Chapter 2

analytic geometry

In this chapter I will describe n-dimensional Euclidean space and its essential properties. Much of this is not much removed from the discussion of vectors in calculus III. However, we will state as many things as possible for arbitrarily many finite dimensions. Also, we will make use of matrices and linear algebra where it is helpful. For those of you who have not yet taken linear algebra, I have included a few exercises in the Problem sets to help elucidate matrix concepts. If you do those exercises it should help. If you need more examples just ask.

2.1 Euclidean space and vectors

Rene Descartes put forth the idea of what we now call *Cartesian coordinates* for the plane several hundred years ago. The Euclidean concept of geometry predating Descartes seems abstract in comparison. Try graphing without coordinates. In any event, the definition of Cartesian coordinates and $\mathbb{R}^n$ are intertwined in these notes. If we talk about $\mathbb{R}^n$ then we have a preferred coordinate system because the zero point is at the origin.[1]

Definition 2.1.1.

We define $\mathbb{R}^n = \{ (x_1, x_2, \dots, x_n) \mid x_i \in \mathbb{R} \text{ for each } i = 1, 2, \dots, n\}$. If $P = (a_1, a_2, \dots, a_n)$ is a **point** in $\mathbb{R}^n$ then the j-th Cartesian coordinate of the point P is a_j.

Notice that[2] in terms of sets we can write $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ and $\mathbb{R}^3 = \mathbb{R} \times \mathbb{R} \times \mathbb{R}$. Since points in $\mathbb{R}^n$ are in 1-1 correspondance with vectors in $\mathbb{R}^n$ we can add vectors and rescale them by scalar multiplication. If I wish to emphasize that we are working with vectors I may use the notation $< a, b, c > \in V^3$ etc... However, we will think of $\mathbb{R}^n$ as both a set of points and a set of vectors, which takes precendence depends on the context.

[1]some other authors might use $\mathbb{R}^n$ is refer to abstract Euclidean space where no origin is given apriori by the mathematics. Given Euclidean space $\mathcal{E}$ and a choice of an origin $\mathcal{O}$, one can always set-up a 1-1 correspondance with $\mathbb{R}^n$ by mapping the origin to zero in $\mathbb{R}^n$.

[2]Technically these are ambiguous since the Cartesian product of sets is nonassociative but in these notes we identify $\mathbb{R} \times (\mathbb{R} \times \mathbb{R})$ and $(\mathbb{R} \times \mathbb{R}) \times \mathbb{R}$ as the same object. Btw, my Math 200 notes have more on basics of Cartesian products.

Definition 2.1.2.

We define $V^n = \{< v_1, v_2, \ldots, v_n > \mid v_i \in \mathbb{R} \text{ for each } i = 1, 2, \ldots, n\}$.
If $v =< v_1, v_2, \ldots, v_n >$ is a **vector** in $\mathbb{R}^n$ then the j-th component of the vector v is v_j.
Let $v, w \in V^n$ with $v =< v_i >, w =< w_i >$ and $c \in \mathbb{R}$ then we define:

$$v + w =< v_1 + w_1, v_2 + w_2, \ldots, v_n + w_n > \qquad cv =< cv_1, cv_2, \ldots, cv_n > .$$

I will refer to V^n as the set of n-dimensional real vectors. The dot-product is used to define angles and lengths of vectors in V^n.

Definition 2.1.3.

If $v =< v_1, v_2, \ldots, v_n >$ and $w =< w_1, w_2, \ldots, w_n >$ are vectors in V^n then the **dot-product** of v and w is a **real number** defined by:

$$v \cdot w = v_1 w_1 + v_1 w_1 + \cdots + v_n w_n .$$

The **length (or norm) of a vector** $v =< v_1, v_2, \ldots, v_n >$ is denoted $||v||$ and is the real number defined by:

$$||v|| = \sqrt{v \cdot v} = \sqrt{v_1^2 + v_1^2 + \cdots + v_n^2}.$$

If $v =< v_1, v_2, \ldots, v_n > \neq 0$ and $w =< w_1, w_2, \ldots, w_n > \neq 0$ are vectors in V^n then the **angle θ between v and w** is defined by:

$$\theta = \cos^{-1}\left(\frac{v \cdot w}{||v||\,||w||} \right)$$

The vectors v, w are said to be **orthogonal** iff $v \cdot w = 0$.

Example 2.1.4. . .

① $V \cdot 0 = V_1(0) + V_2(0) + \cdots + V_n(0) = 0$

$\therefore$ the zero vector is orthogonal to all vectors.

② $\langle a, b, c \rangle \cdot \hat{i} = \langle a, b, c \rangle \cdot \langle 1, 0, 0 \rangle = a + 0 + 0 = a$

we can select the x-component by taking dot-product with $\hat{i}$.

③ if you need to see more

The dot-product has many well-known properties:

Proposition 2.1.5.

Suppose $x, y, z \in \mathbb{R}^{n\times 1}$ and $c \in \mathbb{R}$ then

1. $x \cdot y = y \cdot x$
2. $x \cdot (y + z) = x \cdot y + x \cdot z$
3. $c(x \cdot y) = (cx) \cdot y = x \cdot (cy)$
4. $x \cdot x \geq 0$ and $x \cdot x = 0$ iff $x = 0$

Notice that the formula $\cos^{-1}\left[\frac{x\cdot y}{||x||\,||y||}\right]$ needs to be justified since the domain of inverse cosine does not contain all real numbers. The inequality that we need for it to be reasonable is $\left|\frac{x\cdot y}{||x||\,||y||}\right| \leq 1$, otherwise we would not have a number in the $dom(cos^{-1}) = range(cos) = [-1, 1]$. An equivalent inequality is $|x \cdot y| \leq ||x||\,||y||$ which is known as the **Cauchy-Schwarz** inequality.

Proposition 2.1.6.

If $x, y \in \mathbb{R}^{n\times 1}$ then $|x \cdot y| \leq ||x||\,||y||$

These properties are easy to justify for the norm we defined in this section.

Proposition 2.1.7.

Let $x, y \in \mathbb{R}^{n\times 1}$ and suppose $c \in \mathbb{R}$ then

1. $||cx|| = |c|\,||x||$
2. $||x + y|| \leq ||x|| + ||y||$

Every nonzero vector can be written as a unit vector scalar multiplied by its magnitude.

$$\boxed{v \in V^n \text{ such that } v \neq 0 \;\Rightarrow\; v = ||v||\hat{v} \text{ where } \hat{v} = \frac{1}{||v||}v.}$$

You should recall that we can write any vector in V^3 as

$$v =< a, b, c >= a < 1, 0, 0 > +b < 0, 1, 0 > +c < 0, 0, 1 >= a\hat{i} + b\hat{j} + c\hat{k}$$

where we defined the $\hat{i} =< 1, 0, 0 >$, $\hat{j} =< 0, 1, 0 >$, $\hat{k} =< 0, 0, 1 >$. You can easily verify that distinct Cartesian unit-vectors are orthogonal. Sometimes we need to produce a vector which is orthogonal to a given pair of vectors, it turns out the cross-product is one of two ways to do that in V^3. We will see much later that this is special to three dimensions.

Definition 2.1.8.

If $A =< A_1, A_2, A_3 >$ and $B =< B_1, B_2, B_3 >$ are vectors in V^3 then the **cross-product** of A and B is a **vector** $A \times B$ which is defined by:

$$\vec{A} \times \vec{B} =< A_2B_3 - A_3B_2, A_3B_1 - A_1B_3, A_1B_2 - A_2B_1 > .$$

The magnitude of $\vec{A} \times \vec{B}$ can be shown to satisfy $||\vec{A} \times \vec{B}|| = ||\vec{A}|| \, ||\vec{B}|| \sin(\theta)$ and the direction can be deduced by **right-hand-rule**. The right hand rule for the unit vectors yields:

$$\hat{i} \times \hat{j} = \hat{k}, \quad \hat{k} \times \hat{i} = \hat{j}, \quad \hat{j} \times \hat{k} = \hat{i}$$

If I wish to discuss both the point and the vector to which it corresponds we may use the notation

$$\boxed{P = (a_1, a_2, \dots, a_n) \longleftrightarrow \vec{P} =< a_1, a_2, \dots, a_n >}$$

With this notation we can easily define directed line-segments as the vector which points from one point to another, also the distance bewtween points is simply the length of the vector which points from one point to the other:

Definition 2.1.9.

Let $P, Q \in \mathbb{R}^n$. The directed line segment from P to Q is $\overrightarrow{PQ} = \vec{Q} - \vec{P}$. This vector is drawn from tail Q to the tip P where we denote the direction by drawing an arrowhead. The **distance between** P **and** Q is $d(P, Q) = ||\overrightarrow{PQ}||$.

2.1.1 compact notations for vector arithmetic

I prefer the following notations over the hat-notation of the preceding section because this notation generalizes nicely to n-dimensions.

$$\boxed{e_1 =< 1, 0, 0 > \qquad e_2 =< 0, 1, 0 > \qquad e_3 =< 0, 0, 1 > .}$$

Likewise the Kronecker delta and the Levi-Civita symbol are at times very convenient for abstract calculation:

$$\delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} \qquad \epsilon_{ijk} = \begin{cases} 1 & (i, j, k) \in \{(1, 2, 3), (3, 1, 2), (2, 3, 1)\} \\ -1 & (i, j, k) \in \{(3, 2, 1), (2, 1, 3), (1, 3, 2)\} \\ 0 & \text{if any index repeats} \end{cases}$$

An equivalent definition for the Levi-civita symbol is simply that $\epsilon_{123} = 1$ and it is antisymmetric with respect to the interchange of any pair of indices;

$$\epsilon_{ijk} = \epsilon_{jki} = \epsilon_{kij} = -\epsilon_{kji} = -\epsilon_{jik} = -\epsilon_{ikj}.$$

Now let us restate some earlier results in terms of the Einstein repeated index conventions[3], let $\vec{A}, \vec{B} \in V^n$ and $c \in \mathbb{R}$ then

$\vec{A} = A_k e_k$	standard basis expansion
$e_i \cdot e_j = \delta_{ij}$	orthonormal basis
$(\vec{A} + \vec{B})_i = \vec{A}_i + \vec{B}_i$	vector addition
$(\vec{A} - \vec{B})_i = \vec{A}_i - \vec{B}_i$	vector subtraction
$(c\vec{A})_i = c\vec{A}_i$	scalar multiplication
$\vec{A} \cdot \vec{B} = A_k B_k$	dot product
$(\vec{A} \times \vec{B})_k = \epsilon_{ijk} A_i B_j$	cross product.

All but the last of the above are readily generalized to dimensions other than three by simply increasing the number of components. However, the cross product is special to three dimensions. I can't emphasize enough that the formulas given above for the dot and cross products can be utilized to yield great efficiency in abstract calculations.

Example 2.1.10. . .

Prove $\vec{A} \cdot (\vec{B} \times \vec{C}) = \vec{C} \cdot (\vec{A} \times \vec{B})$.

$$
\begin{aligned}
\vec{A} \cdot (\vec{B} \times \vec{C}) &= A_k \, (\vec{B} \times \vec{C})_k \\
&= A_k \, \epsilon_{ijk} B_i C_j \\
&= \epsilon_{ijk} A_k B_i C_j \\
&= \epsilon_{ijk} C_j A_k B_i \\
&= \epsilon_{kij} C_j A_k B_i \\
&= C_j \, \epsilon_{kij} A_k B_i \\
&= C_j \, (\vec{A} \times \vec{B})_j \\
&= \vec{C} \cdot (\vec{A} \times \vec{B}).
\end{aligned}
$$

notice $\epsilon_{ijk} = -\epsilon_{ikj} = -(-\epsilon_{kij})$

$\therefore \ \epsilon_{ijk} = \epsilon_{kij}$

Another example, you can prove $\epsilon_{ikj}\epsilon_{mpj} = \delta_{im}\delta_{kp} - \delta_{ip}\delta_{km}$ then this identity can be used to evaluate triple cross products,

$$
\begin{aligned}
[\vec{A} \times (\vec{B} \times \vec{C})]_k &= \epsilon_{ijk} A_i \, (\vec{B} \times \vec{C})_j \\
&= \epsilon_{ijk} \epsilon_{mpj} A_i B_m C_p \\
&= -\epsilon_{ikj} \epsilon_{mpj} A_i B_m C_p \\
&= -\delta_{im}\delta_{kp} A_i B_m C_p + \delta_{ip}\delta_{km} A_i B_m C_p
\end{aligned}
$$

[3]there are more details to be seen in the Appendix if you're curious

$$
\begin{aligned}
&= -A_i B_i C_k + A_i B_k C_i \\
&= -(\vec{A} \cdot \vec{B}) C_k + (\vec{A} \cdot \vec{C}) B_k \\
&= \left[(\vec{A} \cdot \vec{C}) \vec{B} - (\vec{A} \cdot \vec{B}) \vec{C} \right]_k
\end{aligned}
$$

$\therefore \ \underline{\vec{A} \times (\vec{B} \times \vec{C}) = (\vec{A} \cdot \vec{C}) \vec{B} - (\vec{A} \cdot \vec{B}) \vec{C}}$.

2.2 matrices

An $m \times n$ matrix is an array of numbers with m-rows and n-columns. We define $\mathbb{R}^{m \times n}$ to be the set of all $m \times n$ matrices. The set of all n-dimensional column vectors is $\mathbb{R}^{n \times 1}$. The set of all n-dimensional row vectors is $\mathbb{R}^{1 \times n}$. A given matrix $A \in \mathbb{R}^{m \times n}$ has mn-components A_{ij}. Notice that the components are numbers; $A_{ij} \in \mathbb{R}$ for all i, j such that $1 \leq i \leq m$ and $1 \leq j \leq n$. We should not write $A = A_{ij}$ because it is nonesense, however $A = [A_{ij}]$ is quite fine.

Suppose $A \in \mathbb{R}^{m \times n}$, note for $1 \leq j \leq n$ we have $col_j(A) \in \mathbb{R}^{m \times 1}$ whereas for $1 \leq i \leq m$ we find $row_i(A) \in \mathbb{R}^{1 \times n}$. In other words, an $m \times n$ matrix has n columns of length m and n rows of length m.

Definition 2.2.1.

