Furthermore, $\widetilde{F}$ is differentiable at 0 and we can show

$$d\widetilde{F}_0 = I$$

Indeed, we can argue $\widetilde{F}$ is continuously differentiable near zero on the basis of the given assumption that F is continuously differentiable on W. Thus, for r sufficiently small we can assume $x \in C_r$ implies

$$\|d\widetilde{F}_x - I\|_0 \leq \varepsilon < 1.$$

Then our little lemma applies and we obtain neighborhoods $\widetilde{U}$ and $\widetilde{V}$ of zero as well as a bijection $\widetilde{G} : \widetilde{V} \to \widetilde{U}$ which is differentiable at zero. Moreover, the sequence

$$\widetilde{G}_0(y) = 0, \quad \widetilde{G}_{k+1}(y) = \widetilde{G}_k(y) - \widetilde{F}(\widetilde{G}_k(y)) + y$$

provides succesive approximations whose limit is $\widetilde{G}(y)$.

The rest of the proof is routine. I will turn it into a homework! (you can find it in a slightly different notation on pages 186-187 of Edwards.) □

Remark 10.5.5. In Math 332 we will also cover the implicit mapping theorem which allows us to locally solve $G(x, y) = k$ for $y = f(x)$ given $(x, y) \in \mathbb{R}^n \times \mathbb{R}^m$ where $G : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$. The condition which is sufficient to allow for such local solutions is that the part of the Jacobian matrix corresponding to partial derivatives in y is invertible. Given that condition, we can construct f as the limit of a sequence of successive approximations stemming from linearizing the nonlinear system of equations $G(x, y) = k$ and applying the contraction mapping theorem. This is made precise in Edward's Theorem 3.4 on page 190. I decided against covering it here in the interest of making time for later topics. Also, you can see me do it in Math 332 if you're interested.

10.6 theory of differential equations

This section is based on Chapter 13 of Nagel Saff and Snider's text *Fundamental of Differential Equations and Boundary Value Problems*, fourth edition. Once again we will see Banach's Fixed Point Theorem playing a major role in proof.

Consider the initial value problem:

$$\boxed{y'(x) = f(x, y(x)), \qquad y(x_0) = y_0.}$$

Notice we can integrate the equation above and reformulate it as an integral equation.

$$\int_{x_0}^{x} y'(t)dt = \int_{x_0}^{x} f(t, y(t))dt \quad \Rightarrow \quad \boxed{y(x) = y_0 + \int_{x_0}^{x} f(t, y(t))dt.}$$

[9] I owe you a picture here

Suppose we have a solution of the integral equation:

$$y(x) = y_0 + \int_{x_0}^{x} f(t, y(t))dt \quad \Rightarrow \quad y'(x) = \frac{d}{dx}\left[y_0 + \int_{x_0}^{x} f(t, y(t))dt\right] = f(x, y(x)).$$

and $y(x_0) = y_0 + \int_{x_0}^{x_0} f(t, y(t))dt = y_0$. Thus the integral equation has solutions which solve the given initial value problem. In short, the integral and the given initial value problem are equivalent, they share the same solution set. This equivalence is important as we will find it more convenient to study the integral equation as an operator equation. Define T by

$$T[y](x) = y_0 + \int_{x_0}^{x_0} f(t, y(t))dt.$$

Then, to solve the differential equation is to find y for which $T[y] = y$. We recognize this as the problem of finding a fixed point for T. We can use T to construct a sequence of approximate solutions whose limit gives the desired fixed point. This is known as the **Picard Iteration** in this context:

$$y_{n+1}(x) = T[y_n](x) = y_0 + \int_{x_0}^{x} f(t, y_n(t))dt$$

where $y_0(x) = y_0$ initiates the sequence.

Example 10.6.1. *Consider $y'(x) = 2y(x)$ with $y(0) = 1$. We begin with $y_0(x) = 1$. Then,*

$$T[y](x) = 1 + \int_0^x 2y(t)dt$$

Hence, we calculate:

$$y_1(x) = T[y_0](x) = 1 + \int_0^x 2dt = 1 + 2x.$$

$$y_2(x) = T[y_1](x) = 1 + \int_0^x 2(1 + 2t)dt = 1 + 2x + 2x^2.$$

$$y_3(x) = T[y_2](x) = 1 + \int_0^x 2(1 + 2t + 2t^2)dt = 1 + 2x + 2x^2 + \frac{4}{3}x^3.$$

Notice $\frac{dy}{dx} = 2y$ gives $\int \frac{dy}{y} = \int 2dx$ hence $\ln|y| = 2x + c$ and so $y = ke^{2x}$. But, $y(0) = 1 = k$ thus

$$y = e^{2x} = 1 + 2x + \frac{1}{2!}(2x)^2 + \frac{1}{3!}(2x)^2 + \cdots$$

Thus the Picard iteration agrees well with the power series expansion of the exact solution. Nagel Saff and Snider warn that the example above is fortunate. Given our experience in exploring the iterations underlying the inverse function theorem, I am inclined to believe them.

Definition 10.6.2. *Let $C[a, b]$ be the set of continuous functions from $[a, b]$ to $\mathbb{R}$. Define*

$$\|y\| = \max_{x \in [a,b]} |y(x)|$$

for $y \in C[a, b]$. Furthermore, a sequence of functions $\{y_n\}$ is said to **converge uniformly** *to y if $\lim_{n\to\infty} \|y_n - y\| = 0$. Likewise, a sequence of function is pointwise convergent to y if $\lim_{n\to\infty} y_n(x) = y(x)$ for each $x \in [a, b]$.*

I invite the students to verify that the terminology above is in good agreement the usual definition of uniform convergence. In particular, given the sequence of functions $\{y_n\}$ with domain $[a,b]$ we say the sequence converges uniformly to $y : [a,b] \to \mathbb{R}$ if for each $\varepsilon > 0$ there exists $N \in \mathbb{N}$ for which $n > N$ implies $|y_n(x) - y(x)| < \varepsilon$ for all $x \in [a,b]$. Given the condition $n > N$ we note

$$\|y_n - y\| = \max_{x\in[a,b]} |y_n(x) - y(x)| < \max_{x\in[a,b]} \varepsilon = \varepsilon$$

Thus it is clear that uniform convergence in the usual sense implies uniform convergence in the sense defined above. I should mention, the definition given here for $[a,b]$ can be generalized to some arbitary subset E of a normed linear space. Moreover, a helpful way to encapsulate uniform convergence is given by the following proposition:

Proposition 10.6.3. Worst Case Estimator Formulation of Uniform Convergence: *Suppose $\{f_n\}$ is a sequence of functions on a subset E of the normed linear space V with norm $\|\cdot\|$. If there exists ε_n with $\sup_{x\in E} \|f_n(x) - f(x)\| \leq \varepsilon_n$ and $\varepsilon_n \to 0$ as $n \to \infty$ then $\{f_n\}$ converges uniformly on E. We call ε_n a worst case estimator for the sequence $\{f_n\}$.*

Proof: this shall be homework. □

A series of functions converges pointwise to its sum function if its sequence of partial sums converges to the sum function. Likewise, a series of functions is uniformly convergent if its sequence of partial sums is uniformly convergent. With this terminology in mind we give a useful theorem for determining the uniform convergence of a sequence of functions. This test is oft stated for functions with values in $\mathbb{R}$ or $\mathbb{C}$, but it holds in a far more general context:

Proposition 10.6.4. Weierstrass M-Test: *Suppose $M_k \geq 0$ and $\sum M_k$ converges. Suppose E is a subset of the normed linear space V with norm $\|\cdot\|$ and suppose $g_k : E \to V$ and $\|g_k(x)\| \leq M_k$ for all $x \in E$ then $\sum g_k$ converges uniformly on E.*

Proof: let $x \in E$ and notice $\|g_k(x)\| \leq M_k$ implies $\sum \|g_k(x)\|$ is convergent by the direct comparison test against $\sum M_k$ which is given as convergent. Furthermore, we[10] can prove absolute convergence implies convergence in this context hence $\sum g_k(x) = g(x)$. Furthermore,

$$\|g(x)\| \leq \sum \|g_k(x)\| \leq \sum M_k$$

Observe,

$$\left\|\sum_{k=0}^{n} g_k(x) - g(x)\right\| = \left\|g(x) - \sum_{k=0}^{n} g_k(x)\right\| = \left\|\sum_{k=n+1}^{\infty} g_k(x)\right\| \leq \sum_{k=n+1}^{\infty} M_k = \varepsilon_n.$$

Observe $\sum_{k=n+1}^{\infty} M_k = \varepsilon_n \to 0$ as $n \to \infty$ since $\sum M_k$ converges. Thus the sequence of partial sums for $\sum_{k=0}^{\infty} g_k$ is uniformly convergent and the Weierstrass M-Test follows. □

I have made some (hopefully) mild assumptions about the theory of series in a normed linear space. You will likely get to prove these assertions in the homework. I suppose I should resist the urge to abstract everything and focus on the task at hand.

[10] yeah, you know what I'm about to say here

Proposition 10.6.5. Exchange of limit for integration:
Suppose $\{y_n\}$ is a uniformly convergent sequence of functions in $C[a,b]$ which converge to the function y. Then

$$\lim_{n\to\infty}\int_a^b y_n(x)dx = \int_a^b \lim_{n\to\infty} y_n(x)dx = \int_a^b y(x)dx.$$

Proof: If y_n converges uniformly to y on $[a,b]$ then there exists a worst case estimator ε_n for which $\max_{x\in[a,b]}|y_n(x)-y(x)| \leq \varepsilon_n$ where $\varepsilon_n \to 0$ as $n\to\infty$. Calculate,

$$\left|\int_a^b (y_n(x)-y(x))dx\right| \leq \int_a^b |y_n(x)-y(x)|dx \leq \int_a^b \max_{x\in[a,b]}|y_n(x)-y(x)| \leq \int_a^b \varepsilon_n = (b-a)\varepsilon_n$$

Therefore,

$$\lim_{n\to\infty}\left|\int_a^b (y_n(x)-y(x))dx\right| \leq \lim_{n\to\infty}(b-a)\varepsilon_n = 0.$$

Consequently,

$$\lim_{n\to\infty}\int_a^b (y_n(x)-y(x))dx = 0 \quad\Rightarrow\quad \lim_{n\to\infty}\int_a^b y_n(x)dx = \int_a^b y(x)dx. \ \Box$$

If $\{y_n\}$ is a sequence of continuous V-valued functions with domain $[a,b]$ and the sequence uniformly converges to $y : E \to V$ then we can give essentially the same argument as given above to show the above proposition extends to vector-valued functions on $[a,b]$.

Theorem 10.6.6. Banach Theorem for Operators:
Let $S = \{y \in C[a,b] \mid \|y-y_0\| \leq \alpha\}$ define the set of functions that are within distance α from a given function y_0. Suppose $G : S \to S$ and suppose there exists K with $0 \leq K < 1$ for which

$$\|G[w]-G[z]\| \leq K\|w-z\|$$

for all $z, w \in S$. Then the equation $y = G[y]$ has a unique solution in S. Moreover, $y_{n+1} = G[y_n]$ defines a sequence which converges uniformly to the solution of $y = G[y]$.

Proof: let $y_0 : [a,b] \to \mathbb{R}$ and let S and G satisfy the conditions of the theorem. Define $y_{n+1} = G[y_n]$ for $n \in \mathbb{N}$. This is reasonable as we suppose $G : S \to S$. Moreover, since G is a contraction we can calculate

$$\|y_{j+1}-y_j\| = \|G[y_j]-G[y_{j-1}]\| < K\|y_j - y_{j-1}\|$$

hence, applying the result above repeatedly,

$$\|y_{j+1}-y_j\| < K\|y_j-y_{j-1}\| < K^2\|y_{j-1}-y_{j-2}\| < K^3\|y_{j-2}-y_{j-3}\| < \cdots < K^j\|y_1-y_0\|.$$

We find $\|y_{j+1}-y_j\| < K^j\|y_1-y_0\|$ for $j = 0,1,2,\dots$. Next, following Nagel Saff and Snider, our proof goes a rather different direction. We notice that the sequence $\{y_n\}$ can be envisioned as the $(n+1)-th$ partial sum of the series $y_0 + \sum_{j=0}^{\infty}(y_{j+1}-y_j)$ since

$$y_n = y_0 + (y_1-y_0) + (y_2-y_1) + (y_3-y_2) + \cdots + (y_n - y_{n-1}) = y_0 + \sum_{j=0}^{n}(y_{j+1}-y_j).$$

Therefore, if the series converges uniformly then the sequence converges uniformly as well. The Weierstrass M test gives the desired uniform convergence. In particular, we note the series is **majorized** by the geometric series since

$$\|y_{j+1}-y_j\| < K^j\|y_1-y_0\|$$

and $\sum_{j=0}^{\infty} K^j \|y_1 - y_0\| = \frac{\|y_1 - y_0\|}{1-K} < \infty$. Therefore, $y_n = y_0 + \sum_{j=0}^{\infty}(y_{j+1} - y_j)$ converges uniformly to a function $y_\infty \in S$. Furthermore, as G is a contraction map,

$$\|G[y_\infty] - G[y_n]\| \leq K\|y_\infty - y_n\| \quad \Rightarrow \quad \lim_{n\to\infty} \|G[y_\infty] - G[y_n]\| \leq \lim_{n\to\infty} K\|y_\infty - y_n\| = 0.$$

Therefore, as $\lim \|\heartsuit\| = 0$ if and only if $\lim \heartsuit = 0$,

$$\lim_{n\to\infty} G[y_\infty] = \lim_{n\to\infty} G[y_n] \quad \Rightarrow \quad G[y_\infty] = \lim_{n\to\infty} G[y_n] = \lim_{n\to\infty} y_{n+1} = y_\infty.$$

Thus y_∞ is a fixed point of G. Suppose z is another fixed point of G then $\|y_\infty - z\| = \|G[y_\infty] - G[z]\| \leq K\|y_\infty - z\|$ thus $\|y_\infty - z\| = 0$ and we conclude $z = y_\infty$. □

Theorem 10.6.7. Picard's Existence and Uniqueness Theorem:
Suppose f and $\frac{\partial f}{\partial y}$ are continuous functions in a rectangle $R = \{(x, y) \mid a < x < b, c < y < d\}$ that contains the point (x_0, y_0). Then the initial value problem

$$y'(x) = f(x, y(x)) \qquad \& \qquad y(x_0) = y_0$$

has a unique solution on some interval $[x_0 - h, x_0 + h]$ where $h > 0$. Moreover, the Picard iterations recursively defined by $y_0(x) = y_0$ and

$$y_{n+1}(x) = y_0 + \int_{x_0}^{x} f(t, y_n(t))dt$$

converge uniformly to the solution on $[x_0 - h, x_0 + h]$.

Proof: our goal is to apply Banach's fixed point theorem for operators to $T[y] = y_0 + \int_{x_0}^{x} f(t, y(t))dt$ to show T is a contraction of a function space S where

$$S = \{y \in C(I) \mid \|y - y_0\| \leq \alpha\}$$

and $I = [x_0 - h, x_0 + h]$ for appropriately constructed $\alpha, h > 0$. Naturally the argument centers on using the given continuity of f and $\partial f/\partial y$ to bound values of f and its change with respect to appropriate inputs. I am following the proof in Nagel Saff and Snider, pages 843-845 in the 4^{th} edition of their text.

Choose $h_1, \alpha_1 > 0$ such that

$$R_1 = \{(x, y) \mid |x - x_0| \leq h_1, \ |y - y_0| \leq \alpha_1\} \subseteq R.$$

Note that R_1 is closed hence the extreme value theorem provides bounds M and L for the continuous functions f and $\partial f/\partial y$ on R_1:

$$|f(x, y)| \leq M \qquad \& \qquad \left|\frac{\partial f}{\partial y}(x, y)\right| \leq L$$

for all $(x, y) \in R_1$.