Two matrices A and B are equal iff $A_{ij} = B_{ij}$ for all i, j. Given matrices A, B with components A_{ij}, B_{ij} and constant $c \in \mathbb{R}$ we define

$$(A + B)_{ij} = A_{ij} + B_{ij} \qquad (cA)_{ij} = cA_{ij} \qquad , \text{ for all } i, j.$$

The **zero matrix** in $\mathbb{R}^{m \times n}$ is denoted 0 and defined by $0_{ij} = 0$ for all i, j. The additive inverse of $A \in \mathbb{R}^{m \times n}$ is the matrix $-A$ such that $A + (-A) = 0$. The components of the additive inverse matrix are given by $(-A)_{ij} = -A_{ij}$ for all i, j. Likewise, if $A \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{n \times p}$ then the product $AB \in \mathbb{R}^{m \times p}$ is defined by:

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

for each $1 \leq i \leq m$ and $1 \leq j \leq p$. In the case $m = p = 1$ the indices i, j are omitted in the equation since the matrix product is simply a number which needs no index. The identity matrix in $\mathbb{R}^{n \times n}$ is the $n \times n$ square matrix I whose components are the Kronecker delta; $I_{ij} = \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$. The notation I_n is sometimes used if the size of the identity matrix needs emphasis, otherwise the size of the matrix I is to be understood from the context.

$$I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Let $A \in \mathbb{R}^{n \times n}$. If there exists $B \in \mathbb{R}^{n \times n}$ such that $AB = I$ and $BA = I$ then we say that A is **invertible** and $A^{-1} = B$. Invertible matrices are also called **nonsingular**. If a matrix has no inverse then it is called a **noninvertible** or **singular** matrix. Let $A \in \mathbb{R}^{m \times n}$ then $A^T \in \mathbb{R}^{n \times m}$ is called the **transpose** of A and is defined by $(A^T)_{ji} = A_{ij}$ for all $1 \leq i \leq m$ and $1 \leq j \leq n$. Note **dot-product** of $v, w \in V^n$ is given by $v \cdot w = v^T w$.

Remark 2.2.2.

We will use the convention that points in $\mathbb{R}^n$ are column vectors. However, we will use the somewhat subtle notation $(x_1, x_2, \dots x_n) = [x_1, x_2, \dots x_n]^T$. This helps me write $\mathbb{R}^n$ rather than $\mathbb{R}^{n\times 1}$ and I don't have to pepper transposes all over the place. If you've read my linear algebra notes you'll appreciate the wisdom of our convention. Likewise, for the sake of matrix multiplication, we adopt the subtle convention $< x_1, x_2, \dots x_n >= [x_1, x_2, \dots x_n]^T$ for vectors in V^n. Worse yet I will later in the course fail to distinguish between V^n and $\mathbb{R}^n$. Most texts adopt the view that points and vectors can be identified so there is no distinction made between these sets. We also follow that view, however I reserve the right to use V^n if I wish to emphasize that I am using vectors.

Definition 2.2.3.

Let $e_i \in \mathbb{R}^n$ be defined by $(e_i)_j = \delta_{ij}$. The size of the vector e_i is determined by context. We call e_i the i-th standard basis vector.

Example 2.2.4. . .

$$e_1 \in \mathbb{R}^2 \Rightarrow e_1 = (1,0)$$

$$e_1 \in \mathbb{R}^3 \Rightarrow e_1 = (1,0,0)$$

Also, note $\vec{A} = \langle A_1, A_2, \dots, A_n \rangle = \sum_{i=1}^{n} A_i e_i$

and $\vec{A} \cdot e_j = \sum_{i=1}^{n} A_i e_i \cdot e_j = \sum_{i=1}^{n} A_i \delta_{ij} = A_j$.

Definition 2.2.5.

The ij-th **standard basis matrix** for $\mathbb{R}^{m\times n}$ is denoted E_{ij} for $1 \leq i \leq m$ and $1 \leq j \leq n$. The matrix E_{ij} is zero in all entries except for the (i,j)-th slot where it has a 1. In other words, we define $(E_{ij})_{kl} = \delta_{ik}\delta_{jl}$.

Theorem 2.2.6.

Assume $A \in \mathbb{R}^{m\times n}$ and $v \in \mathbb{R}^{n\times 1}$ and define $(E_{ij})_{kl} = \delta_{ik}\delta_{jl}$ and $(e_i)_j = \delta_{ij}$ as before then,

$$v = \sum_{i=1}^{n} v_n e_n \qquad A = \sum_{i=1}^{m}\sum_{j=1}^{n} A_{ij}E_{ij}.$$

$$[e_i{}^T A] = row_i(A) \qquad [Ae_i] = col_i(A) \qquad A_{ij} = (e_i)^T A e_j$$

$$E_{ij}E_{kl} = \delta_{jk}E_{il} \qquad E_{ij} = e_i e_j{}^T \qquad e_i{}^T e_j = e_i \cdot e_j = \delta_{ij}$$

You can look in my linear algebra notes for the details of the theorem. I'll just expand one point here: Let $A \in \mathbb{R}^{m\times n}$ then

$$\begin{aligned} A &= \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ A_{m1} & A_{m2} & \cdots & A_{mn} \end{bmatrix} \\ &= A_{11}\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & 0 \\ 0 & 0 & \cdots & 0 \end{bmatrix} + A_{12}\begin{bmatrix} 0 & 1 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & 0 \\ 0 & 0 & \cdots & 0 \end{bmatrix} + \cdots + A_{mn}\begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & 0 \\ 0 & 0 & \cdots & 1 \end{bmatrix} \\ &= A_{11}E_{11} + A_{12}E_{12} + \cdots + A_{mn}E_{mn}. \end{aligned}$$

The calculation above follows from repeated mn-applications of the definition of matrix addition and another mn-applications of the definition of scalar multiplication of a matrix.

Example 2.2.7. . . .

Let $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ if I want to select the (ij)-component then I can multiply by e_i^T on the left & e_j on right ; $A_{ij} = e_i^T A e_j$

For example,

$$\begin{aligned} e_1^T A e_2 &= [1, 0]\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} \\ &= [1, 0]\begin{bmatrix} b \\ d \end{bmatrix} \\ &= \underline{b = A_{12}}. \end{aligned}$$

this idea is useful when we study quadratic forms $Q(x) = x^T A x$.

2.3 linear transformations

We should recall the precise definition of a linear combination: A **linear combination** of objects $A_1, A_2, \dots, A_k$ is a sum $c_1A_1 + c_2A_2 + \cdots + c_kA_k = \sum_{i=1}^{k} c_iA_i$ where $c_i \in \mathbb{R}$ for each i. Essentially, a vector space is simply a set of objects called "vectors" for which any linear combination of the vectors is again in the set. In other words, vectors in a vector space can be added by "vector addition" or rescaled by a so-called "scalar multiplication". A linear transformation is a mapping from one vector space to another which preserves linear combinations.

Definition 2.3.1.

Let V, W be vector spaces. If a mapping $L : V \to W$ satisfies

1. $L(x + y) = L(x) + L(y)$ for all $x, y \in V$,
2. $L(cx) = cL(x)$ for all $x \in V$ and $c \in \mathbb{R}$

then we say L is a linear transformation.

Example 2.3.2. . .

$L(x) = x \cdot x$ for $x \in \mathbb{R}^n$, this means that $L : \mathbb{R}^n \to \mathbb{R}$. Notice that

$$\begin{aligned} L(x+y) &= (x+y)\cdot(x+y) \\ &= x\cdot x + x\cdot y + y\cdot x + y\cdot y \\ &= L(x) + 2x\cdot y + L(y) \end{aligned}$$

$\Rightarrow L$ <u>not</u> linear since we can find x, y with $x \cdot y \neq 0$.

Example 2.3.3. . .

$L(x) = mx + b$ for all $x \in \mathbb{R}$ and m, b are fixed constants.

Note $L(0+0) = b$

but $L(0) + L(0) = b + b = 2b$

The function $L : \mathbb{R} \to \mathbb{R}$ whose graph is a line need not be a linear transformati

Example 2.3.4. . .

$L(x_1, x_2, x_3) = x_1$ is a mapping from $\mathbb{R}^3 \to \mathbb{R}$. We can show this is linear

$$L(x+y) = (x+y)_1 = x_1 + y_1 = L(x) + L(y)$$

$$L(cx) = (cx)_1 = cx_1 = cL(x)$$

for all $x, y \in \mathbb{R}^3$.

Definition 2.3.5.

Let $L : \mathbb{R}^{n\times 1} \to \mathbb{R}^{m\times 1}$ be a linear transformation, the matrix $A \in \mathbb{R}^{m\times n}$ such that $L(x) = Ax$ for all $x \in \mathbb{R}^{n\times 1}$ is called the **standard matrix** of L. We denote this by $[L] = A$ or more compactly, $[L_A] = A$, we say that L_A is the linear transformation induced by A.

Example 2.3.6. . .

$$L(x, y) = (x+y,\ x-y,\ y) = \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$$

note $L : \mathbb{R}^2 \to \mathbb{R}^3$ need a 3×2 matrix.

$$\therefore \quad [L] = \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 0 & 1 \end{bmatrix}$$

Example 2.3.7. . .

$$L(x_1, x_2, x_3, x_4) = (x_1 + x_4,\ x_2 - 3x_3,\ x_4 + 3) = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & -3 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 3 \end{bmatrix}$$

This mapping $L : \mathbb{R}^4 \to \mathbb{R}^3$ is <u>not</u> linear due to the $(0,0,0,3)$ vector.

$$L(0+0) = L(0) = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 3 \end{bmatrix}$$

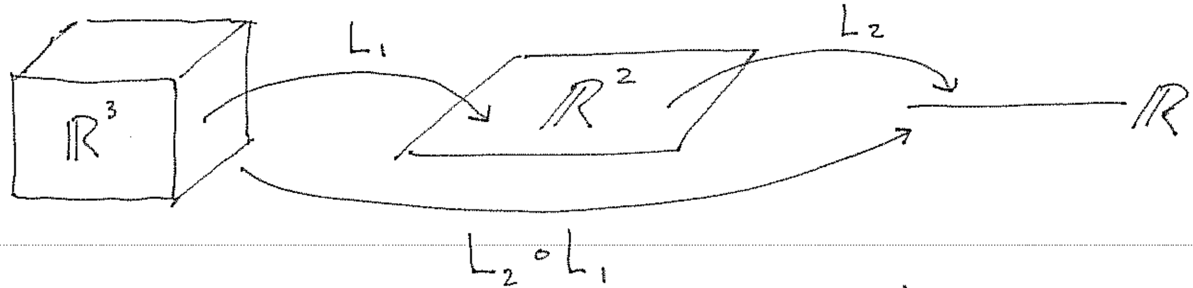
$$L(0) + L(0) = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 3 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 6 \end{bmatrix} \neq L(0+0)$$

<u>Remark</u>: if $L(0) \neq 0$ then L is <u>not</u> linear transformation. This is a useful criteria since 0 is easy to check.

Proposition 2.3.8.

> Let V_1, V_2, V_3 be vector spaces and suppose $L_1 : V_1 \to V_2$ and $L_2 : V_2 \to V_3$ are linear transformations then $L_2 \circ L_1 : V_1 \to V_3$ is a linear transformation and if V_1, V_2 are column spaces then $[L_2 \circ L_1] = [L_2][L_1]$.

Example 2.3.9. . .



If $L_1(x) = (x_1, x_2)$ and $L_2(x_1, x_2) = x_1 + x_2$

then $L_1(x_1, x_2, x_3) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ & $L_2(x_1, x_2) = [1, 1]\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

Hence, $[L_1] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$ and $[L_2] = [1, 1]$.

$$\Rightarrow \underline{[L_2][L_1] = [1, 1]\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} = [1, 1, 0]}. \quad \text{I}$$

In contrast we can calculate the composite directly,

$$\begin{aligned} (L_2 \circ L_1)(x_1, x_2, x_3) &= L_2(L_1(x_1, x_2, x_3)) \\ &= L_2(x_1, x_2) \\ &= x_1 + x_2 \\ &= [1 \quad 1 \quad 0]\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \end{aligned}$$

$$\therefore \underline{[L_2 \circ L_1] = [1, 1, 0]}. \quad \text{II}$$

Comparing I & II you see $[L_2 \circ L_1] = [L_2][L_1]$.

this algebra $\Rightarrow$ chain rule of multivariate calculus, we'll see it.

2.4 orthogonal transformations

Orthogonal transformations play a central role in the study of geometry.

Definition 2.4.1.

If $T : \mathbb{R}^{n\times 1} \rightarrow \mathbb{R}^{n\times 1}$ is a linear transformation such that $T(x) \cdot T(y) = x \cdot y$ for all $x, y \in \mathbb{R}^{n\times 1}$ then we say that T is an **orthogonal transformation**. The matrix R of an orthogonal transformation is called an **orthogonal matrix** and it satisfies $R^T R = I$. The set of orthogonal matrices is $O(n)$ and the subset of rotation matrices is denoted $SO(n) = \{R \in O(n) | det(R) = 1\}$.

The definition above is made so that an orthogonal transformation preserves the lengths of vectors and the angle between pairs of vectors. Since both of those quantities are defined in terms of the dot-product it follows that lengths and angles are invariant under a linear transformation since the dot-product is unchanged. In particular,

$$||T(x)||^2 = T(x) \cdot T(x) = x \cdot x = ||x||^2 \quad \Rightarrow \quad \boxed{||T(x)|| = ||x||}$$

Likewise, defining θ to be the angle between x, y and θ_T the angle between $T(x), T(y)$:

$$T(x) \cdot T(y) = x \cdot y \;\Rightarrow\; ||T(x)||||T(y)|| \cos\theta_T = ||x||||y|| \cos\theta \;\Rightarrow\; \cos\theta_T = \cos\theta \;\Rightarrow\; \boxed{\theta_T = \theta}$$

2.5 orthogonal bases

Definition 2.5.1.

A set S of vectors in $\mathbb{R}^{n\times 1}$ is **orthogonal** iff every pair of vectors in the set is orthogonal. If S is orthogonal and all vectors in S have length one then we say S is **orthonormal**.

It is easy to see that an orthogonal transformation maps an orthonormal set to another orthonormal set. Observe that the standard basis $\{e_1, e_2, \ldots, e_n\}$ is an orthonormal set of vectors since $e_i \cdot e_j = \delta_{ij}$. When I say the set is a **basis** for $\mathbb{R}^n$ this simply means that it is a set of vectors which **spans** $\mathbb{R}^n$ by finite linear combinations and is also **linearly independent**. In case you haven't had linear,

Definition 2.5.2.

1. $S = \{v_1, v_2 \ldots, v_k\}$ is **linearly independent** iff $\sum_{i=1}^k c_i v_i = 0$ implies $c_i = 0$ for $i = 1, 2, \ldots k$.

2. $S = \{v_1, v_2 \ldots, v_k\}$ is **spans** W iff for each $w \in W$ there exist constants $w_1, w_2, \ldots, w_k$ such that $w = \sum_{i=1}^k w_i v_i$.

3. β is a **basis** for a vector space V iff it is a linearly independent set which spans V. Moreover, if there are n vectors in β then we say $dim(V) = n$.