Choose h such that $0 < h < \min\{h_1, \alpha_1/M, 1/L\}$. Let $I = [x_0 - h, x_0 + h]$ and define

$$S = \{y \in C(I) \mid \|y - y_0\|_I \leq \alpha$$

where $\alpha = \alpha_1$ and $||g||_I = \max_{x \in I} |g(x)|$ for $g \in S$. Notice we may use maximum since we know by the extreme value theorem continuous functions on the closed interval I attain a maximum on I. Let $g \in S$ and observe

$$T[g](x) = y_0 + \int_{x_0}^{x} f(t, g(t))dt$$

shows $T[g]$ defines a continuous function on I. Furthermore, for $x \in I$,

$$|T[g](x) - y_0| = \left|\int_{x_0}^{x} f(t, g(t))dt\right| \leq \left|\int_{x_0}^{x} |f(t, g(t))|dt\right| \leq M \left|\int_{x_0}^{x} dt\right| = M|x - x_0| < M\left(\frac{\alpha_1}{M}\right) = \alpha_1.$$

Therefore, if $g \in S$ then $||T[g] - y_0|| \leq \alpha_1$ hence $T[g] \in S$. We find T maps S into S. Next we show T is a contraction on S, let $g, w \in S$ and $x \in I$,

$$\begin{aligned} |T[g](x) - T[w](x)| &= \left|\int_{x_0}^{x} \left[f(t, g(t)) - f(t, w(t))\right] dt\right| \\ &= \left|\int_{x_0}^{x} \frac{\partial f}{\partial y}(t, z(t)) \left[g(t) - w(t)\right] dt\right| \quad \text{using MVT} \\ &= L\left|\int_{x_0}^{x} \left[g(t) - w(t)\right] dt\right| \\ &\leq L||g - w||_I |x - x_0| \\ &\leq Lh||g - w||_I. \end{aligned}$$

Thus T is a contraction on S with contraction constant $K = Lh < 1$ by construction. Therefore, by Banach's fixed point theorem for operators the sequence defined recursively by $y_0(x) = y_0$ and $y_{n+1}(x) = y_0 + \int_{x_0}^{x} f(t, y_n(t))dt$ converges uniformly to the solution y for which $T[y] = y$ on I.

It remains to show any other solution of the initial value problem on I must reduce to y for x sufficiently close to x_0. We encourage the reader to see Nagel Saff and Snider for the remaining details. I have not left much out here, but there is a graph you might appreciate to follow the argument I omit here. □

Theorem 10.6.8. Picard's Existence and Uniqueness Theorem for Systems:
Suppose F and $\frac{\partial F}{\partial x_i}$ are continuous functions in a rectangle

$$R = \{(t, x_1, \ldots, x_n) \mid a < t < b, c_1 < x_i < d_i, i = 1, \ldots, n\}$$

that contains the point $(t_0, \vec{r}_0)$. Then the initial value problem

$$\frac{d\vec{r}}{dt} = F(t, \vec{r}(t)) \qquad \& \qquad \vec{r}(t_0) = \vec{r}_0$$

has a unique solution in the interval $[t_0 - h, t_0 + h]$ where h is some positive constant.

Proof: if $\vec{y} = (y_1, \ldots, y_n)$ is a vector of real-valued continuous functions on $[a, b]$ then $\vec{y} \in C_n[a, b]$ and we may define $||\vec{y}|| = \max_{t \in [a,b]} ||\vec{y}(t)||_2$ where

$$||\vec{y}(t)||_2 = \sqrt{|y_1(t)|^2 + \cdots + |y_n(t)|^2}.$$

Then we can study $T[\vec{r}] = \vec{r}_0 + \int_{t_0}^{t} F(u, \vec{r}(u))du$ and show it serves to define a contraction mapping on an appropriate subset of $C_n[a, b]$. Once more the fixed point of the map will serve to give a unique solution to the initial value problem described in the theorem. I leave the details to the reader. □

Remark 10.6.9. You might be bothered by the sudden appearance of the Euclidean norm in the proof sketch above. Why not use

$$\|\vec{y}(t)\|_1 = |y_1(t)| + \cdots + |y_n(t)|$$

instead ? Well, honestly, if I assign this as homework, feel free. The truth of the matter is that the choice of norm in finite dimensions is largely a matter of taste. Convergence in the Euclidean norm implies convergence in the taxicab norm just the same. So, don't read too much into it, it is largely a matter of style. Proof of the equivalence of norms is found in many texts, I work through the proof from *Introduction to Hilbert Spaces with Applications* by Lokenath Debnath and Piotr Mikusinski in this video: this lecture from my Hilbert Spaces special topics course of Spring 2023.

Theorem 10.6.10. Continuation of Solution:
Suppose F and $\frac{\partial F}{\partial x_i}$ are continuous on the strip

$$R = \{(t, \vec{r}) \mid a \leq t \leq b, \vec{r} \in \mathbb{R}^n\}$$

that contains the point $(t_0, \vec{r}_0)$. Assume also that there exists a positive constant L such that, for $i = 1, \dots, n$, $\left\| \frac{\partial F}{\partial x_i}(t, \vec{r}) \right\| \leq L$ for all $(t, \vec{r}) \in R$. Then the initial value problem

$$\frac{d\vec{r}}{dt} = F(t, \vec{r}(t)) \qquad \& \qquad \vec{r}(t_0) = \vec{r}_0$$

has a unique solution on the entire interval $a \leq t \leq b$.

Proof: following Nagel Saff and Snider, we prove the $n = 1$ case here[11]. Suppose f and $\frac{\partial f}{\partial y}$ are continuous on the strip

$$R = \{(t, y) \mid a \leq t \leq b, y \in \mathbb{R}\}$$

that contains the point (t_0, y_0). Assume also that there exists a positive constant L such that, $\left|\frac{\partial f}{\partial y}\right| \leq L$ on R. Construct

$$T[y](t) = y_0 + \int_{t_0}^{t} f(u, y(u))du$$

Notice if y is continuous on $[a, b]$ then so is $T[y]$ as defined above. Furthermore, if we recursively define $y_0(t) = y_0$ and $y_{n+1} = T[y_n]$ for each $n \in \mathbb{N}$ then we find y_n is continuous on $[a, b]$ by induction. We seek to show uniform convergence of $\{y_n\}$ on $[a, b]$. Once more, note

$$\begin{aligned} y_n(t) &= y_0 + (y_1(t) - y_0(t)) + (y_2(t) - y_1(t)) + \cdots + (y_n(t) - y_{n-1}(t)) \\ &= y_0 + \sum_{j=0}^{n-1} (y_{j+1}(t) - y_j(t)). \end{aligned}$$

Thus uniform convergence of $\{y_n\}$ on $[a, b]$ can be established by showing that the series $\sum_{j=0}^{\infty} (y_{j+1}(t) - y_j)$ converges uniformly on $[a, b]$. This requires a somewhat lengthy argument.

Notice $f(t, y_0)$ is continuous on $[a, b]$ hence there exists M for which $|f(t, y_0)| \leq M$ for all $t \in [a, b]$. We suppose $t \in [t_0, b]$ to simplify the argument a bit[12] Calculate,

$$|y_1(t) - y_0(t)| = |T[y_0](t) - y_0| = \left| \int_{t_0}^{t} f(u, y_0)du \right| \leq \int_{t_0}^{t} |f(u, y_0)|du \leq M(t - t_o). \tag{10.5}$$

[11] homework will look at the higher order proof naturally.

[12] to be fair, we ought to supply similar arguments in the case $t \in [a, t_0]$, however, that seems like homework.

Next, if w, v are two continuous functions on $[t_0, b]$ then using the MVT[13] and that $\left|\frac{\partial f}{\partial y}\right| \leq L$ on R we find for $t \in [t_0, b]$,

$$|f(t, w(t)) - f(t, v(t))| = \left|\frac{\partial f}{\partial y}(t, z(t))|w(t) - v(t)|\right| \leq L|w(t) - v(t)| \tag{10.6}$$

where $z(t)$ is between $w(t)$ and $v(t)$ for each t. Next, consider, applying Equations 10.5 and 10.6,

$$\begin{aligned} |y_2(t) - y_1(t)| &\leq \int_{t_0}^{t} |f(u, y_1(u)) - f(u, y_0)|\, du \\ &\leq L \int_{t_0}^{t} |y_1(u) - y_0|\, du \\ &\leq L \int_{t_0}^{t} M(u - u_0) du \\ &= LM\frac{(t-t_0)^2}{2!} \end{aligned}$$

for t in $[t_0, b]$. Likewise, for $n \in \mathbb{N}$ suppose inductively that $|y_n(t) - y_{n-1}(t)| \leq L^{n-1}M\frac{(t-t_0)^n}{n!}$

$$\begin{aligned} |y_{n+1}(t) - y_n(t)| &\leq \int_{t_0}^{t} |f(u, y_n(u)) - f(u, y_{n-1}(u))|\, du \\ &\leq L \int_{t_0}^{t} |y_{n+1}(u) - y_n(u)|\, du \quad \text{(by Equation 10.5)} \\ &\leq LL^{n-1}M \int_{t_0}^{t} \frac{(u-u_0)^n}{n!} du \quad \text{(by induction hypothesis)} \\ &= L^n M\frac{(t-t_0)^{n+1}}{(n+1)!} \end{aligned}$$

Therefore, for $t \in [t_0, b]$ we find

$$|y_{n+1}(t) - y_n(t)| \leq \frac{M}{L}\frac{(L(t-t_0))^{n+1}}{(n+1)!}$$

Notice that

$$\frac{M}{L}\left(e^{L(t-t_0)} - 1\right) = \frac{M}{L}\left(L(t-t_0) + \frac{(L(t-t_0))^2}{2} + \cdots\right) = \sum_{n=0}^{\infty} \frac{M}{L}\frac{(L(t-t_0))^{n+1}}{(n+1)!}$$

Thus the convergent series of positive terms above serves to majorize the $\sum y_n$ on $[t_0, b]$ where each y_n is continuous hence $\sum y_n$ converges uniformly to the continuous function y_∞ on $[t_0, b]$. Likewise, the students show that the nearly the same argument applies to $[a, t_0]$ thus we find continuous y_∞

[13] notice, here for higher n we'd need to use the generalized mean value estimate theorem we proved earlier in this article

on $[a,b]$ as the limit of the iteration $T[y_n] = y_{n+1}$. Notice,

$$\begin{aligned}
y_\infty(t) &= \lim_{n\to\infty} y_{n+1}(t) \\
&= \lim_{n\to\infty} T[y_n](t) \\
&= \lim_{n\to\infty} \left[y_0 + \int_{t_0}^{t} f(u, y_n(u))du \right] \\
&= y_0 + \lim_{n\to\infty} \int_{t_0}^{t} f(u, y_n(u))du \\
&= y_0 + \int_{t_0}^{t} \lim_{n\to\infty} f(u, y_n(u))du \quad (\text{ by Proposition 10.6.5 }) \\
&= y_0 + \int_{t_0}^{t} f\left(u, \lim_{n\to\infty} y_n(u)\right) du \\
&= y_0 + \int_{t_0}^{t} f(u, y_\infty(u))du
\end{aligned}$$

thus $T[y_\infty] = y_\infty$ so y_∞ solves the given initial value problem on $[a,b]$. It remains to show the solution is unique. Suppose z is a solution hence

$$z(t) = y_0 + \int_{t_0}^{t} f(u, z(u))du.$$

Therefore, if $t \in [a,b]$

$$\begin{aligned}
|z(t) - y_\infty(t)| &\leq \left| \int_{t_0}^{t} |f(u, z(u)) - f(u, y_\infty(u))| \, du \right| \\
&\leq L \left| \int_{t_0}^{t} |z(u) - y_\infty(u)| du \right| \\
&\leq L \|z - y_\infty\| |t - t_0|
\end{aligned}$$

where $\|z - y_\infty\| = \max_{t\in[a,b]} |z(t) - y_\infty(t)|$. Applying what we just derived, similar calculation yields

$$|z(t) - y_\infty(t)| \leq L^2 \|z - y_\infty\| \frac{|t - t_0|^2}{2}$$

then continuing in this fashion we find

$$|z(t) - y_\infty(t)| \leq L^n \frac{|t - t_0|^n}{n!} \|z - y_\infty\|.$$

As $n \to \infty$ we deduce $\|z - y_\infty\| = 0$ hence $z = y_\infty$. $\square$

The conditions of the continuation of solution theorem are naturally met by linear systems.

Theorem 10.6.11. Existence and Uniqueness for Systems of Linear ODEs:

Suppose $t \mapsto A(t)$ defines a continuous function from (a,b) to $\mathbb{R}^{n\times n}$. In other words, suppose $A_{ij}(t)$ defines a continuous function on (a,b) for each i,j with $1 \leq i,j \leq n$. Also, suppose $\vec{f} : (a,b) \to \mathbb{R}^n$ is continuous. Suppose further that $t_0 \in (a,b)$. Then there exists a unique solution of the initial value problem:

$$\frac{d\vec{r}}{dt} = A(t)\vec{r}(t) + \vec{f}(t) \qquad \& \qquad \vec{r}(t_0) = \vec{r}_0$$

on (a,b) for any choice of $\vec{r}_0$.

Proof: suppose $[a', b'] \subset (a, b)$ with $t_0 \in [a', b']$. Define

$$F(t, \vec{r}) = A(t)\vec{r}(t) + \vec{f}(t)$$

and note F is continuous on the strip $R' = [a', b'] \times \mathbb{R}^n$. Furthermore, if we use $\vec{r} = (x_1, x_2, \dots, x_n)$ to denote the component functions of $\vec{r}$ then

$$\frac{\partial F}{\partial x_i}(t, \vec{r}(t)) = \frac{\partial}{\partial x_i}\Big[A(t)\vec{r}(t) + \vec{f}(t)\Big] = A(t)\frac{\partial \vec{r}}{\partial x_i} = A(t)e_i = col_i(A(t)).$$

Thus F and $\frac{\partial F}{\partial x_i}$ are continuous on R'. Furthermore, as $A_{ij} : [a', b'] \to \mathbb{R}$ is continuous have $|A_{ij}(t)| \leq M_{ij}$ for each $t \in [a', b']$. Let $L = \frac{1}{n}\max_{1 \leq i,j \leq n} M_{ij}$ then clearly

$$\left\| \frac{\partial F}{\partial x_i}(t, \vec{r}(t)) \right\| = \|col_i(A(t))\| \leq L$$

thus $\left\| \frac{\partial F}{\partial x_i} \right\| \leq L$ on R' for $i = 1, \dots, n$. Then, by Theorem 10.6.10 we find a unique solution to the initial value problem $\frac{d\vec{r}}{dt} = A(t)\vec{r}(t) + \vec{f}(t)$ with $\vec{r}(t_0) = \vec{r}_0$ on $[a', b']$. But, as $[a', b']$ is an arbitrary we find the solution extends to all of (a, b). □

Theorem 10.6.12. Existence and Uniqueness for n-th order Linear ODE: *Suppose $p_1, \dots, p_n, g$ are continuous real-valued functions on (a, b) and $t_0 \in (a, b)$. For any choice of initial values $y_0, y_1 \dots, y_{n-1}$ there exists a unique solution on (a, b) of the intial value problem given below:*

$$\frac{d^n y}{dt^n} + p_{n-1}\frac{d^{n-1}y}{dt^{n-1}} + \cdots + p_2(t)\frac{d^2 y}{dt^2} + p_1(t)\frac{dy}{dt} + p_0(t)y(t) = g(t)$$

with $y(t_0) = y_0$, $y'(t_0) = y_1, \dots, y^{(n-1)}(t_0) = y_{n-1}$.

Proof: any n-th order linear ODE can be rewritten as n-first order linear ODEs by the reduction of order technique. In particular,

$$x_1 = y, \quad x_2 = \frac{dy}{dt}, \quad x_3 = \frac{d^2y}{dt^2}, \quad \dots \quad x_n = \frac{d^{n-1}y}{dt^{n-1}}$$

Notice the given differential equation yields:

$$\frac{dx_n}{dt} + p_{n-1}x_n + \cdots + p_2x_3 + p_1x_2 + p_0x_1 = g(t)$$

Furthermore,

$$\frac{dx_1}{dt} = x_2, \quad \frac{dx_2}{dt} = x_3, \quad \dots \quad \frac{dx_{n-1}}{dt} = x_n, \quad \frac{dx_n}{dt} = \frac{d^n y}{dt^n}.$$

Therefore,

$$\underbrace{\begin{bmatrix} dx_1/dt \\ dx_2/dt \\ \vdots \\ dx_{n-2}/dt \\ dx_{n-1}/dt \\ dx_n/dt \end{bmatrix}}_{\frac{d\vec{r}}{dt}} = \underbrace{\begin{bmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ -p_0 & -p_1 & -p_2 & \cdots & -p_{n-2} & -p_{n-1} \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-2} \\ x_{n-1} \\ x_n \end{bmatrix}}_{\vec{r}} + \underbrace{\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ g \end{bmatrix}}_{\vec{f}}$$

Thus we find the given n-th order ODE can be expressed as the system of first order ODEs with the form

$$\frac{d\vec{r}}{dt} = A\vec{r} + \vec{f}$$

and $A, \vec{f}$ are continuous on (a, b) hence Theorem 10.6.11 applies and gives a unique solution to the initial value problem with $\vec{r}(t_0) = \vec{r}_0 = (y_0, y_1, \dots, y_{n-1})$ where $t_0 \in (a, b)$. Identify the first component of $\vec{r}$ serves as the solution of the given n-th order problem. Furthermore, as

$$\frac{d\vec{r}}{dt}(t_0) = \big(y(t_0), y'(t_0), \dots, y'_{n-1}(t_0)\big) = (y_0, y_1, \dots, y_{n-1})$$

we find $y(t_0) = y_0$, $y'(t_0) = y_1, \dots, y^{(n-1)}(t_0) = y_{n-1}$ as desired. □

Remark 10.6.13. There is much more to say, but I think I'll stop here for this article. What comes next is the proof of the existence and uniqueness theorem for differential equations in the $\mathcal{A}$-Calculus. What follows next is §3 from my 2017 paper *Introduction to the Theory of $\mathcal{A}$-ODEs* https://arxiv.org/pdf/1708.04137 with Nathan BeDell. That paper needs to be published somewhere.