In fact, since the dimension of $\mathbb{R}^n$ is known to be n either spanning or linear independence of a set of n vectors is a sufficient condition to insure a given set of vectors is a basis for $\mathbb{R}^n$. In any event, we can prove that an orthonormal set of vectors is linearly independent. So, to summarize, if we have a linear transformation T we can construct a new orthonormal basis from the standard basis:

$$T(\{e_1, \dots, e_n\}) = \{T(e_1), \dots T(e_n)\}$$

Example 2.5.3. *In calculus III you hopefully observed (perhaps not in this langauge, but the patterns were there just waiting to be noticed):*

1. *a line through the origin is spanned by its direction vector.*
2. *a plane through the origin is spanned by any two non-paralell vectors that lie in that plane.*
3. *three dimensional space is spanned by three non-coplanar vectors. For example, $\hat{i}, \hat{j}, \hat{k}$ span $\mathbb{R}^3$.*

2.6 coordinate systems

Definition 2.6.1.

A **coordinate system** of $\mathbb{R}^n$ is a set of n functions $\bar{x}_i : \mathbb{R}^n \to \mathbb{R}$ for $i = 1, 2, \dots, n$ such that we can invert the equations

$$\bar{x}_i = \bar{x}_i(x_1, x_2, \dots, x_n) \qquad \text{to obtain} \qquad x_i = x_i(\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n)$$

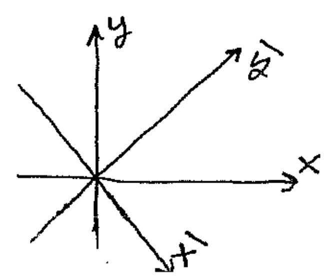
on most of $\mathbb{R}^n$ In other words, we can group the functions into a coordinate map $\Phi = \bar{x} = (\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n)$ and $\bar{x}$ is a 1-1 correspondance on most of $\mathbb{R}^n$. We call $\bar{x}_j$ the **j-th coordinate** of the $\bar{x}$ coordinate system. For a particular coordinate system we also define the **j-th coordinate axis** to be the set of points such that all the other coordinates are zero. If the coordinate axis is a line for each coordinate then the coordinate system is said to be **rectilinear**. If the coordinate axis is not a line for all the coordinates then the coordinate system is said to be **curvelinear**. If the coordinate axes have orthgonal unit vectors then the coordinate system is said to be an **orthogonal coordinate system**. Likewise, if the coordinate curves of a curvelinear coordinate system have orthogonal tangent vectors at all points then the curvelinear coordinate system is said to give an **orthogonal coordinate system**.

Example 2.6.2. .

Let $\bar{x} = x - y$ and $\bar{y} = x + y$.
Notice that $\bar{x} + \bar{y} = 2x$ whereas $\bar{x} - \bar{y} = -2y$ hence we have inverse relations: $x = \frac{1}{2}(\bar{x} + \bar{y})$ & $y = \frac{1}{2}(-\bar{x} + \bar{y})$.
The coordinate axes are found by setting $\bar{x} = 0$ or $\bar{y} = 0$,

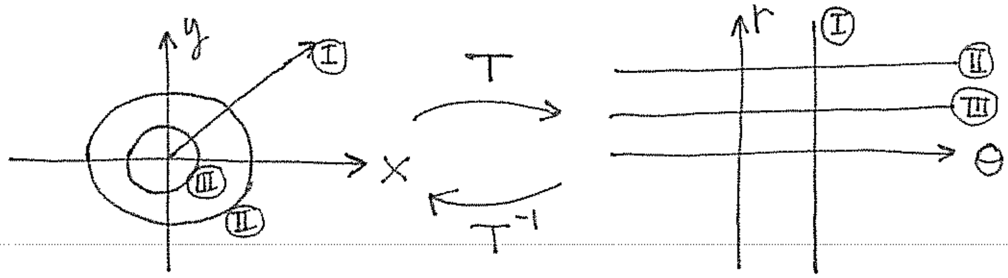
$\bar{x}$ axis has $\bar{y} = 0$ $\to$ $x = (1/2)\bar{x}$, $y = (-1/2)\bar{x}$ $\to$ $y = -x$ eq^{n} of $\bar{x}$ axis

$\bar{y}$ axis has $\bar{x} = 0$ $\to$ $x = \frac{1}{2}\bar{y}$, $y = \frac{1}{2}\bar{y}$ $\to$ $y = x$ eq^{n} of $\bar{y}$ axis



The case of Cartesian coordinates has $\Phi = Id$. Conceptually we think of the codomain as a different space than the domain in general. For example, in the case of polar coordinates on the plane we have a mapping $\Phi : \mathbb{R}^2 \to \mathbb{R}^2$ where a circle in the domain becomes a line in the range. The line in $r\theta$ space is a representation of the circle in the view of polar coordinates. Students often confuse themselves by implicitly insisting that the domain and range of the coordinate map are the same copy of $\mathbb{R}^n$ but this is the wrong concept. Let me illustrate with a few mapping pictures:

Example 2.6.3. .



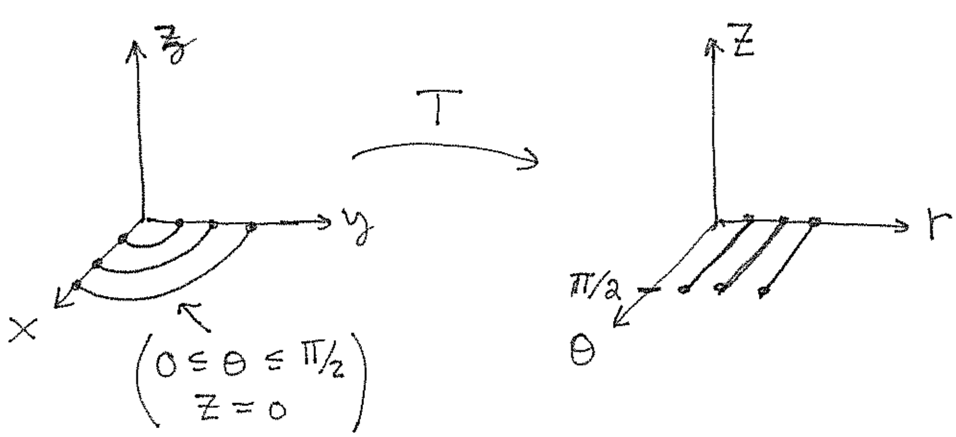
$$T(x,y) = (\theta, r) = \left(\tan^{-1}(y/x),\ \sqrt{x^2+y^2}\right)$$
$$T^{-1}(\theta, r) = (r\cos\theta, r\sin\theta)$$
$$T : \mathbb{R}^2_{xy} \longrightarrow \mathbb{R}^2_{\theta r}$$, polar coordinates are curvelinear.

Generally I admit that I'm being a bit vague here because the common useage of the term coordinate system is a bit vague. Later I'll define a patched manifold and that structure will give a refinement of the coordinate concept which is unambiguous. That said, common coordinate systems such as polar, spherical coordinates fail to give coordinates for manifolds unless we add restrictions on the domain of the coordinate which are not typically imposed in applications. Let me give a few coordinate systems commonly used in applications so we can constrast those against the coordinate systems given from orthonormal bases of $\mathbb{R}^n$.

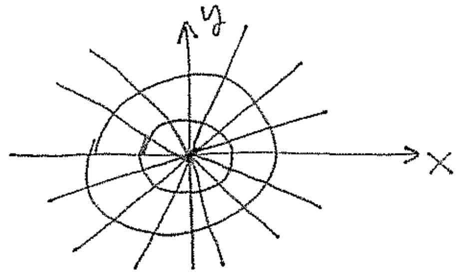
Example 2.6.4. . Cylindrical Coordinates

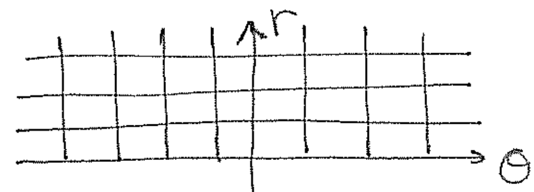
$$T : \mathbb{R}^3_{xyz} \longrightarrow \mathbb{R}^3_{\theta r z}$$

$T(x,y,z) = (\theta, r, z)$ where $\overbrace{\theta = \tan^{-1}(y/x), \quad r = \sqrt{x^2+y^2}}^{\text{(Sometimes)}}$



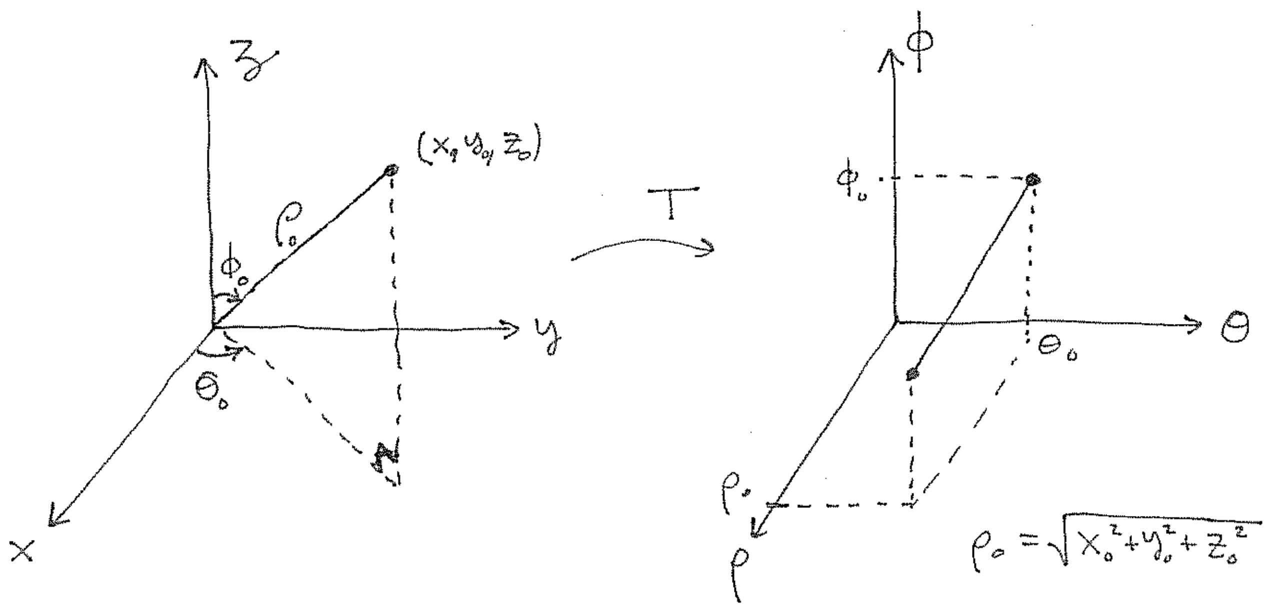
Example 2.6.5. *Consider $\mathbb{R}^2$ with the usual x, y coordinates,* **polar coordinates** *r, θ are given by the* **polar radius** *$r = \sqrt{x^2+y^2}$ and* **polar angle** *$\theta = \tan^{-1}(y/x)$. These are inverted to give $x = r\cos(\theta)$ and $y = r\sin(\theta)$. Notice that θ is not well defined along $x = 0$ if we take the given formula as the definition. Even so the angle at the origin is not well-defined no matter how you massage the equations. Polar coordinates are curvelinear coordinates, setting $\theta = 0$ yields a ray along the postive x-axis whereas setting $r = 0$ just yields the origin.*





note: in calculus II we allow $r<0$ for convenience of polar graphing.

Example 2.6.6. *Consider $\mathbb{R}^3$ with the usual x, y, z coordinates,* **spherical coordinates** *ρ, θ, ϕ are given by* **spherical radius** *$\rho = \sqrt{x^2+y^2+z^2}$,* **polar angle** *$\theta = \tan^{-1}(y/x)$ and* **azimuthial angle** *$\phi = \cos^{-1}(z/\sqrt{x^2+y^2+z^2})$. These are inverted to give $x = \rho\cos(\theta)\sin(\phi)$ and $y = \rho\sin(\theta)\sin(\phi)$ and $z = \rho\cos(\phi)$. Even so the angles can't be well-defined everywhere. The function of inverse tangent can never return a polar angle in quadrants* **II** *or* **III** *because range$(\tan^{-1}) = (-\pi/2, \pi/2)$. In order to find angles in the quadrants with $x < 0$ we have to adjust the equations by hand as we are taught in trigonmetry. Spherical coordinates are also curvelinear, there is no coordinate axis for the spherical radius and the angles have rays rather than lines for their coordinate axes.*



quick notation

$$T : \underbrace{\mathbb{R}^3_{xyz}}_{\text{tell us use } xyz \text{ variables for domain}} \longrightarrow \underbrace{\mathbb{R}^3_{\rho\theta\phi}}_{\text{tells us use } \rho\theta\phi \text{ for variables in range}}$$

Example 2.6.7. *Consider $\mathbb{R}^n$ with the usual Cartesian coordinates $x = (x_1, x_2, \dots, x_n)$. If $p \in \mathbb{R}^n$ then we can write*

$$p = x_1 e_1 + x_2 e_2 + \cdots + x_n e_n = [e_1|e_2|\cdots|e_n][x_1, x_2, \dots, x_n]^T$$

Let T be an orthogonal transformation and define a rotated basis f_i by $[f_1|\cdots|f_n] = [e_1|\cdots|e_n]R = R$ where $R \in SO(n)$. Since $R^T R = I$ it follows that $R^{-1} = R^T$ and so $[e_1|\cdots|e_n] = [f_1|\cdots|f_n]R^T$. Note that $p = [f_1|\cdots|f_n]R^T p$. However, the y-coordinates will satisfy $p = [f_1|\cdots|f_n]y$ where $y = [y_1, y_2, \dots, y_n]^T$. We deduce,

$$y = R^T x.$$

We find that if we set up a rotated coordinate system where the new basis is formed by rotating the standard basis by R then the new coordinates relate to the old coordinates by the inverse rotation $R^T = R^{-1}$.

Let me break down the example in the $n = 2$ case.

Example 2.6.8. *Let $\{e_1, e_2\}$ be the standard basis for $\mathbb{R}^2$. In invite the reader to check that $R(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \in SO(2)$. If our calculation is correct in the previous example the new coordinate axes should be obtained from the standard basis by the inverse transformation.*

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x\cos\theta + y\sin\theta \\ -x\sin\theta + y\cos\theta \end{bmatrix}$$

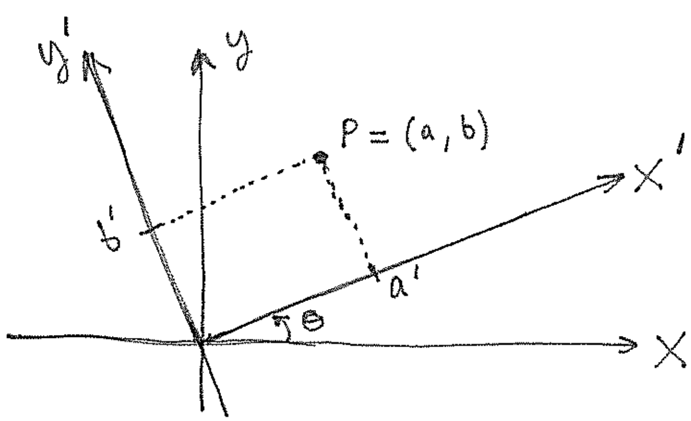
The inverse transformations to give x, y in terms of x', y' are similar

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} x'\cos\theta - y'\sin\theta \\ x'\sin\theta + y'\cos\theta \end{bmatrix}$$

Let's find the equations of the primed coordinate axes.