10.6.1 existence and uniqueness theory for differential equations in the $\mathcal{A}$-Calculus

We assume $\mathcal{A}$ is commutative throughout this Section. If $\vec{g} : \mathcal{A}^m \to \mathcal{A}^k$ then $\vec{g} = (g_1, \dots, g_k)$ is $\mathcal{A}$-differentiable if each component function g_i is $\mathcal{A}$-differentiable in the sense that $\vec{g}$ is real differentiable and has a right-$\mathcal{A}$-linear differential. Let $\mathcal{A}^m$ have algebra variables $z_1, z_2, \dots, z_m$. If $\vec{g}$ is $\mathcal{A}$-differentiable at p then define

$$\frac{\partial \vec{g}}{\partial z_i}(p) = d_p\vec{g}(e_i) \tag{10.7}$$

where $e_1 = (1, 0, \dots, 0), \dots, e_m = (0, \dots, 0, 1)$. Hence generally,

$$d_p\vec{g}(h_1, h_2, \dots, h_m) = h_1\frac{\partial \vec{g}}{\partial z_1}(p) + h_2\frac{\partial \vec{g}}{\partial z_2}(p) + + \cdots + h_m\frac{\partial \vec{g}}{\partial z_m}(p). \tag{10.8}$$

These partial derivatives are used in what follows.

Theorem 10.6.14. *Let I be compact and star-shaped in $\mathcal{A}$ and let $R = I \times \mathcal{A}^k$. Suppose $\vec{f} : R \to \mathcal{A}^k$ is $\mathcal{A}$-differentiable with $||\frac{\partial \vec{f}}{\partial y_i}(z, \vec{y})|| \leq L$ for each $(z, \vec{y}) \in R$. Let $(z_o, \vec{w}_o) \in R$. The initial value problem $\frac{d\vec{y}}{dz} = \vec{f}(z, \vec{y})$ with $\vec{y}(z_o) = \vec{w}_o$ has a unique solution on I.*

Proof: we intend to define the solution as the limit function of a Picard iteration. Begin by setting $\vec{y}_o = \vec{w}_o$ and for $n = 0, 1, \dots$

$$\vec{y}_{n+1}(z) = \vec{w}_o + \int_{[z_o,z]} \vec{f}(\zeta, \vec{y}_n(\zeta)) \star d\zeta. \tag{10.9}$$

Since I is star-shaped we know $[z_o, z] = \{z_o + t(z - z_o) \mid 0 \leq t \leq 1\} \subseteq I$ and consequently the integral is well-defined. In particular, we define for $\vec{f} = (f_1, f_2, \dots, f_k)$,

$$\int_{[z_o,z]} \vec{f} \star d\zeta = \left(\int_{[z_o,z]} f_1 \star d\zeta, \int_{[z_o,z]} f_2 \star d\zeta, \dots, \int_{[z_o,z]} f_k \star d\zeta\right). \tag{10.10}$$

Notice,

$$\vec{y}_o + \sum_{j=0}^{n-1} [\vec{y}_{j+1} - \vec{y}_j] = \vec{y}_o + [\vec{y}_1 - \vec{y}_o] + [\vec{y}_2 - \vec{y}_1] + \cdots + [\vec{y}_n - \vec{y}_{n-1}] = \vec{y}_n. \tag{10.11}$$

Thus uniform convergence of $\sum_{j=0}^{\infty} [\vec{y}_{j+1} - \vec{y}_j]$ provides uniform convergence of $\{\vec{y}_n\}$. We will show that $\sum_{j=0}^{\infty} [\vec{y}_{j+1} - \vec{y}_j]$ can be majorized over I by a convergent series. Then, using [**?**], we deduce the convergence of the series is uniform.

Observe I compact implies there exists $M > 0$ for which $||f(\zeta, \vec{w}_o)|| \leq M$ for all $\zeta \in I$. Thus for $[z_o, z] \subset I$,

$$||\vec{y}_1(z) - \vec{y}_o(z)|| = \left|\left| \int_{[z_o,z]} \vec{f}(\zeta, \vec{w}_o) \star d\zeta \right|\right| \leq M m_{\mathcal{A}} ||z - z_o|| \tag{10.12}$$

as the length of $[z_o, z]$ is simply $||z - z_o||$. We need a generalization of the mean value theorem for our current context to make further progress in the proof:

Lemma 10.6.15. *With $\vec{f}$ and R as in preceding discussion, there exists $l > 0$ for which $||\vec{f}(\zeta, \vec{v}) - \vec{f}(\zeta, \vec{w})|| \leq l||\vec{v} - \vec{w}||$ for $(\zeta, \vec{v}), (\zeta, \vec{w}) \in R$.*

Proof: Notice, if $D\vec{f}$ denotes the Frechet derivative for $\vec{f} : R \to \mathcal{A}^k$ then Theorem 1 on page 73 of [**?**] gives that

$$||\vec{f}(A + H) - \vec{f}(A)|| \leq ||H|| \sup_{x \in [A, A+H]} \{||D\vec{f}(x)||\} \tag{10.13}$$

where we suppose R is given norm by $||(z, \vec{y})|| = \sqrt{||z||^2 + ||y_1||^2 + \cdots + ||y_k||^2}$ for each $(z, \vec{y}) \in R$ and $||D\vec{f}||$ denotes the operator norm defined by

$$||D\vec{f}(x)|| = \sup_{||W||=1}(||D\vec{f}(x)(W)||) \tag{10.14}$$

where $D\vec{f}(x)(W) = (y_1, \ldots, y_k) \in \mathcal{A}^k$ has norm $||(y_1, \ldots, y_k)|| = \sqrt{||y_1||^2 + \cdots + ||y_k||^2}$. Notice, from our initial discussion leading to Equation 10.8,

$$(D\vec{f})(P)(w_o, \vec{w}) = \frac{\partial \vec{f}}{\partial z}(P)w_o + \frac{\partial \vec{f}}{\partial y_1}(P)w_1 + \frac{\partial \vec{f}}{\partial y_2}(P)w_2 + \cdots + \frac{\partial \vec{f}}{\partial y_k}(P)w_k. \tag{10.15}$$

If $P \in [A, A + H] \subset R$ where $H = (0, \vec{h})$ then $w_o = 0$ whereas $w_i = h_i$ for $i = 1, 2, \ldots, k$ so

$$(D\vec{f})(P)(0, \vec{h}) = \frac{\partial \vec{f}}{\partial y_1}(P)h_1 + \frac{\partial \vec{f}}{\partial y_2}(P)h_2 + \cdots + \frac{\partial \vec{f}}{\partial y_k}(P)h_k. \tag{10.16}$$

We assumed $\left|\left|\frac{\partial \vec{f}}{\partial y_i}(z, \vec{y})\right|\right| \leq L$ for each $(z, \vec{y}) \in R$ hence by the triangle inequality and submultiplicativity of the norm on $\mathcal{A}$,

$$||(D\vec{f})(P)(0, \vec{h})|| \leq m_{\mathcal{A}}||h_1||L + \cdots + m_{\mathcal{A}}||h_k||L \leq kLm_{\mathcal{A}}||\vec{h}|| \tag{10.17}$$

and $||D\vec{f}(P)|| \leq kLm_{\mathcal{A}}$ for $P \in [A, A + H] \subset R$ where $H = (0, \vec{h})$. Thus from 10.13 we find

$$||\vec{f}(A + H) - \vec{f}(A)|| \leq ||H||kLm_{\mathcal{A}} \tag{10.18}$$

for $H = (0, \vec{h})$ with $[A, A + H] \subset R$. If $(\zeta, \vec{v}), (\zeta, \vec{w}) \in R$ then set $A = (\zeta, \vec{w})$ and $A + H = (\zeta, \vec{v})$ hence $H = (0, \vec{v} - \vec{w}) \in R$ and for $l = kLm_{\mathcal{A}}$ we find $||\vec{f}(\zeta, \vec{v}) - \vec{f}(\zeta, \vec{w})|| \leq l||\vec{v} - \vec{w}||$. □