1. *The y' axis has equation $x' = 0$ hence $x = -y'\sin(\theta)$ and $y = y'\cos(\theta)$ which yields $y = -\cot(\theta)x$ for $y' \neq 0$*
2. *Likewise, the x' axis has equation $y' = 0$ hence $x = x'\cos(\theta)$ and $y = x'\sin(\theta)$ which yields $y = \tan(\theta)x$ for $x' \neq 0$.*

Therefore the new primed axes are perpendicular to one another and are apparently rotated by angle θ in the clockwise direction as illustrated below.



rotated coordinates
Could also
draw two
planes to
see transformation.

2.7 orthogonal complements

Perhaps you've seen part of this Theorem before:

Proposition 2.7.1. *Pythagorean Theorem in n-dimensions*

If $x, y \in \mathbb{R}^{n\times 1}$ are orthogonal vectors then $||x||^2+||y||^2 = ||x+y||^2$. Moreover, if $x_1, x_2, \dots x_k$ are orthogonal then

$$||x_1||^2 + ||x_2||^2 + \cdots + ||x_k||^2 = ||x_1 + x_2 + \cdots + x_k||^2$$

The notation $W \leq V$ is meant to read "W is a **subspace** of V". A subspace is a subset of a vector space which is again a vector space with respect to the operations of V

Proposition 2.7.2. Existence of Orthonormal Basis

If $W \leq \mathbb{R}^{n\times 1}$ then there exists an orthonormal basis of W

The proof of the proposition above relies on an algorthim called **Gram-Schmidt orthogonalization**. That algorithm allows you to take any set of linearly indepedent vectors and replace it with a new set of vectors which are pairwise orthogonal.

Example 2.7.3. *For the record, the standard basis of $\mathbb{R}^{n\times 1}$ is an orthonormal basis and*

$$v = (v \cdot e_1)e_1 + (v \cdot e_2)e_2 + \cdots + (v \cdot e_n)e_n$$

for any vector v in $\mathbb{R}^{n\times 1}$.

Definition 2.7.4.

Suppose $W_1, W_2 \subseteq \mathbb{R}^{n\times 1}$ 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 $W_1 \perp W_2$.

Definition 2.7.5.

Let V be a vector space and $W_1, W_2 \leq V$. If every $v \in V$ can be written as $v = w_1 + w_2$ for a unique pair of $w_1 \in W_1$ and $w_2 \in W_2$ then we say that V is the **direct sum** of W_1 and W_2. Moreover, we denote the statement "V is a direct sum of W_1 and W_2" by $V = W_1 \oplus W_2$.

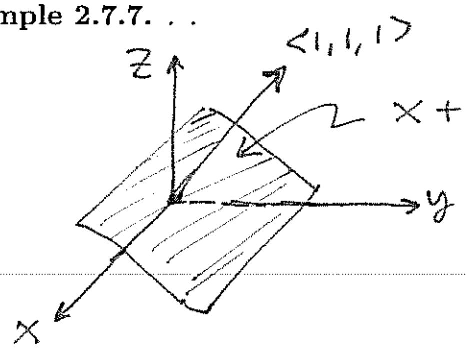
Proposition 2.7.6.

Let $W \leq \mathbb{R}^{n\times 1}$ then

1. $\mathbb{R}^{n\times 1} = W \oplus W^{\perp}$.
2. $dim(W) + dim(W^{\perp}) = n$,
3. $(W^{\perp})^{\perp} = W$,

Basically the cross-product is used in V^3 to select the perpendicular to a plane formed by two vectors. The theorem above tells us that if we wished to choose a perpendicular direction for a 2-dimensional plane inside V^5 then we would have a $5 - 2 = 3$-dimensional orthogonal complement to choose a "normal" for the plane. In other words, the concept of a normal vector to a plane is not so simple in higher dimensions. We could have a particular plane with two different "normal" vectors which were orthogonal!

Example 2.7.7. . . .



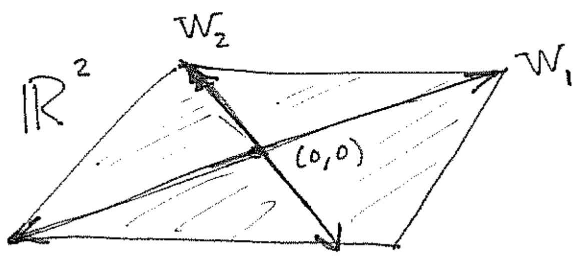
$x+y+z=0$ has normal $\langle 1,1,1\rangle$

$W_1 = \{(x,y,z) \mid x+y+z = 0\} =$ plane

$W_2 = \{t\langle 1,1,1\rangle \mid t \in \mathbb{R}\} =$ normal line.

$$\mathbb{R}^3 = W_1 \oplus W_2$$

Example 2.7.8. . . .



$W_1 = \text{span}\{\langle 1,1\rangle\}$

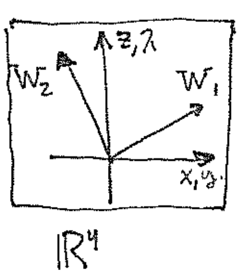
$W_2 = \text{span}\{\langle -1,1\rangle\}$

$\mathbb{R}^2 = W_1 \oplus W_2$

Example 2.7.9. . . .

Let $W_1 = \text{span}\{(1,1,1,1), (1,0,0,0)\}$

$\Rightarrow W_1 = \{s(1,1,1,1) + t(1,0,0,0) \mid s,t \in \mathbb{R}\}$



This is a plane in 4-dim'l space. Let's find its orthogonal complement.

$$W_2 = \{v \in \mathbb{R}^4 \mid v \cdot w = 0 \ \ \forall w \in W_1\}$$

need $v \cdot (1,1,1,1) = 0$ & $v \cdot (1,0,0,0) = 0$

Let $v = (x,y,z,\lambda)$ need $x+y+z+\lambda = 0$ & $x = 0$

Hence $y+z+\lambda = 0$ for $v = (x,y,z,\lambda) \in W_2$.

Thus $\underbrace{W_2 = \{(0,y,z,-y-z) \mid y,z \in \mathbb{R}\}}_{\text{this is normal to } W_1}$, $\underline{W_1 \oplus W_2 = \mathbb{R}^4}$.

Chapter 3

topology and mappings

We begin this chapter by briefly examining all the major concepts of the metric topology for $\mathbb{R}^n$. Then we discuss limits for functions and mappings from using the rigorous $\epsilon - \delta$ formulation. For this chapter and course a "function" has range which is a subset of $\mathbb{R}$. In contrast, a mapping has a range which is in some subset of $\mathbb{R}^n$ for $n \geq 2$ if we want to make it interesting[1]. Continuity is defined and a number of basic theorems are either proved by me or you. Finally I quote a few important (and less trivial) theorems about topology and mappings in $\mathbb{R}^n$.

3.1 functions and mappings

In this section we disucss basic vocabulary for functions and mappings.

Definition 3.1.1.

Let $U \subseteq \mathbb{R}^n$ and $V \subseteq \mathbb{R}$ then we say that $f : U \to V$ is a **function** iff $f(x)$ assigns a single value in V for each input $x \in U$. We say a function is **single-valued** from **domain** U to **codomain** V. We denote $dom(f) = U$. The **range** or **image** of the function is defined by:

$$range(f) = f(D) = \{y \in \mathbb{R} \mid \exists x \in U \text{ such that } f(x) = y\}$$

We can also say that "f is a real-valued function of U".

Example 3.1.2. . .

① $f(x) = x^2$ for $x \in \mathbb{R}$, $dom(f) = \mathbb{R}$

② $g(\vec{x}) = \vec{x} \cdot \vec{x} = x_1^2 + x_2^2 + \cdots + x_n^2$ for $\vec{x} \in \mathbb{R}^n$, $dom(g) = \mathbb{R}^n$

[1] I generally prefer the term function for a more abstract concept: I would like to say $f : A \to B$ is an B-valued function of A and I don't make any restriction except that A, B must be sets. Anyhow, I'll try to respect the custom of calculus for this course because it saves us a lot of talking. I will use the term "abstract function" if I don't wish to presuppose the codmain contains only real numbers.

A mapping is an abstract function with codmain in $\mathbb{R}^n$

Definition 3.1.3.

Let $U \subset \mathbb{R}^n$ and $V \subset \mathbb{R}^m$ then we say that $f : U \to V$ is a **mapping** iff $f(x)$ assigns a single value in V for each input $x \in U$. We say a f is a single-value mapping from **domain** U to **codomain** V. We mean for $dom(f) = U$ to be read that the domain of f is U. The **range** or **image** of the mapping is the set of all possible outputs: we denote

$$range(f) = f(D) = \{y \in \mathbb{R}^m \mid \exists x \in U \text{ such that } f(x) = y\}$$

Suppose that $x =\in dom(f)$ and $f(x) = (f_1(x), f_2(x), \ldots, f_m(x))$ then we say that $f_1, f_2, \ldots, f_m$ are the **component functions** of f and $f = (f_i) = (f_1, f_2, \ldots f_m)$.

In the case $m = 1$ we find that the concept of a mapping reduces to a plain-old-function.

Example 3.1.4. . .

$\Phi : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ where $\Phi(r, \theta) = \langle r\cos\theta, r\sin\theta \rangle$

we have component functions

$\Phi_1(r,\theta) = r\cos\theta$ & $\Phi_2(r,\theta) = r\sin\theta$.

Definition 3.1.5.

A mapping $f : U \subseteq \mathbb{R}^n \to V \subseteq \mathbb{R}^m$ is said to be **injective** or **1-1** on $S \subseteq U$ iff $f(x) = f(y)$ implies $x = y$ for all $x, y \in S$. If a mapping is 1-1 on its domain then it is said to be 1-1 or injective. The mapping f is said to be **surjective** or **onto** $T \subseteq V$ iff for each $v \in T$ there exists $u \in U$ such that $f(u) = v$; in set notation we can express: f is onto T iff $f(U) = T$. A mapping is said to be surjective or onto iff it is onto its codomain. A mapping is a **bijection** or **1-1 correspondance** of U and V iff f is injective and surjective.

Example 3.1.6. . .

① $f(x,y) = (e^x, y^2, 2)$ is onto $(0,\infty) \times [0,\infty) \times \{2\}$ given that $dom(f) = \mathbb{R}^2$. To prove this let $(a,b,c) \in (0,\infty) \times [0,\infty) \times \{2\}$ and observe that $\ln(a), \sqrt{b} \in \mathbb{R}$ since $a > 0$ and $b \geq 0$ hence

$$f(\ln(a), \sqrt{b}, 2) = (e^{\ln(a)}, (\sqrt{b})^2, 2) = (a, b, 2) = (a,b,c).$$

② $g : \mathbb{R} \longrightarrow (0,\infty)$ defined by $g(x) = e^x$ is a <u>bijection</u>.

③ $h : \mathbb{R} \longrightarrow \mathbb{R}$ defined by $h(x) = x^2$ is neither 1-1 nor onto.

We can also adjust the domain of a given mapping by restriction and extension.

Definition 3.1.7.

Let $f : U \subseteq \mathbb{R}^n \to V \subseteq \mathbb{R}^m$ be a mapping. If $R \subset U$ then we define the **restriction of** f **to** R to be the mapping $f|_R : R \to V$ where $f|_R(x) = f(x)$ for all $x \in R$. If $U \subseteq S$ and $V \subset T$ then we say a mapping $g : S \to T$ is an **extension of** f iff $g|_{dom(f)} = f$.

When I say $g|_{dom(f)} = f$ this means that these functions have matching domains and they agree at each point in that domain; $g|_{dom(f)}(x) = f(x)$ for all $x \in dom(f)$. Once a particular subset is chosen the restriction to that subset is a unique function. Of course there are usually many susbets of $dom(f$ so you can imagine many different restictions of a given function. The concept of extension is more vague, once you pick the enlarged domain and codomain it is not even necessarily the case that another extension to that same pair of sets will be the same mapping. To obtain uniqueness for extensions one needs to add more stucture. This is one reason that complex variables are interesting, there are cases where the structure of the complex theory forces the extension of a complex-valued function of a complex variable to be unique. This is very surprising.

Example 3.1.8. . .

① Let $\underset{dom(f) = \mathbb{R}}{f(x) = \sqrt{x^2}}$ then $f|_{[0,\infty)}(x) = x$ whereas $f|_{(-\infty,0]}(x) = -x$.

② Let $f(x) = \ln(x)$ for $x \in (0,\infty)$. If $g(x) = \ln|x|$ for $x \in \mathbb{R} - \{0\}$ then $g|_{(0,\infty)} = f$ so g is an extension of f.

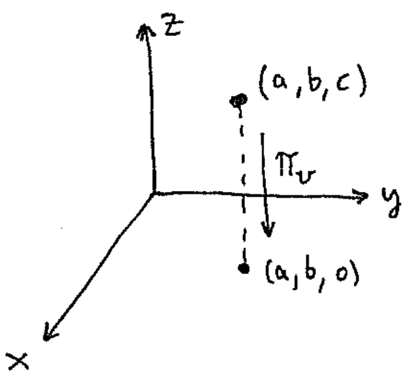
Definition 3.1.9.

Let $\pi_U : \mathbb{R}^n \to U \subseteq \mathbb{R}^n$ be a mapping such that $\pi_U(x) = x$ for all $x \in U$. We say that pi_U is a **projection** onto U. The **identity mapping** on $U \subseteq \mathbb{R}^n$ is defined by $Id_U : U \to U$ with $Id_U(x) = x$ for all $x \in U$. We may also denote $Id_{\mathbb{R}^n} = Id_n = Id$ where convenient. The j**-th projection function** is $\pi_j : \mathbb{R}^n \to \mathbb{R}$ defined by $\pi_j(x_1, x_2, \dots, x_n) = x_j$

Notice that every identity map is a projection however not every projection is an identity.

Example 3.1.10. . .

Let $U = \mathbb{R}^2 \times \{0\} = \{ (x, y, 0) \mid x, y \in \mathbb{R} \}$, this is the xy-plane.



$\pi_U(a, b, c) = (a, b, 0)$.

Definition 3.1.11.

Let $f : V \subseteq \mathbb{R}^n \to W \subseteq \mathbb{R}^m$ and $g : U \subseteq \mathbb{R}^n \to V \subseteq \mathbb{R}^m$ are mappings such that $g(U) \subseteq dom(f)$ then $f \circ g : U \to W$ is a mapping defined by $(f \circ g)(x) = f(g(x))$ for all $x \in U$. We say f is the outside function and g is the inside function.