We now continue the proof of Theorem 10.6.14. Let $l = kLm_{\mathcal{A}}$ and inductively suppose

$$||\vec{y}_n(z) - \vec{y}_{n-1}(z)|| \leq \frac{Ml^{n-1}m_{\mathcal{A}}^n||z - z_o||^n}{n!} \tag{10.19}$$

for $[z_o, z] \subset I$. Notice Equation 10.12 gives the induction claim for $n = 1$. Consider, for $z \in I$,

$$\begin{aligned} ||\vec{y}_{n+1}(z) - \vec{y}_n(z)|| &= \left|\left| \int_{[z_o,z]} \left(\vec{f}(\zeta, \vec{y}_n(\zeta)) - \vec{f}(\zeta, \vec{y}_{n-1}(\zeta))\right) \star d\zeta \right|\right| \qquad (10.20) \\ &\leq m_{\mathcal{A}} l \int_{[z_o,z]} ||\vec{y}_n(\zeta) - \vec{y}_{n-1}(\zeta)|| \, ||d\zeta|| \quad (\text{ by Lemma 10.6.15, }) \\ &\leq m_{\mathcal{A}} l \int_{[z_o,z]} \frac{Ml^{n-1}m_{\mathcal{A}}^n||\zeta - z_o||^n}{n!} ||d\zeta|| \quad (\text{ by induction claim of 10.19, }) \\ &= \frac{Ml^n m_{\mathcal{A}}^{n+1}}{n!} \int_{[z_o,z]} s^n \, ds \\ &= \frac{Ml^n m_{\mathcal{A}}^{n+1}}{n!} \cdot \frac{||z - z_o||^{n+1}}{n+1} \end{aligned}$$

hence $||\vec{y}_{n+1}(z) - \vec{y}_n(z)|| \leq \frac{Ml^n m_{\mathcal{A}}^{n+1}||z-z_o||^{n+1}}{(n+1)!}$ and we find estimate 10.19 is true for all $n \in \mathbb{N}$ by induction. Furthermore, if s denotes the distance from z_o to z and $\beta = m_{\mathcal{A}}sl$ then we may reformulate the bound of 10.19 as

$$||\vec{y}_n(z) - \vec{y}_{n-1}(z)|| \leq \frac{M}{l} \cdot \frac{\beta^n}{n!}. \tag{10.21}$$

Since I compact we know there exists $s_o > 0$ for which the distance $s = ||z - z_o|| \leq s_o$. Let $\beta_o = m_{\mathcal{A}}ls_o$ and note that

$$||\vec{y}_n(z) - \vec{y}_{n-1}(z)|| \leq \frac{M}{l} \cdot \frac{\beta_o^n}{n!} \tag{10.22}$$

for all $z \in I$. Since $\sum_{n=0}^{\infty} \frac{\beta_o^n}{n!} = e^{\beta_o}$ we have majorized the series $\sum_{n=0}^{\infty} ||\vec{y}_n(z) - \vec{y}_{n-1}(z)||$ on I. Thus, $\sum_{n=0}^{\infty} ||\vec{y}_n(z) - \vec{y}_{n-1}(z)||$ is uniformly convergent on I and we deduce from Equation 10.11 that $\{\vec{y}_n\}$ converges uniformly to $\vec{y}_*$ on I.

Let us examine why $\vec{y}_*$ is a solution to the initial value problem. First, note $\vec{y}_n(z_o) = \vec{w}_o$ and as uniform convergence implies pointwise convergence we have

$$\vec{y}_*(z_o) = \left(\lim_{n\to\infty} \vec{y}_n\right)(z_o) = \lim_{n\to\infty} (\vec{y}_n(z_o)) = \lim_{n\to\infty} \vec{w}_o = \vec{w}_o. \tag{10.23}$$

Second, to see $\vec{y}_*$ is a solution for $z \in I$, consider

$$\vec{y}_*(z) = \lim_{n\to\infty} \vec{y}_n(z) = \lim_{n\to\infty} \left(\vec{w}_o + \int_{[z_o,z]} \vec{f}(\zeta, \vec{y}_{n-1}(\zeta)) \star d\zeta\right). \tag{10.24}$$

However, uniform convergence of $\{\vec{y}_n\}$ and continuity of $\vec{f}$ imply uniform convergence of $\{\vec{f}(\zeta, \vec{y}_{n-1}(\zeta))\}$ therefore we can exchange the order of integration and the limit to deduce

$$\vec{y}_*(z) = \vec{w}_o + \int_{[z_o,z]} \left(\lim_{n\to\infty} \vec{f}(\zeta, \vec{y}_{n-1}(\zeta))\right) \star d\zeta = \vec{w}_o + \int_{[z_o,z]} \vec{f}(\zeta, \vec{y}_*(\zeta)) \star d\zeta. \tag{10.25}$$

Thus, $\frac{d\vec{y}_*}{dz} = \vec{f}(z, \vec{y}_*(z))$ for each $z \in I$.

Finally, to see the solution is unique, suppose $\vec{y}_{**}$ is a solution on I of $\frac{d\vec{y}}{dz} = \vec{f}(z, \vec{y})$ with $\vec{y}_{**}(z_o) = \vec{w}_o$. Let $z \in I$, by Lemma 10.6.15, $||\vec{f}(\zeta, \vec{y}_{**}(\zeta)) - \vec{f}(\zeta, \vec{y}_*(\zeta))|| \leq l||\vec{y}_{**}(\zeta) - \vec{y}_*(\zeta)||$ where $l > 0$. Moreover, $\Upsilon = \sup\{||\vec{y}_{**}(\zeta) - \vec{y}_*(\zeta)|| \mid \zeta \in [z_o, z]\}$ provides a bound for $||\vec{y}_{**}(\zeta) - \vec{y}_*(\zeta)||$ on $[z_o, z]$ hence

$$||\vec{y}_{**}(z) - \vec{y}_*(z)|| = \left|\left| \int_{[z_o,z]} \left(\vec{f}(\zeta, \vec{y}_{**}(\zeta)) - \vec{f}(\zeta, \vec{y}_*(\zeta))\right) \star d\zeta \right|\right| \leq l \cdot \Upsilon \cdot ||z - z_o||. \tag{10.26}$$

Thus, by Lemma 10.6.15 and the estimate above,

$$||\vec{f}(\zeta, \vec{y}_{**}(\zeta)) - \vec{f}(\zeta, \vec{y}_*(\zeta))|| \leq l||\vec{y}_{**}(\zeta) - \vec{y}_*(\zeta)|| \leq l^2\Upsilon||\zeta - z_o||. \tag{10.27}$$

Thus,

$$||\vec{y}_{**}(z) - \vec{y}_*(z)|| \leq l^2\Upsilon \int_{[z_o,z]} ||\zeta - z_o||\, d\zeta = l^2\Upsilon \int_{[z_o,z]} s\, ds = \frac{l^1\Upsilon||z - z_o||}{2}. \tag{10.28}$$

Continuing in the above fashion we find

$$||\vec{y}_{**}(z) - \vec{y}_*(z)|| \leq \frac{l^{n-1}\Upsilon||z - z_o||^n}{n!}. \tag{10.29}$$

for $n \in \mathbb{N}$. As $n \to \infty$ we find $||\vec{y}_{**}(z) - \vec{y}_*(z)|| \to 0$ for each $z \in I$. Thus $\vec{y}_{**}(z) = \vec{y}_*(z)$ for each $z \in I$ and the proof of Theorem 10.6.14 is complete. □

With the Theorem above in hand the remaining theory of linear $\mathcal{A}$-ODEs follows easily.