Notice that the definition of the composite assumes that the range of the inside function fits nicely in the domain of the outside function. If domains are not explicitly given then it is customary to choose the domain of the composite of two functions to be as large as possible. Indeed, the typical pattern in calculus is that the domain is implicitly indicated by some formula. For example, $g(x) = e^x \frac{x-4}{x-4}$ has implied domain $dom(g) = (-\infty, 4) \cup (4, \infty)$ however if we simply the formula to give $g(x) = e^x$ then the implied domain of $\mathbb{R}$ is not correct. Of course we can not make that simplification unless $x \neq 4$. In short, when we do algebra for variables we should be careful to consider the values which the variables may assume. Often one needs to break a calculation into cases to avoid division by zero.

Example 3.1.12. . .

Let $\Phi : (0,\infty) \times (-\pi/2, \pi/2) \to \mathbb{R}^2$ be defined by $\Phi(r,\theta) = (r\cos\theta, r\sin\theta)$.

and $F : (0,\infty) \times \mathbb{R} \longrightarrow \mathbb{R}^2$ be defined by $F(x,y) = \left(\sqrt{x^2+y^2}, \tan^{-1}(y/x)\right)$.

Let $(r,\theta) \in dom(\Phi)$, notice that $\underbrace{\cos(-\pi/2,\pi/2) = (0,1) \text{ whereas } \sin(-\pi/2,\pi/2) = (-1,1)}$,

$$
\begin{aligned}
(F \circ \Phi)(r,\theta) &= F(\Phi(r,\theta)) \\
&= F(r\cos\theta, r\sin\theta) \quad \longleftarrow \text{need these to be sure } \Phi(r,\theta) \text{ is in } dom(F) \\
&= \left(\sqrt{r^2\cos^2\theta + r^2\sin^2\theta}, \tan^{-1}\left(\frac{r\sin\theta}{r\cos\theta}\right)\right) \\
&= \left(\sqrt{r^2}, \tan^{-1}(\tan\theta)\right) \\
&= (r,\theta)
\end{aligned}
$$

Definition 3.1.13.

Let $f : U \subseteq \mathbb{R}^n \to V \subseteq \mathbb{R}^m$ be a mapping, if there exists a mapping $g : f(U) \to U$ such that $f \circ g = Id_{f(U)}$ and $g \circ f = Id_U$ then g is the inverse mapping of f and we denote $g = f^{-1}$.

If a mapping is injective then it can be shown that the inverse mapping is well defined. We define $f^{-1}(y) = x$ iff $f(x) = y$ and the value x must be a single value if the function is one-one. When a function is not one-one then there may be more than one point which maps to a particular point in the range.

Example 3.1.14. . .

$\Phi : (0,\infty) \times (-\pi/2, \pi/2) \longrightarrow (0,\infty) \times \mathbb{R}$ be defined by $\Phi(r,\theta) = (r\cos\theta, r\sin\theta)$. We can show Φ is injective and onto $(0,\infty) \times \mathbb{R}$ thus there exists $\Phi^{-1} : (0,\infty) \times \mathbb{R} \longrightarrow dom(\Phi)$. In particular,

$$\Phi^{-1}(x,y) = \left(\sqrt{x^2+y^2}, \tan^{-1}(y/x)\right)$$

Note the Ex 3.1.12 shows $\Phi \circ \Phi^{-1} = Id\big|_{(0,\infty)\times\mathbb{R}}$.

Definition 3.1.15.

Let $f : U \subseteq \mathbb{R}^n \to V \subseteq \mathbb{R}^m$ be a mapping. We define a **fiber** of f over $y \in range(f)$ as

$$f^{-1}\{y\} = \{x \in U | f(x) = y\}$$

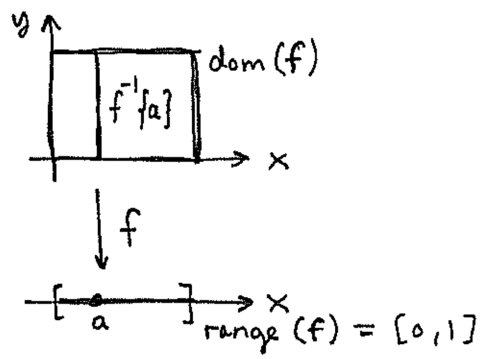
Notice that the inverse image of a set is well-defined even if there is no inverse mapping. Moreover, it can be shown that the fibers of a mapping are disjoint and their union covers the domain of the mapping:

$$f(y) \neq f(z) \quad \Rightarrow \quad f^{-1}\{y\} \cap f^{-1}\{z\} = \emptyset \qquad\qquad \bigcup_{y \, \in \, range(f)} f^{-1}\{y\} = dom(f).$$

This means that the fibers of a mapping *partition* the domain.

Example 3.1.16. . .

Let $f(x,y) = x$ for all $(x,y) \in dom(f) = [0,1] \times [0,1]$



Definition 3.1.17.

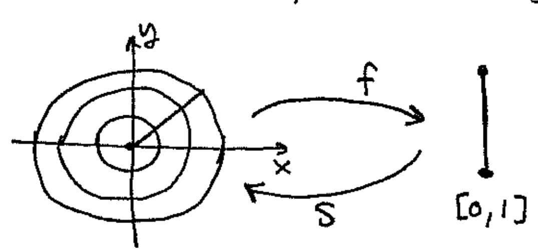
Let $f : U \subseteq \mathbb{R}^n \to V \subseteq \mathbb{R}^m$ be a mapping. Furthermore, suppose that $s : \mathbf{V} \to U$ is a mapping which is constant on each fiber of f. In other words, for each fiber $f^{-1}\{y\} \subseteq U$ we have some constant $u \in U$ such that $s(f^{-1}\{y\}) = u$. The subset $s^{-1}(U) \subseteq U$ is called a **cross section** of the fiber partition of f.

How do we construct a cross section for a particular mapping? For particular examples the details of the formula for the mapping usually suggests some obvious choice. However, in general if you accept the **axiom of choice** then you can be comforted in the existence of a cross section even in the case that there are infinitely many fibers for the mapping.

Example 3.1.18. . .

Let $f(x,y) = x^2+y^2$ for all $(x,y) \in \mathbb{R}^2$ such that $x^2+y^2 \leq 1$.

The fibers are circles, and the origin $f^{-1}\{0\} = (0,0)$. Let us



define a <u>section</u> by

$s(r) = (\sqrt{r}/\sqrt{2}, \sqrt{r}/\sqrt{2})$

for each $r \in [0,1]$. Notice

$s[0,1] = $ ray at $\theta = \pi/4$.

Notice $f(s(r)) = f(r/\sqrt{2}, r/\sqrt{2}) = \sqrt{\left(\frac{r}{\sqrt{2}}\right)^2 + \left(\frac{r}{\sqrt{2}}\right)^2} = \sqrt{r^2} = |r| = r.$

Note $f|_{s[0,1]}$ is injective because each fiber is reduced to a point.

Proposition 3.1.19.

Let $f : U \subseteq \mathbb{R}^n \to V \subseteq \mathbb{R}^m$ be a mapping. The restriction of f to a cross section S of U is an injective function. The mapping $\tilde{f} : U \to f(U)$ is a surjection. The mapping $\tilde{f}|_S : S \to f(U)$ is a bijection.

The proposition above tells us that we can take any mapping and cut down the domain and/or codomain to reduce the function to an injection, surjection or bijection. If you look for it you'll see this result behind the scenes in other courses. For example, in linear algebra if we throw out the kernel of a linear mapping then we get an injection. The idea of a local inverse is also important to the study of calculus.

Example 3.1.20. . . . Let $f : [0,1]\times[0,1] \longrightarrow \mathbb{R}$ be defined by $f(x,y) = x$ is not onto $\mathbb{R}$, and it's not injective since $f^{-1}\{x\} = \{x\}\times[0,1]$. You can check $s : [0,1] \longrightarrow dom(f)$ with $s(x) = (x, 1/2)$ is a section of f. Moreover, $\tilde{f} : [0,1]\times[0,1] \longrightarrow [0,1]$ is onto and $\tilde{f}|_{s[0,1]}$ is a bijection.

Definition 3.1.21.

Let $f : U \subseteq \mathbb{R}^n \to V \subseteq \mathbb{R}^m$ be a mapping then we say a mapping g is a **local inverse** of f iff there exits $S \subseteq U$ such that $g = (f|_S)^{-1}$.

Usually we can find local inverses for functions in calculus. For example, $f(x) = \sin(x)$ is not 1-1 therefore it is not invertible. However, it does have a local inverse $g(y) = \sin^{-1}(y)$. If we were more pedantic we wouldn't write $\sin^{-1}(y)$. Instead we would write $g(y) = \left(\sin|_{[\frac{-\pi}{2},\frac{\pi}{2}]}\right)^{-1}(y)$ since the inverse sine is actually just a local inverse. To construct a local inverse for some mapping we must locate some subset of the domain upon which the mapping is injective. Then relative to that subset we can reverse the mapping. The inverse mapping theorem (which we'll study mid-course) will tell us more about the existence of local inverses for a given mapping.

Definition 3.1.22.

Let $f : U_1 \subseteq \mathbb{R}^n \to V_1 \subseteq \mathbb{R}^p$ and $g : U_1 \subseteq \mathbb{R}^n \to V_2 \subseteq \mathbb{R}^q$ be a mappings then (f,g) is a mapping from U_1 to $V_1 \times V_2$ defined by $(f,g)(x) = (f(x), g(x))$ for all $x \in U_1$.

There's more than meets the eye in the definition above. Let me expand it a bit here:

$$(f,g)(x) = (f_1(x), f_2(x), \dots, f_p(x), g_1(x), g_2(x), \dots, g_q(x)) \quad \text{where} \quad x = (x_1, x_2, \dots, x_n)$$

You might notice that Edwards uses π for the identity mapping whereas I use Id. His notation is quite reasonable given that the identity is the cartesian product of all the projection maps:

$$\pi = (\pi_1, \pi_2, \dots, \pi_n)$$

I've had courses where we simply used the coordinate notation itself for projections, in that notation have formulas such as $x(a,b,c) = a$, $x_j(a) = a_j$ and $x_j(e_i) = \delta_{ji}$.

Example 3.1.23. . . .

① $f : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ be defined by $f(x,y,z) = (x, yz^2)$

$g : \mathbb{R}^3 \longrightarrow \mathbb{R}$ be defined by $g(x,y,z) = \|(x,y,z)\|$

$(f, g) : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ has $(f, g)(\vec{x}) = (f(\vec{x}), g(\vec{x}))$

② $\pi_{xy} : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ be defined by $\pi_{xy}(x,y,z) = (x,y)$

$\pi_z : \mathbb{R}^3 \longrightarrow \mathbb{R}$ be defined by $\pi_z(x,y,z) = z$

You can see $(\pi_{xy}, \pi_z) = Id_{\mathbb{R}^3}$.

The constructions thus far in this section have not relied on the particular properties of real vectors. If you look at the definitions they really only depend on an understanding of sets, points and subsets. In contrast, the definition given below defines the sum of two mappings, the scalar product of a mapping and a constant or a function, and the dot-product of two mappings.

Definition 3.1.24.

Let $f, g : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ be a mappings and $c \in \mathbb{R}$ and $h : U \to \mathbb{R}$ a function. We define:

1. $f + g$ is a mapping from U to $\mathbb{R}^m$ where $(f+g)(x) = f(x) + g(x)$ for all $x \in U$.
2. hf is a mapping from U to $\mathbb{R}^m$ where $(hf)(x) = h(x)f(x)$ for all $x \in U$.
3. cf is a mapping from U to $\mathbb{R}^m$ where $(cf)(x) = cf(x)$ for all $x \in U$.
4. $f \cdot g$ is a function of U where $(f \cdot g)(x) = f(x) \cdot g(x)$ for all $x \in U$.

We cannot hope to define the product and quotient of mappings to be another new mapping because we do not know how to define the product or quotient of vectors for arbitrary dimensions. In contrast, we can define the product of matrix-valued maps or of complex-valued maps because we have a way to multiply matrices and complex numbers. If the range of a function allows for some type of product it generally makes sense to define a corresponding operation on functions which map into that range.

Definition 3.1.25.

Let $f, g : U \subseteq \mathbb{R}^n \to \mathbb{C}$ be complex-valued functions. We define:

1. fg is a complex-valued function defined by $(fg)(x) = f(x)g(x)$ for all $x \in U$.
2. If $0 \notin g(U)$ then f/g is a complex-valued function defined by $(f/g)(x) = f(x)/g(x)$ for all $x \in U$.

Example 3.1.26. . .

$f(\theta) = e^{i\theta} = \cos\theta + i\sin\theta$ defines $f : \mathbb{R} \to \mathbb{C}$

$g(\theta) = 3 + i\theta$

$$(fg)(\theta) = f(\theta)g(\theta) = (\cos\theta + i\sin\theta)(3 + i\theta)$$
$$= 3\cos\theta + 3i\sin\theta + i\,\theta\cos\theta - \theta\sin\theta$$
$$= 3\cos\theta - \theta\cos\theta + i\,(3\sin\theta + \theta\cos\theta).$$

Definition 3.1.27.

Let $A, B : U \subseteq \mathbb{R} \to \mathbb{R}^{m \times n}$ and $X : U \subseteq \mathbb{R} \to \mathbb{R}^{n \times p}$ be matrix-valued functions and $f : U \subseteq \mathbb{R} \to \mathbb{R}$. We define:

1. $A + B$ is a matrix-valued function defined by $(A + B)(x) = A(x) + B(x)$ for all $x \in U$.
2. AX is a matrix-valued function defined by $(AX)(x) = A(x)B(x)$ for all $x \in U$.
3. fA is a matrix-valued function defined by $(fA)(x) = f(x)A(x)$ for all $x \in U$.

The calculus of matrices is important to physics and differential equations.

Example 3.1.28. . .

Let $A, B : \mathbb{R} \to \mathbb{R}^{2\times 2}$ be defined by $A(t) = \begin{bmatrix} 1 & 2 \\ t & t^2 \end{bmatrix}$ & $B(t) = \begin{bmatrix} e^t & t \\ t^2 & t^3 \end{bmatrix}$

$$(A+B)(t) = A(t) + B(t) = \begin{bmatrix} 1 & 2 \\ t & t^2 \end{bmatrix} + \begin{bmatrix} e^t & t \\ t^2 & t^3 \end{bmatrix} = \begin{bmatrix} 1+e^t & 2+t \\ t+t^2 & t^2+t^3 \end{bmatrix}.$$

$$(AB)(t) = A(t)B(t) = \begin{bmatrix} 1 & 2 \\ t & t^2 \end{bmatrix}\begin{bmatrix} e^t & t \\ t^2 & t^3 \end{bmatrix} = \begin{bmatrix} e^t + 2t^2 & t + 2t^3 \\ te^t + t^4 & t^2 + t^5 \end{bmatrix}.$$

Let $f(t) = \sin t$,

$$(fA)(t) = f(t)A(t) = \sin t \begin{bmatrix} 1 & 2 \\ t & t^2 \end{bmatrix} = \begin{bmatrix} \sin t & 2\sin t \\ t\sin t & t^2\sin t \end{bmatrix}.$$

3.2 elementary topology and limits

In this section we describe the *metric topology* for $\mathbb{R}^n$. In the study of functions of one real variable we often need to refer to open or closed intervals. The definition that follows generalizes those concepts to n-dimensions. I have included a short discussion of general topology in the Appendix if you'd like to learn more about the term.