Theorem 10.6.16. *Let $L = D^k + a_{k-1}D^{k-1} + \cdots + a_2D^2 + a_1D + a_o$ where $a_o, a_1, \dots, a_{k-1}$ are $\mathcal{A}$-differentiable functions on a compact and star-shaped domain I and $D = d/dz$. Also, suppose g is an $\mathcal{A}$-differentiable function on I. The k-th order $\mathcal{A}$-ODE $L[y] = g$ with initial conditions $y(z_o) = y_o, y'(z_o) = y_1, \dots, y^{(k-1)}(z_o) = y_{k-1}$ for $z_o \in I$ has unique solution on I.*

Proof: the proof is by the usual reduction of order. Let $w_1 = y, w_2 = y', \dots, w_k = y^{(k-1)}$. Observe, $\frac{dw_j}{dz} = \frac{dy^{(j-1)}}{dz} = y^{(j)} = w_{j+1}$ for $j = 1, 2, \dots, k-1$. Observe $L[y] = 0$ provides:

$$\begin{aligned} y^{(k)} &= g - a_{k-1}y^{(k-1)} - \cdots - a_2y'' - a_1y' - a_oy \\ &= g - a_ow_1 - a_1w_2 - a_2w_3 - \cdots - a_{k-1}w_k. \end{aligned} \tag{10.30}$$

Thus, as $w_k' = y^{(k)}$, we calculate the reduced system has a coefficient matrix which is a complementary matrix[14] to the characteristic polynomial of the given k-th order $\mathcal{A}$-ODE,

$$\frac{d\vec{w}}{dz} = A\vec{w} + \vec{b} \quad \text{where} \quad A = \begin{bmatrix} 0 & 1 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 1 \\ -a_o & -a_1 & \cdots & -a_{k-2} & -a_{k-1} \end{bmatrix} \quad \& \quad \vec{b} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ g \end{bmatrix}. \tag{10.31}$$

Notice $\vec{f}(z, \vec{w}) = A\vec{w} + \vec{b}$ is $\mathcal{A}$-differentiable since we suppose the coefficient functions $a_o, \dots, a_{k-1}$ and forcing term g are $\mathcal{A}$-differentiable on I. Thus, by Theorem 10.6.14 we find a unique solution to $\frac{d\vec{w}}{dz} = A\vec{w} + \vec{b}$ for a given an initial condition vector $\vec{w}(z_o) = (y_o, y_1, \dots, y_k) \in \mathcal{A}^k$. By construction, $w_1 = y$ of the solution provides the solution to the initial value problem $L[y] = g$ where $y(z_o) = y_o, y'(z_o) = y_1, \dots, y^{(k-1)}(z_o) = y_{k-1}$. □

The set of $\mathcal{A}$-differentiable functions has a natural $\mathcal{A}$-module structure. Hence define:

[14] or the transpose of a complementary matrix if you prefer

Definition 10.6.17. *Let I be a connected subset of $\mathcal{A}$. Suppose $f_j : I \to \mathcal{A}$ are functions. We say the set of functions $\{f_1, f_2, f_3, \dots, f_m\}$ are* **linearly independent (LI)** *on I if and only if for $c_1, \dots, c_m \in \mathcal{A}$*

$$c_1 f_1(z) + c_2 f_2(z) + \cdots + c_m f_m(z) = 0$$

for all $z \in I$ implies $c_1 = c_2 = \cdots = c_m = 0$. Conversely, if $\{f_1, f_2, f_3, \dots, f_m\}$ are not linearly independent on I then they are said to be **linearly dependent** *on I.*

The Wronskian generalizes for suitably differentiable functions on $\mathcal{A}$ in the natural fashion.

Definition 10.6.18. **Wronskian** *of functions $y_1, y_2, \dots, y_m$ at least $(m-1)$ times differentiable at z is given by:*

$$W(y_1, y_2, \dots, y_m; z) = det \begin{bmatrix} y_1(z) & y_2(z) & \cdots & y_m(z) \\ y_1'(z) & y_2'(z) & \cdots & y_m'(z) \\ \vdots & \vdots & \cdots & \vdots \\ y_1^{(m-1)}(z) & y_2^{(m-1)}(z) & \cdots & y_m^{(m-1)}(z) \end{bmatrix}.$$

Notice that the Wronskian is formed by the determinant of a matrix of $\mathcal{A}$-elements for a given z. Fortunately, linear algebra over a commutative ring allows the usual theory of determinants. In particular, $\det : \mathcal{A}^{m \times m} \to \mathcal{A}$ and $M \in \mathcal{A}^{m \times m}$ is invertible if and only if $\det(M) \in \mathcal{A}^\times$. Furthermore, $Mx = 0$ has nontrivial solutions if and only if $\det(M) \in \mathbf{zd}(\mathcal{A})$.

Theorem 10.6.19. *Let $I \subseteq \mathcal{A}$ and suppose $y_1, \dots, y_m : I \to \mathcal{A}$ are at least $(m-1)$-times $\mathcal{A}$-differentiable. If $W(y_1, \dots, y_m; z) \in \mathcal{A}^\times$ for each $z \in I$ then $\{y_1, \dots, y_m\}$ is linearly independent on I.*

Proof: Suppose for all $z \in I$

$$c_1 y_1(z) + c_2 y_2(z) + \cdots + c_m y_m(z) = 0. \tag{10.32}$$

Differentiate $(m-1)$ times to produce the following system of equations over $\mathcal{A}$:

$$\underbrace{\begin{bmatrix} y_1(z) & y_2(z) & \cdots & y_m(z) \\ y_1'(z) & y_2'(z) & \cdots & y_m'(z) \\ \vdots & \vdots & \cdots & \vdots \\ y_1^{(m-1)}(z) & y_2^{(m-1)}(z) & \cdots & y_m^{(m-1)}(z) \end{bmatrix}}_{Y(z)} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \tag{10.33}$$

Thus, Equation 10.32 has only the zero solution if and only if $\det(Y(z)) \in \mathcal{A}^\times$ for each $z \in I$. But, this means $\{y_1, \dots, y_m\}$ is linearly independent on I if and only if $W(y_1, \dots, y_n; z)$ is a unit for each $z \in I$. □

In practice, we are primarily interested in solution sets to linear n-th order $\mathcal{A}$-ODEs where the study of linear independence is greatly simplified by **Abel's formula**. In particular, this formula forces the Wronskian of a full solution set to remain in either $\mathcal{A}^\times$ or $\mathbf{zd}(\mathcal{A})$ throughout the entirety of a connected subset.

Theorem 10.6.20. *Suppose $a_o, a_1, \dots, a_n$ are continuous functions on the connected set $I \subseteq \mathcal{A}$ where $a_o(z) \in \mathcal{A}^\times$ for each $z \in I$. If $y_1, y_2, \dots, y_n$ are solutions of $a_o y^{(n)} + a_1 y^{(n-1)} + \cdots + a_{n-1} y' + a_n y = 0$ then $W(y_1, \dots, y_n; z) = C\, exp \left[\int \frac{a_1}{a_o}\, dz \right]$ for each $z \in I$.*

Proof: suppose $y_1, y_2, \dots, y_n$ are solutions on I for $a_o y^{(n)} + a_1 y^{(n-1)} + \cdots + a_{n-1}y' + a_n y = 0$. Let $Y = [y_1, y_2, \dots, y_n]$ thus $Y' = [y_1', y_2', \dots, y_n']$ and $Y^{(n-1)} = [y_1^{(n-1)}, y_2^{(n-1)}, \dots, y_n^{(n-1)}]$. The determinant which forms Wronskian is given by

$$W = \sum_{i_1, i_2, \dots, i_n = 1}^{n} \epsilon_{i_1 i_2 \dots i_n} Y_{i_1} Y'_{i_2} \cdots Y^{(n-1)}_{i_n} \tag{10.34}$$

where $\epsilon_{i_1 i_2 \dots i_n}$ denoted the completely antisymmetric symbol where $\epsilon_{12\dots n} = 1$. Apply the product rule for n-fold products on each summand in the above sum,

$$W' = \sum_{i_1, \dots, i_n = 1}^{n} \epsilon_{i_1 i_2 \dots i_n} \big(Y'_{i_1} Y'_{i_2} \cdots Y^{(n-1)}_{i_n} + Y_{i_1} Y''_{i_2} Y''_{i_3} \cdots Y^{(n-1)}_{i_n} + \cdots + Y_{i_1} Y'_{i_2} \cdots Y^{(n-2)}_{i_{n-1}} Y^{(n)}_{i_n} \big). \tag{10.35}$$

The term $Y'_{i_1} Y'_{i_2} \cdots Y^{(n-1)}_{i_n} = Y'_{i_2} Y'_{i_1} \cdots Y^{(n-1)}_{i_n}$ hence is symmetric in the pair of indices i_1, i_2. Next, the term $Y_{i_1} Y''_{i_2} Y''_{i_3} \cdots Y^{(n-1)}_{i_n}$ is symmetric in the pair of indices i_2, i_3. This pattern continues up to the term $Y_{i_1} Y'_{i_2} \cdots Y^{(n-1)}_{i_{n-2}} Y^{(n-2)}_{i_{n-1}} Y^{(n-1)}_{i_n}$ which is symmetric in the i_{n-2}, i_{n-1} indices. Thus all the terms vanish when contracted against the antisymmetric symbol. Only one term remains in calculation of W':

$$W' = \sum_{i_1, i_2, \dots, i_n = 1}^{n} \epsilon_{i_1 i_2 \dots i_n} Y_{i_1} Y'_{i_2} \cdots Y^{(n-2)}_{i_{n-1}} Y^{(n)}_{i_n} \tag{10.36}$$

Recall that $y_1, y_2, \dots, y_n$ are solutions of $a_o y^{(n)} + a_1 y^{(n-1)} + \cdots + a_{n-1} y' + a_n y = 0$ hence

$$Y^{(n)} = -\frac{a_1}{a_o} Y^{(n-1)} - \cdots - \frac{a_{n-1}}{a_o} Y' - \frac{a_n}{a_o} Y \tag{10.37}$$

Substitute this into Equation 10.36,

$$\begin{aligned} W' &= \sum_{i_1, i_2, \dots, i_n = 1}^{n} \epsilon_{i_1 i_2 \dots i_n} Y_{i_1} Y'_{i_2} \cdots Y^{(n-2)}_{i_{n-1}} \left[-\frac{a_1}{a_o} Y^{(n-1)} - \cdots - \frac{a_{n-1}}{a_o} Y' - \frac{a_n}{a_o} Y \right]_{i_n} \qquad (10.38) \\ &= \sum_{i_1, i_2, \dots, i_n = 1}^{n} \epsilon_{i_1 i_2 \dots i_n} \left(-\frac{a_1}{a_o} Y_{i_1} Y'_{i_2} \cdots Y^{(n-1)}_{i_n} - \cdots - \frac{a_{n-1}}{a_o} Y_{i_1} Y'_{i_2} \cdots Y'_{i_n} - \frac{a_n}{a_o} Y_{i_1} Y'_{i_2} \cdots Y_{i_n} \right) \\ &= -\frac{a_1}{a_o} \left(\sum_{i_1, i_2, \dots, i_n = 1}^{n} \epsilon_{i_1 i_2 \dots i_n} Y_{i_1} Y'_{i_2} \cdots Y^{(n-1)}_{i_n} \right) \qquad \star \\ &= -\frac{a_1}{a_o} W. \end{aligned}$$

The $\star$ step is based on the observation that the index pairs i_1, i_n and i_2, i_n etc... are symmetric in the line above it hence as they are summed against the completely antisymmetric symbol those terms vanish. Finally, we find $W' = -\frac{a_1}{a_o} W$ and conclude Abel's formula $W(y_1, \dots, y_n; z) = C\, exp\left[-\int \frac{a_1}{a_o}\, dz \right]$ follows by integration since I is connected[15]. □

In a field the only divisor of zero is zero itself hence we need only worry the Wronskian be zero in the ordinary theory. In $\mathcal{A}$-calculus we must also beware of nontrivial divisors of zero.

[15] If I was formed by several connected components then we could have different values for C in different components.

Corollary 10.6.21. *Suppose $a_o, a_1, \dots, a_n$ are continuous functions on the connected set $I \subseteq \mathcal{A}$ where $a_o(z) \in \mathcal{A}^\times$ for each $z \in I$. Let $y_1, y_2, \dots, y_n$ be solutions of $a_oy^{(n)} + a_1y^{(n-1)} + \cdots + a_{n-1}y' + a_ny = 0$. There exists $z_o \in I$ such that $W(y_1, y_2, \dots, y_n; z_o) \in \mathcal{A}^\times$ if and only if $\{y_1, y_2, \dots, y_n\}$ is linearly independent on I. Likewise, there exists $z_o \in I$ such that $W(y_1, y_2, \dots, y_n; z_o) \in \mathbf{zd}(\mathcal{A})$ if and only if $\{y_1, y_2, \dots, y_n\}$ is linearly dependent on I.*

Proof: Theorem 10.6.20 provides $W(y_1, \dots, y_n; z) = C\, exp\left[-\int \frac{a_1}{a_o}\, dz\right]$ for all $z \in I$. If there exists $z_o \in I$ such that $W(y_1, y_2, \dots, y_n; z_o) = Cexp\left[-\int \frac{a_1}{a_o}\, dz\right]\Big|_{z=z_o} \in \mathbf{zd}(\mathcal{A})$ then we find $C \in \mathbf{zd}(\mathcal{A})$ since the image of the exponential is in $\mathcal{A}^\times$. Likewise, if $W(y_1, y_2, \dots, y_n; z_o) \in \mathcal{A}^\times$ then $C \in \mathcal{A}^\times$. Thus the Wronskian of a solution set on a connected subset is either always a zero divisor or always a unit. □

Definition 10.6.22. *Suppose $L[y] = f$ is an n-th order linear differential equation on connected $I \subseteq \mathcal{A}$. We say $S = \{y_1, y_2, \dots, y_n\}$ is a* **fundamental solution set** *of $L[y] = f$ if and only if S is a linearly independent set of solutions to the homogeneous equation; $L[y_j] = 0$ for $j = 1, 2, \dots n$.*

Note the fundamental solution set of $L[y] = f \neq 0$ does not solve $L[y] = f$. We should mention the usual theory for nonhomogeneous differential equations is also naturally generalized to $\mathcal{A}$-calculus. We leave explicit discussion to a future work.

Theorem 10.6.23. *If $L[y] = f$ is an n-th order linear differential equation with continuous coefficient functions on connected $I \subseteq \mathcal{A}$ then there exists a fundamental solution set $S = \{y_1, y_2, \dots, y_n\}$ on I.*

Proof: Apply Theorem 10.6.16 n-times as to select $z_o \in I$ and unique solutions $y_1, \dots, y_n$ for which $y_i^{(j)}(z_o) = \delta_{i,j-1}$ for $0 \leq i \leq n-1$ and $i = 1, \dots, n$. Let the Wronskian at $z = z_o$ for the solution set $\{y_1, y_2, \dots, y_n\}$ be $W(z)$ for the remainder of this proof:

$$W(z_o) = det\begin{bmatrix} y_1(z_o) & y_2(z_o) & \cdots & y_m(z_o) \\ y_1'(z_o) & y_2'(z_o) & \cdots & y_m'(z_o) \\ \vdots & \vdots & \cdots & \vdots \\ y_1^{(n-1)}(z_o) & y_2^{(n-1)}(z_o) & \cdots & y_n^{(n-1)}(z_o) \end{bmatrix} = det\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} = 1.$$

Therefore, by Corollary 10.6.21 the solution set is linearly independent on I. □

Theorem 10.6.24. *If $L[y] = 0$ is an n-th order linear differential equation with continuous coefficient functions and fundamental solution set $S = \{y_1, y_2, \dots, y_n\}$ on connected $I \subseteq \mathcal{A}$. Then if y solves $L[y] = 0$ then there exist unique constants $c_1, c_2, \dots, c_n \in \mathcal{A}$ such that:*

$$y = c_1y_1 + c_2y_2 + \cdots + c_ny_n.$$

Proof: Suppose $L[y] = 0$. If $y = c_1y_1 + c_2y_2 + \cdots c_ny_n$ and $z_o \in I$ then note $y^{(j)}(z_o) = (c_1y_1 + c_2y_2 + \cdots c_ny_n)^{(j)}(z_o)$ for $j = 0, 1, \dots, n-1$. That is, we must solve:

$$\begin{bmatrix} y(z_o) \\ y'(z_o) \\ \vdots \\ y^{(n-1)}(z_o) \end{bmatrix} = \begin{bmatrix} y_1(z_o) & y_2(z_o) & \cdots & y_n(z_o) \\ y_1'(z_o) & y_2'(z_o) & \cdots & y_n'(z_o) \\ \vdots & \vdots & \cdots & \vdots \\ y_1^{(n-1)}(z_o) & y_2^{(n-1)}(z_o) & \cdots & y_n^{(n-1)}(z_o) \end{bmatrix}\begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} \tag{10.39}$$

for $c_1, \dots, c_n \in \mathcal{A}$. Since $\{y_1, \dots, y_n\}$ is a fundamental solution set we know the determinant of the coefficient matrix above is a unit (it is the Wronskian of $y_1, \dots, y_n$ at z_o) hence this system of equations has a unique solution. □