Definition 3.2.1.

An **open ball** of radius ϵ centered at $a \in \mathbb{R}^n$ is the subset all points in $\mathbb{R}^n$ which are less than ϵ units from a, we denote this open ball by

$$B_\epsilon(a) = \{x \in \mathbb{R}^n \mid ||x - a|| < \epsilon\}$$

The **closed ball** of radius ϵ centered at $a \in \mathbb{R}^n$ is likewise defined

$$\overline{B}_\epsilon(a) = \{x \in \mathbb{R}^n \mid ||x - a|| \leq \epsilon\}$$

Notice that in the $n = 1$ case we observe an open ball is an open interval: let $a \in \mathbb{R}$,

$$B_\epsilon(a) = \{x \in \mathbb{R} \mid ||x - a|| < \epsilon\} = \{x \in \mathbb{R} \mid |x - a| < \epsilon\} = (a - \epsilon, a + \epsilon)$$

In the $n = 2$ case we observe that an open ball is an open disk: let $(a, b) \in \mathbb{R}^2$,

$$B_\epsilon((a,b)) = \{(x,y) \in \mathbb{R}^2 \mid ||\,(x,y) - (a,b)\,|| < \epsilon\} = \{(x,y) \in \mathbb{R}^2 \mid \sqrt{(x-a)^2 + (y-b)^2} < \epsilon\}$$

For $n = 3$ an open-ball is a sphere without the outer shell. In contrast, a closed ball in $n = 3$ is a solid sphere which includes the outer shell of the sphere.

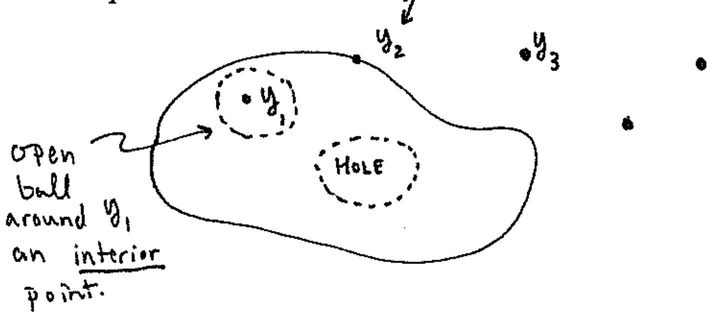
Definition 3.2.2.

Let $D \subseteq \mathbb{R}^n$. We say $y \in D$ is an **interior point** of D iff there exists some open ball centered at y which is completely contained in D. We say $y \in \mathbb{R}^n$ is a **limit point** of D iff every open ball centered at y contains points in $D - \{y\}$. We say $y \in \mathbb{R}^n$ is a **boundary point** of D iff every open ball centered at y contains points not in D and other points which are in $D - \{y\}$. We say $y \in D$ is an **isolated point** of D if there exist open balls about y which do not contain other points in D. The set of all interior points of D is called the **interior of** D. Likewise the set of all boundary points for D is denoted ∂D. The **closure** of D is defined to be $\overline{D} = D \cup \{y \in \mathbb{R}^n \mid y \text{ a limit point}\}$

If you're like me the paragraph above doesn't help much until I see the picture below. All the terms are aptly named. The term "limit point" is given because those points are the ones for which it is natural to define a limit.

y_2 is not an interior pt.

Example 3.2.3. . .



① y_1 & y_2 are limit points of pictured subset, $D \subset \mathbb{R}^2$

② y_3 is isolated point

③ $y_2 \in \partial D$, y_2 is boundary pt.

④ $\overline{D}$ formed by connecting dots in interior hole.

Definition 3.2.4.

Let $A \subseteq \mathbb{R}^n$ is an **open set** iff for each $x \in A$ there exists $\epsilon > 0$ such that $x \in B_\epsilon(x)$ and $B_\epsilon(x) \subseteq A$. Let $B \subseteq \mathbb{R}^n$ is an **closed set** iff its complement $\mathbb{R}^n - B = \{x \in \mathbb{R}^n \mid x \notin B\}$ is an open set.

Notice that $\mathbb{R} - [a, b] = (\infty, a) \cup (b, \infty)$. It is not hard to prove that open intervals are open hence we find that a closed interval is a closed set. Likewise it is not hard to prove that open balls are open sets and closed balls are closed sets. I may ask you to prove the following proposition in the homework.

Proposition 3.2.5.

A closed set contains all its limit points, that is $A \subseteq \mathbb{R}^n$ is closed iff $A = \overline{A}$.

Example 3.2.6. . .

① (a,b) has limit points $x=a$ & $x=b$. However, $a,b \notin (a,b)$ and we note that (a,b) is <u>not</u> a closed set.

② $\overline{(a,b)} = [a,b]$ which is a closed set. Closed intervals are closed sets.

In calculus I the limit of a function is defined in terms of deleted open intervals centered about the limit point. We can define the limit of a mapping in terms of deleted open balls centered at the limit point.

Definition 3.2.7.

Let $f : U \subseteq \mathbb{R}^n \to V \subseteq \mathbb{R}^m$ be a mapping. We say that f has limit $b \in \mathbb{R}^m$ at limit point a of U iff for each $\epsilon > 0$ there exists a $\delta > 0$ such that $x \in \mathbb{R}^n$ with $0 < ||x - a|| < \delta$ implies $||f(x) - b|| < \epsilon$. In such a case we can denote the above by stating that

$$\lim_{x \to a} f(x) = b$$

In calculus I the limit of a function is defined in terms of deleted open intervals centered about the limit point. We now define the limit of a mapping in terms of deleted open balls centered at the limit point. The term "deleted" refers to the fact that we assume $0 < ||x - a||$ which means we do not consider $x = a$ in the limiting process. In other words, the limit of a mapping considers values close to the limit point but not necessarily the limit point itself. The case that the function is defined at the limit point is special, when the limit and the mapping agree then we say the mapping is continuous at that point.

Example 3.2.8. . .

In calculus I we prove that elementary functions are continuous on the interior of their domains. For example $f(x) = e^x$, $\cos(x)$, $\sin(x)$, $P(x)$ (polynomial) are continuous on $\mathbb{R}$ whereas $r(x) = \frac{P(x)}{q(x)}$ is continuous for $x \in \mathbb{R}$ such that $q(x) \neq 0$.

Ok, to be honest we don't <u>prove</u> all these things, hopefully we at least show the students how they might try to prove these assertions...

Definition 3.2.9.

Let $f : U \subseteq \mathbb{R}^n \to V \subseteq \mathbb{R}^m$ be a mapping. If $a \in U$ is a limit point of f then we say that f is **continuous at** a iff

$$\lim_{x \to a} f(x) = f(a)$$

If $a \in U$ is an isolated point then we also say that f is continous at a. The mapping f is **continous on** S iff it is continous at each point in S. The **mapping** f **is continuous** iff it is continuous on its domain.

Notice that in the $m = n = 1$ case we recover the definition of continuous functions from calc. I.

Proposition 3.2.10.

Let $f : U \subseteq \mathbb{R}^n \to V \subseteq \mathbb{R}^m$ be a mapping with component functions $f_1, f_2, \dots, f_m$ hence $f = (f_1, f_2, \dots, f_m)$. If $a \in U$ is a limit point of f then

$$\lim_{x \to a} f(x) = b \quad \Leftrightarrow \quad \lim_{x \to a} f_j(x) = b_j \text{ for each } j = 1, 2, \dots, m.$$

.

We can analyze the limit of a mapping by analyzing the limits of the component functions:

Example 3.2.11. . .

Let $f(x) = \left(\sqrt{x^2}, \sin(x), \frac{\sin x}{x}\right)$ thus $f = (f_1, f_2, f_3)$

where $f_1(x) = \sqrt{x^2}$, $f_2(x) = \sin(x)$, $f_3(x) = \frac{\sin x}{x}$ for $x \in \mathbb{R} - \{0\}$.

$$\left.\begin{array}{l} \lim_{x \to 0} f_1(x) = \sqrt{0^2} = 0 \\ \lim_{x \to 0} (\sin(x)) = 0 \\ \lim_{x \to 0} \left(\frac{\sin x}{x}\right) = 1 \end{array}\right\} \lim_{x \to 0} \left(\sqrt{x^2}, \sin x, \frac{\sin x}{x}\right) = (0, 0, 1).$$

The following follows immediately from the preceding proposition.

Proposition 3.2.12.

Suppose that $f : U \subseteq \mathbb{R}^n \to V \subseteq \mathbb{R}^m$ is a mapping with component functions $f_1, f_2, \dots, f_m$. Let $a \in U$ be a limit point of f then f is continous at a iff f_j is continuous at a for $j = 1, 2, \dots, m$. Moreover, f is continuous on S iff all the component functions of f are continuous on S. Finally, a mapping f is continous iff all of its component functions are continuous. .

The proof of the proposition is in Edwards, it's his Theorem 7.2. It's about time I proved something.

Proposition 3.2.13.

> The projection functions are continuous. The identity mapping is continuous.

Proof: Let $\epsilon > 0$ and choose $\delta = \epsilon$. If $x \in \mathbb{R}^n$ such that $0 < ||x - a|| < \delta$ then it follows that $||x - a|| < \epsilon$.. Therefore, $\lim_{x \to a} x = a$ which means that $\lim_{x \to a} Id(x) = Id(a)$ for all $a \in \mathbb{R}^n$. Hence Id is continuous on $\mathbb{R}^n$ which means Id is continuous. Since the projection functions are component functions of the identity mapping it follows that the projection functions are also continuous (using the previous proposition). □

Definition 3.2.14.

> 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 3.2.15.

> The sum and product functions are continuous.

Preparing for the proof: Let the limit point be (a, b). Consider what we wish to show: given a point (x, y) such that $0 < ||(x, y) - (a, b)|| < \delta$ we wish to show that

$$|s(x, y) - (a + b)| < \epsilon \qquad \text{or for the product} \qquad |p(x, y) - (ab)| < \epsilon$$

follow for appropriate choices of δ. Think about the sum for a moment,

$$|s(x, y) - (a + b)| = |x + y - a - b| \leq |x - a| + |y - b|$$

I just used the triangle inequality for the absolute value of real numbers. We see that if we could somehow get control of $|x - a|$ and $|y - b|$ then we'd be getting closer to the prize. We have control of $0 < ||(x, y) - (a, b)|| < \delta$ notice this reduces to

$$||(x - a, y - b)|| < \delta \quad \Rightarrow \quad \sqrt{(x - a)^2 + (y - b)^2} < \delta$$

it is clear that $(x - a)^2 < \delta^2$ since if it was otherwise the inequality above would be violated as adding a nonegative quantity $(y - b)^2$ only increases the radicand resulting in the squareroot to be larger than δ. Hence we may assume $(x - a)^2 < \delta^2$ and since $\delta > 0$ it follows $\boxed{|x - a| < \delta}$. Likewise, $\boxed{|y - b| < \delta}$. Thus

$$|s(x, y) - (a + b)| = |x + y - a - b| < |x - a| + |y - b| < 2\delta$$

We see for the sum proof we can choose $\delta = \epsilon/2$ and it will work out nicely.

Proof: Let $\epsilon > 0$ and let $(a, b) \in \mathbb{R}^2$. Choose $\delta = \epsilon/2$ and suppose $(x, y) \in \mathbb{R}^2$ such that $||(x, y) - (a, b)|| < \delta$. Observe that

$$||(x, y) - (a, b)|| < \delta \Rightarrow ||(x - a, y - b)||^2 < \delta^2 \Rightarrow |x - a|^2 + |y - b|^2 < \delta^2.$$

It follows $|x - a| < \delta$ and $|y - b| < \delta$. Thus

$$|s(x, y) - (a + b)| = |x + y - a - b| \leq |x - a| + |y - b| < \delta + \delta = 2\delta = \epsilon.$$

Therefore, $\lim_{(x,y)\to(a,b)} s(x, y) = a + b$. and it follows that the sum function if continuous at (a, b). But, (a, b) is an arbitrary point thus s is continuous on $\mathbb{R}^2$ hence the sum function is continuous. $\square$.

Preparing for the proof of continuity of the product function: I'll continue to use the same notation as above. We need to study $|p(x, y) - (ab)| = |xy - ab| < \epsilon$. Consider that

$$|xy - ab| = |xy - ya + ya - ab| = |y(x - a) + a(y - b)| \leq |y||x - a| + |a||y - b|$$

We know that $|x-a| < \delta$ and $|y-b| < \delta$. There is one less obvious factor to bound in the expression. What should we do about $|y|$?. I leave it to the reader to show that:

$$\boxed{|y - b| < \delta \qquad \Rightarrow \qquad |y| < |b| + \delta}$$

Now put it all together and hopefully we'll be able to "solve" for ϵ.

$$|xy - ab| =\leq |y||x - a| + |a||y - b| < (|b| + \delta)\delta + |a|\delta = \delta^2 + \delta(|a| + |b|) \text{ "} = \text{" } \epsilon$$

I put solve in quotes because we have considerably more freedom in our quest for finding δ. We could just as well find δ which makes the " = " become an <. That said let's pursue equality,

$$\delta^2 + \delta(|a| + |b|) - \epsilon = 0 \qquad \delta = \frac{-|a| - |b| \pm \sqrt{(|a| + |b|)^2 + 4\epsilon}}{2}$$

Since $\epsilon, |a|, |b| > 0$ it follows that $\sqrt{(|a| + |b|)^2 + 4\epsilon} < \sqrt{(|a| + |b|)^2} = |a|+|b|$ hence the (+) solution to the quadratic equation yields a positive δ namely:

$$\boxed{\delta = \frac{-|a| - |b| + \sqrt{(|a| + |b|)^2 + 4\epsilon}}{2}}$$

Yowsers, I almost made this a homework. There may be an easier route. You might notice we have run across a few little lemmas (I've boxed the punch lines for the lemmas) which are doubtless useful in other $\epsilon - \delta$ proofs. We should collect those once we're finished with this proof.

Proof: Let $\epsilon > 0$ and let $(a, b) \in \mathbb{R}^2$. By the calculations that prepared for the proof we know that the following quantity is positive, hence choose

$$\delta = \frac{-|a| - |b| + \sqrt{(|a| + |b|)^2 + 4\epsilon}}{2} > 0.$$

Note that[2],

$$\begin{aligned} |xy - ab| = |xy - ya + ya - ab| &= |y(x-a) + a(y-b)| && \text{algebra} \\ &\leq |y||x-a| + |a||y-b| && \text{triangle inequality} \\ &< (|b| + \delta)\delta + |a|\delta && \text{by the boxed lemmas} \\ &= \delta^2 + \delta(|a| + |b|) && \text{algebra} \\ &= \epsilon \end{aligned}$$

where we know that last step follows due to the steps leading to the boxed equation in the proof preparation. Therefore, $\lim_{(x,y)\to(a,b)} p(x,y) = ab$. and it follows that the product function if continuous at (a,b). But, (a,b) is an arbitrary point thus p is continuous on $\mathbb{R}^2$ hence the product function is continuous. □.

Lemma 3.2.16.

Assume $\delta > 0$.

1. If $a, x \in \mathbb{R}$ then $|x - a| < \delta \;\Rightarrow\; |x| < |a| + \delta$.
2. If $x, a \in \mathbb{R}^n$ then $||x - a|| < \delta \;\Rightarrow\; |x_j - a_j| < \delta$ for $j = 1, 2, \dots n$.

The proof of the proposition above is mostly contained in the remarks of the preceding two pages.

Example 3.2.17. . .

Let $f(x,y) = x^2 + y^2$. We seek to show $\lim_{(x,y)\to(a,b)} (x^2+y^2) = a^2 + b^2$.

Notice $f(\vec{x}) = \vec{x}\cdot\vec{x} = ||\vec{x}||^2$. Recall the Cauchy Schwartz inequality $|\vec{v}\cdot\vec{w}| \leq ||\vec{v}||\,||\vec{w}||$. Calculate,

$$|f(\vec{x}) - f(\vec{A})| = |\vec{x}\cdot\vec{x} - \vec{A}\cdot\vec{A}| = |(\vec{x}-\vec{A})\cdot(\vec{x}+\vec{A})|$$

You can check, $(\vec{x}-\vec{A})\cdot(\vec{x}+\vec{A}) = \vec{x}\cdot\vec{x} - \vec{A}\cdot\vec{x} + \vec{x}\cdot\vec{A} - \vec{A}\cdot\vec{A} = \vec{x}\cdot\vec{x} - \vec{A}\cdot\vec{A}$.

Let $\epsilon > 0$ and choose $\delta = \dfrac{A + \sqrt{A^2+4\epsilon}}{2}$ where $A = |\vec{A}|$.

Suppose $\vec{x} \in \mathbb{R}^2$ such that $|\vec{x} - \vec{A}| < \delta$ then

$$|f(\vec{x}) - f(\vec{A})| \leq ||\vec{x}-\vec{A}||\,||\vec{x}+\vec{A}|| \leq \delta(\delta + A) = \epsilon.$$

$$\therefore \lim_{\vec{x}\to\vec{A}} f(\vec{x}) = f(\vec{A}).$$

[2]my notation is that when we stack inequalities the inequality in a particular line refers only to the immediate vertical successor.

<u>Note</u> : $||\vec{x}+\vec{A}|| \leq ||\vec{x}|| + ||\vec{A}||$

Proposition 3.2.18.

Let $f : V \subseteq \mathbb{R}^p \to \mathbb{R}^m$ and $g : U \subseteq \mathbb{R}^n \to \mathbb{R}^p$ be mappings. Suppose that $\lim_{x \to a} g(x) = b$ and suppose that f is continuous at b then

$$\lim_{x \to a} (f \circ g)(x) = f\big(\lim_{x \to a} g(x)\big).$$

The proof is in Edwards, see pages 46-47. Notice that the proposition above immediately gives us the important result below:

Proposition 3.2.19.

Let f and g be mappings such that $f \circ g$ is well-defined. The composite function $f \circ g$ is continuous for points $a \in dom(f \circ g)$ such that the following two conditions hold:

1. g is continuous at a
2. f is continuous at $g(a)$.

I make use of the earlier proposition that a mapping is continuous iff its component functions are continuous throughout the examples that follow. For example, I know (Id, Id) is continuous since Id was previously proved continuous.

Example 3.2.20. *Note that if $f = p \circ (Id, Id)$ then $f(x) = \big(p \circ (Id, Id)\big)(x) = p\big((Id, Id)(x)\big) = p(x, x) = x^2$. Therefore, the quadratic function $f(x) = x^2$ is continuous on $\mathbb{R}$ as it is the composite of continuous functions.*

Example 3.2.21. *Note that if $f = p \circ (p \circ (Id, Id), Id)$ then $f(x) = p(x^2, x) = x^3$. Therefore, the cubic function $f(x) = x^3$ is continuous on $\mathbb{R}$ as it is the composite of continuous functions.*

Example 3.2.22. *The power function is inductively defined by $x^1 = x$ and $x^n = xx^{n-1}$ for all $n \in \mathbb{N}$. We can prove $f(x) = x^n$ is continous by induction on n. We proved the $n = 1$ case previously. Assume inductively that $f(x) = x^{n-1}$ is continuous. Notice that*

$$x^n = xx^{n-1} = xf(x) = p(x, f(x)) = (p \circ (Id, f))(x).$$

Therefore, using the induction hypothesis, we see that $g(x) = x^n$ is the composite of continuous functions thus it is continuous. We conclude that $f(x) = x^n$ is continuous for all $n \in \mathbb{N}$.

We can play similar games with the sum function to prove that sums of power functions are continuous. In your homework you will prove constant functions are continuous. Putting all of these things together gives us the well-known result that polynomials are continuous on $\mathbb{R}$.

Proposition 3.2.23.

Let a be a limit point of mappings $f, g : U \subseteq \mathbb{R}^n \to V \subseteq \mathbb{R}$ and suppose $c \in \mathbb{R}$. If $\lim_{x\to a} f(x) = b_1 \in \mathbb{R}$ and $\lim_{x\to a} g(x) = b_2 \in \mathbb{R}$ then

1. $\lim_{x\to a}(f(x) + g(x)) = \lim_{x\to a} f(x) + \lim_{x\to a} g(x)$.
2. $\lim_{x\to a}(f(x)g(x)) = \big(\lim_{x\to a} f(x)\big)\big(\lim_{x\to a} g(x)\big)$.
3. $\lim_{x\to a}(cf(x)) = c\lim_{x\to a} f(x)$.

Moreover, if f, g are continuous then $f + g, fg$ and cf are continuous.

Proof: Edwards proves (1.) carefully on pg. 48. I'll do (2.) here: we are given that If $\lim_{x\to a} f(x) = b_1 \in \mathbb{R}$ and $\lim_{x\to a} g(x) = b_2 \in \mathbb{R}$ thus by Proposition 3.2.11 we find $\lim_{x\to a}(f, g)(x) = (b_1, b_2)$. Consider then,

$$
\begin{aligned}
\lim_{x\to a}(f(x)g(x)) &= \lim_{x\to a}\big(p(f, g)\big) && \text{defn. of product function} \\
&= p\big(\lim_{x\to a}(f, g)\big) && \text{since } p \text{ is continuous} \\
&= p(b_1, b_2) && \text{by Proposition 3.2.11.} \\
&= b_1 b_2 && \text{definition of product function} \\
&= \big(\lim_{x\to a} f(x)\big)\big(\lim_{x\to a} g(x)\big).
\end{aligned}
$$

In your homework you proved that $\lim_{x\to a} c = c$ thus item (3.) follows from (2.). $\square$.

The proposition that follows does follow immediately from the proposition above, however I give a proof that again illustrates the idea we used in the examples. Reinterpreting a given function as a composite of more basic functions is a useful theoretical and calculational technique.

Proposition 3.2.24.

Assume $f, g : U \subseteq \mathbb{R}^n \to V \subseteq \mathbb{R}$ are continuous functions at $a \in U$ and suppose $c \in \mathbb{R}$.

1. $f + g$ is continuous at a.
2. fg is continuous at a
3. cf is continuous at a.

Moreover, if f, g are continuous then $f + g, fg$ and cf are continuous.

Proof: Observe that $(f + g)(x) = (s \circ (f, g))(x)$ and $(fg)(x) = (p \circ (f, g))(x)$. We're given that f, g are continuous at a and we know s, p are continuous on all of $\mathbb{R}^2$ thus the composite functions $s \circ (f, g)$ and $p \circ (f, g)$ are continuous at a and the proof of items (1.) and (2.) is complete. To prove (3.) I refer the reader to their homework where it was shown that $h(x) = c$ for all $x \in U$ is a continuous function. We then find (3.) follows from (2.) by setting $g = h$ (function multiplication commutes for real-valued functions). $\square$.

We can use induction arguments to extend these results to arbitrarily many products and sums of power functions.To prove continuity of algebraic functions we'd need to do some more work with quotient and root functions. I'll stop here for the moment, perhaps I'll ask you to prove a few more fundamentals from calculus I. I haven't delved into the definition of exponential or log functions not to mention sine or cosine. We will assume that the basic functions of calculus are continuous on the interior of their respective domains. Basically if the formula for a function can be evaluated at the limit point then the function is continuous.

It's not hard to see that the comments above extend to functions of several variables and mappings. If the formula for a mapping is comprised of finite sums and products of power functions then we can prove such a mapping is continuous using the techniques developed in this section. If we have a mapping with a more complicated formula built from elementary functions then that mapping will be continuous provided its component functions have formulas which are sensibly calculated at the limit point. In other words, if you are willing to believe me that $\sin(x), \cos(x), e^x, \ln(x), \cosh(x), \sinh(x), \sqrt{x}, \frac{1}{x^n}, \dots$ are continuous on the interior of their domains then it's not hard to prove:

$$f(x,y,z) = \left(\sin(x) + e^x + \sqrt{\cosh(x^2) + \sqrt{y+e^x}},\ \cosh(xyz),\ xe^{\sqrt{x+\frac{1}{yz}})}\right)$$

is a continuous mapping at points where the radicands of the square root functions are nonnegative. It wouldn't be very fun to write explicitly but it is clear that this mapping is the Cartesian product of functions which are the sum, product and composite of continuous functions.

Definition 3.2.25.

A **polynomial in n-variables** has the form:

$$f(x_1, x_2, \dots, x_n) = \sum_{i_1, i_2, \dots, i_k = 0}^{\infty} c_{i_1, i_2, \dots, i_n} x_1^{i_1} x_2^{i_2} \cdots x_n^{i_k}$$

where only finitely many coefficients $c_{i_1, i_2, \dots, i_n} \neq 0$. We denote the set of multinomials in n-variables as $\mathbb{R}(x_1, x_2, \dots, x_n)$.

Polynomials are $\mathbb{R}(x)$. Polynomials in two variables are $\mathbb{R}(x,y)$, for example,

$$\begin{array}{lll} f(x,y) & = ax + by & deg(f) = 1, \text{ linear function} \\ f(x,y) & = ax + by + c & deg(f) = 1, \text{ affine function} \\ f(x,y) & = ax^2 + bxy + cy^2 & \text{deg(f)=2, quadratic form} \\ f(x,y) & = ax^2 + bxy + cy^2 + dx + ey + g & \text{deg(f)=2} \end{array}$$

If all the terms in the polynomial have the same number of variables then it is said to be **homogeneous**. In the list above only the linear function and the quadratic form were homogeneous. Returning to the topic of the previous chapter for a moment we should note that a linear transformation has component functions which are homogeneous linear polynomials: suppose that

$L : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation with matrix $A \in \mathbb{R}^{m\times n}$ then in the notation of this chapter we have $L = (L_1, L_2, \dots, L_m)$ where

$$L_j(x) = (Ax) \cdot e_j = A_{j1}x_1 + A_{j2}x_2 + \cdots + A_{jn}x_n$$

It is clear that such functions are continuous since they are the sum of products of continuous functions. Therefore, linear transformations are continuous with respect to the usual metric topology on $\mathbb{R}^n$.

Remark 3.2.26.

There are other topologies possible for $\mathbb{R}^n$. For example, one can prove that

$$||v||_1 = |v_1| + |v_2| + \cdots + |v_n|$$

gives a norm on $\mathbb{R}^n$ and the theorems we proved transfer over almost without change by just trading $||\cdot||$ for $||\cdot||_1$. The unit "ball" becomes a diamond for the 1-norm. There are many other norms which can be constructed, infinitely many it turns out. However, it has been shown that the topology of all these different norms is equivalent. This means that open sets generated from different norms will be the same class of sets. For example, if you can fit an open disk around every point in a set then it's clear you can just as well fit an open diamond and vice-versa. One of the things that makes infinite dimensional linear algebra more fun is the fact that the topology generated by distinct norms need not be equivalent for infinite dimensions. There is a difference between the open sets generated by the Euclidean norm verses those generated by the 1-norm. Incidentally, my thesis work is mostly built over the 1-norm. It makes the supernumbers happy.

3.3 compact sets and continuous images

It should be noted that the sets $\mathbb{R}^n$ and the empty set $\emptyset$ are both open and closed (these are the only such sets in the metric topology, other sets are either open, closed or neither open nor closed).

Theorem 3.3.1.

The mapping $f : dom(f) \subset \mathbb{R}^n \to \mathbb{R}^m$ is continuous iff $f^{-1}(U)$ is open in $dom(f)$ for all open sets $U \subset \mathbb{R}^m$. Additionally, f is continuous iff $f^{-1}(U)$ is closed for each closed set U in $\mathbb{R}^m$.

Notice this theorem makes no explicit reference to the norm. It turns out this theorem is used as the very definition of continuity in more abstract topological settings.

I leave the proof of the closed case to the reader. I tackel the open case here:
Proof: ($\Rightarrow$) Suppose f is continuous and U is open in $\mathbb{R}^m$ then for each $z \in U$ there exists an open ball $B_\epsilon(z) \subset U$. If $x \in f^{-1}(U)$ then there exists $y \in U$ such that $f(x) = y$ and hence there exists an open ball about $B_\epsilon(y) \subset U$. I propose that $f^{-1}(B_\epsilon(y))$ is an open subset of $f^{-1}(U)$ which contains

x. Note that $y \in B_\epsilon(y)$ thus $f(x) = y$ implies $x \in f^{-1}(B_\epsilon(y))$ as according to the definition of inverse image. We seek to show $f^{-1}(B_\epsilon(y)) \subset f^{-1}(U)$. Suppose $v \in f^{-1}(B_\epsilon(y))$. It follows that there exists $w \in B_\epsilon(y)$ such that $f(w) = v$. Note that $B_\epsilon(y) \subset U$ therefore $w \in B_\epsilon(y)$ implies $w \in U$ and so $v \in f^{-1}(U)$ as $w \in U$ has $f(w) = v$. We have shown that an arbitrary element in $f^{-1}(B_\epsilon(y))$ is also in $f^{-1}(U)$ hence $f^{-1}(B_\epsilon(y)) \subseteq f^{-1}(U)$.

($\Leftarrow$) Assume that $f^{-1}(U)$ is open in $dom(f)$ for each open set $U \subset \mathbb{R}^m$. Let $a \in dom(f)$. Assume $\epsilon > 0$ and note that $B_\epsilon(f(a))$ is an open set in $\mathbb{R}^m$ therefore $f^{-1}(B_\epsilon(f(a)))$ is open in $dom(f)$. Note $a \in f^{-1}(B_\epsilon(f(a)))$ since $f(a) \in B_\epsilon(f(a))$. Thus a is a point in the open set $f^{-1}(B_\epsilon(f(a)))$ so there exists a $\delta > 0$ such that $B_\delta(a) \subset f^{-1}(B_\epsilon(f(a))) \subset dom(f)$. Suppose that $x \in B_\delta(a)$ note that $B_\delta(a) \subset f^{-1}(B_\epsilon(f(a)))$ hence $x \in f^{-1}(B_\epsilon(f(a)))$. It follows that there exists $y \in B_\epsilon(f(a))$ such that $f(x) = y$ thus $||f(x) - f(a)|| < \epsilon$. Thus, $\lim_{x \to a} f(x) = f(a)$ for each $a \in dom(f)$ and we conclude that f is continuous. $\square$

Definition 3.3.2.

A mapping S from $\mathbb{N}$ to $\mathbb{R}^n$ is called a **sequence** and we usually denote $S(n) = S_n$ for all $n \in \mathbb{N}$. If $\{a_n\}_{n=1}^{\infty}$ is a sequence then we say $\lim_{n\to\infty} a_n = L$ iff for each $\epsilon > 0$ there exists $N \in \mathbb{N}$ such that for all $n > N$ we have $||a_n - L|| < \epsilon$.

A sequence of vectors is not so different than a sequence of numbers. A sequence in $\mathbb{R}^n$ is just a list of vectors instead of a list of numbers and our concept of distance is provided by the norm rather than the absolute value function.

Example 3.3.3. . .

$$a_n = \langle \tfrac{1}{n}, \tan^{-1}(n), 3 \rangle$$

$$a_n \longrightarrow \langle 0, \tfrac{\pi}{2}, 3 \rangle \quad \text{as} \quad n \longrightarrow \infty$$

You can calculate the limit of a vector-valued sequence by taking limits of the component sequences.

Definition 3.3.4.

A set $C \subset \mathbb{R}^n$ is said to be **compact** iff every sequence of points in C contains a convergent subseqence in C which converges to a point in C

The Bolzano-Weiierstrauss theorem says that every closed interval is compact. It's not hard to see that every closed ball in $\mathbb{R}^n$ is compact. I now collect the interesting results from pg. 52 of Edwards' text: note that to say a set is **bounded** simply means that it is possible to surround the whole set with some sufficiently large open ball.

Proposition 3.3.5.

1. Compact subsets of $\mathbb{R}^n$ are closed and bounded.
2. Closed subsets of a compact set are compact.
3. The cartesian product of compact sets gives a compact set.
4. A subset of $\mathbb{R}^n$ is compact iff it is closed and bounded.

The proof in Edwards is very understandable and the idea of a compact set is really encapsulated by item (4.).

Proposition 3.3.6.

Let C be a compact subset of $\mathbb{R}^n$ and $f : dom(f) \to \mathbb{R}^m$ a continuous mapping with $C \subset dom(f)$, it follows that $f(C)$ is a compact subset of $\mathbb{R}^m$.

The proposition above simply says that the continuous image of compact sets is compact. We finally come to the real reason I am mentioning these topological theorems in this course.

Proposition 3.3.7.

If D is a compact set in $\mathbb{R}^n$ and $f : D \to \mathbb{R}$ is a continuous function then f attains a minimum and maximum value on D. In other words, there exist at least two points $a, b \in D$ such that $f(a) \leq f(x) \leq f(b)$ for all $x \in D$.

Since a closed ball is bounded we have that it is compact and the theorem above tells us that if we take any continuous function then the image of a closed ball under the continuous function will have absolute extreme values relative to the closed ball. This result is important to our later efforts to locate min/max values for functions of several variables. The idea will be that we can approximate the function locally by a quadratic form and the local extreme values will be found by evaluating the quadratic form over the unit-n-sphere.

Definition 3.3.8.

Let $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ be a mapping. We say f is **uniformly continuous** iff for each $\epsilon > 0$ there exists a $\delta > 0$ such that $x, y \in U$ with $||x - y|| < \delta$ we find $||f(x) - f(y)|| < \epsilon$.

Proposition 3.3.9.

If $f : C \to \mathbb{R}$ is a continuous mapping and C is compact then f is uniformly continuous.

The Heine-Borel theorem gives a topological refinement of the definition of compactness we gave earlier in this section. Our definition is equivalent to the following: **a compact set is a set for which every open cover has a finite subcover.** An *open cover* of a set is simply a family of open sets whose unions cover the given set. Theorem 8.10 in Edwards states that if we have a sequence of nested subset in $\mathbb{R}^n$ which contains a compact set:

$$V_1 \subset V_2 \subset V_3 \subset \dots \qquad \text{where} \qquad C \subset \bigcup_{n=1}^{\infty} V_n$$

then if we go far enough out in the sequence we'll be able to find V_N such that $C \subset V_N$. In other words, we can find a finite cover for C. The finite-cover definition is prefered in the abstract setting because it makes no reference to the norm or distance function. In graduate topology you'll learn how to think about open sets and continuity without reference to a norm or distance function. Of course it's better to use the norm and distance function in this course because not using it would just result in a silly needless abstraction which made all the geometry opaque. We have an idea of distance and we're going to use it in this course.

3.4 continuous surfaces

We are often interested in a subset of $\mathbb{R}^m$. A particular subset may be a set of points, a curve, a two-dimensional surface, or generally a p-dimensional surface for $p \leq m$. There are more pathological subsets in general, you might have a subset which is one-dimensional in one sector and two-dimensional in another; for example, $S = (\{0\} \times \mathbb{R}) \cup B_1(0) \subset \mathbb{R}^2$. What dimensionality would you ascribe to S? I give the following defintion to help refine our idea of a p-dimensional continuous surface inside $\mathbb{R}^m$.

Definition 3.4.1.

Let $S \subseteq \mathbb{R}^m$. We say S is **continuous surface of dimension** p iff there exists a finite covering of S say $\bigcup_{i=1}^k V_i = S$ such that $V_i = \Phi_i(U_i)$ for a continuous bijection $\Phi_i : U_i \to V_i$ with continuous inverse and U_i homeomorphic to $\mathbb{R}^p$ for all $i = 1, 2, \dots, k$. We define **homeomorphic to** $\mathbb{R}^p$ to mean that there exists a continuous bijection with continuous inverse from U_i to $\mathbb{R}^p$. In addition, we insist that on the intersections $V_j \cap V_k \neq \emptyset$ the mappings Φ_j, Φ_k are **continuously compatible**. If $V_j \cap V_k \neq \emptyset$ then the mappings Φ_j, Φ_k are said to be **continuously compatible** iff $\Phi_j^{-1} \circ \Phi_k$ is continuous when restricted to $\Phi_k^{-1}(V_j \cap V_k)$. Finally we say two subsets $V \subseteq \mathbb{R}^n$ and $W \subseteq \mathbb{R}^m$ are **homeomorphic** iff there exists a continuous bijection from V to W and we write $V \approx W$ in this case.

You might expect we could just use bijectivity to define dimension of a subset but there are some very strange constructions that forbid such simple thinking. For example, Cantor showed that there is one-one mapping of $\mathbb{R}$ onto $[0, 1] \times [0, 1]$-the unit square. The existence of such a mapping prompts us to state that $\mathbb{R}$ and $\mathbb{R}^2$ share the same *cardnality*. The concept of cardnality ignores dimensionality, it purely focuses on the more basic set-theoretic nature of a given set. Cardnality[3] ignores the difference between $\mathbb{R}$ and $\mathbb{R}^n$. Later Netto showed that such mappings were not continuous. So, you might be tempted to say that a p-dimensional surface is a continuous

[3]I have an introductory chapter on this topic in my math 200 notes

image of $\mathbb{R}^p$. However, in 1890 Peano was able to construct a (!!!) continuous mapping of the unit-interval $[0,1]$ onto the unit square $[0,1] \times [0,1]$. Peano's construction was not a one-one mapping. You can gather from these results that we need both bijectivity and continuity to capture our usual idea of dimensionality. The curves that Cantor and Peano constructed are called **space filling curves**. You might look in Han Sagan's text *Space Filling Curves* if you'd like to see more on this topic.

Example 3.4.2. *Lines are one-dimensional surfaces. A line in $\mathbb{R}^m$ with direction $v \neq 0 \in \mathbb{R}^m$ passing through $a \in \mathbb{R}^m$ has the form $L_v = \{a + tv \mid t \in \mathbb{R}\}$. Note $F(t) = a + tv$ is a continuous mapping from $\mathbb{R}$ into $\mathbb{R}^m$. In this silly case we have $U_1 = \mathbb{R}$ and $\Phi_1 = Id$ so clearly Φ_1 is a continuous bijection and the image $F(\mathbb{R}) = L_v$ is a continuous one-dimensional surface.*

Example 3.4.3. *A plane Pin $\mathbb{R}^m$ with point $a \in \mathbb{R}^m$ containing linearly independent vectors $\vec{u}, \vec{v} \in \mathbb{R}^m$ has the form $P = \{a + s\vec{u} + t\vec{v} \mid (s,t) \in \mathbb{R}^2\}$. Notice that $F(s,t) = a + s\vec{u} + t\vec{v}$ provides a continuous bijection from $\mathbb{R}^2$ to P hence P is a two-dimensional continuous surface in $\mathbb{R}^m$.*

Example 3.4.4. *Suppose that $L : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation. I claim that $range(L) \leq \mathbb{R}^m$ is a continuous surface of dimension $rank(L)$. If the matrix of L is A then the dimension of the surface $L(\mathbb{R})$ is precisely the number of linearly independent column vectors.*

All the examples thus far were examples of *flat surfaces.* Usually *curved surfaces* require more attention.

Example 3.4.5. *The open ball of radius one in $\mathbb{R}^n$ centered at the origin is homeomorphic to $\mathbb{R}^n$. To prove this assertion we need to provide a continuous bijection with continuous inverse from $B_1(0)$ to $\mathbb{R}^n$. A moments thought suggests*

$$\Phi(x) = \begin{cases} \frac{x}{||x||}\tan\frac{\pi||x||}{2} & x \in B_1(0) \text{ such that } x \neq 0 \\ 0 & x = 0 \end{cases}$$

might work. The idea is that the point $x \in B_1(0)$ maps to the point which lies along the same ray eminating from the origin but a distance $\tan\frac{\pi||x||}{2}$ along the ray. Note that as $||x|| \to 1$ we find $\tan\frac{\pi||x||}{2} \to \infty$. This map takes the unit-ball and stretches it to cover $\mathbb{R}^n$. It is clear that Φ is continuous since each component function of Φ is the product and composite of continuous functions. It is clear that $\Phi(x) = 0$ iff $x = 0$. Thus, to prove $1-1$ suppose that $\Phi(x) = \Phi(y)$ for $x, y \in B_1(0)$ such that $x, y \neq 0$. It follows that $||\Phi(x)|| = ||\Phi(y)||$. Hence,

$$\tan\frac{\pi||x||}{2} = \tan\frac{\pi||y||}{2}.$$

But $x, y \in B_1(0)$ thus $||x||, ||y|| < 1$ so we find $0 < \frac{\pi||x||}{2}, \frac{\pi||y||}{2} < \frac{\pi}{2}$. Tangent is one-one on the open interval $(0, \pi/2)$ hence $\frac{\pi||x||}{2} = \frac{\pi||y||}{2}$ therefore $||x|| = ||y||$. Consider the vector equation $\Phi(x) = \Phi(y)$, replace $||y||$ with $||x||$ since we proved they're equal,

$$\frac{x}{||x||}\tan\frac{\pi||x||}{2} = \frac{y}{||x||}\tan\frac{\pi||x||}{2}$$

multiply both sides by the nonzero quantity $||x||/\tan\frac{\pi||x||}{2}$ *to find* $x = y$. *We have shown that* Φ *is injective. The inverse mapping is given by*

$$\Phi^{-1}(v) = \frac{2\tan^{-1}(||v||)}{\pi}\frac{v}{||v||}$$

for $v \in \mathbb{R}^n$ *such that* $v \neq 0$ *and* $\Phi^{-1}(0) = 0$. *This map takes the vector* v *and compresses it into the unit-ball. Notice that the vector length approaches infinitity the inverse maps closer and closer to the boundary of the ball as the inverse tangent tends to* $\pi/2$ *as its input tends to infinity. I invite the reader to verify that this is indeed the inverse of* Φ, *you need to show that* $\Phi(\Phi^{-1}(v) = v$ *for all* $v \in \mathbb{R}^n$ *and* $\Phi^{-1}(\Phi(x)) = x$ *for all* $x \in B_1(0)$.

Remark 3.4.6.

The example above gives us license to use open balls as the domains for the mappings which define a continuous surface. You could call the Φ_i **continuous patches** if you wish. A smooth surface will be defined in terms of smooth patches or perhaps in terms of the inverse maps Φ_i^{-1}:which are called **coordinate maps.** We need to define a few ideas about differentiability before we can give the definition for a smooth surface. In fact the concept of a surface and the definition of the derivative in some sense are inseparable. For this reason I have begun the discussion of surfaces in this chapter.

I assume that the closed unit-ball $\overline{B_1(0)}$ is homeomorphic to $\mathbb{R}^2$ in the example below. I leave it to the reader supply proof of that claim.

Example 3.4.7. *I claim that the unit-two-sphere* S^2 *is a two-dimensional continuous surface in* $\mathbb{R}^3$. *We define*

$$S^2 = \partial B_1(0) = \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 = 1\}$$

We can write $S^2 = S^+ \cup S^-$ *where we define the upper hemisphere* S^+ *and the lower hemisphere* S^- *in the usual manner:*

$$S^+ = \{(x, y, z) \in S^2 \mid z \geq 0\} \qquad S^- = \{(x, y, z) \in S^2 \mid z \leq 0\}$$

The equator is at the intersection,

$$E = S^+ \cap S^- = \{(x, y, z) \in S^2 | z = 0\} = S^1 \times \{0\}$$

Define mappings $\Phi_\pm : \overline{B_1(0)} \subset \mathbb{R}^2 \to S^\pm$ *as follows:*

$$\Phi_\pm(x, y) = (\ x,\ y,\ \pm\sqrt{1 - x^2 - y^2}\ )$$

where $(x, y) \in \mathbb{R}^2$ *such that* $x^2 + y^2 \leq 1$. *I claim the inverse mappings are*

$$\Phi_\pm^{-1}(x, y, z) = (\ x,\ y\ ).$$

for all $(x,y,z) \in S^{\pm}$. *Let's check to see if my claim is correct in the (+) case. Let* $(x,y,z) \in S^+$,

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

since $(x,y,z) \in S^+$ *implies* $z = \sqrt{1-x^2-y^2}$. *The (-) case is similar. Likewise let* $(x,y) \in B_1(0) \subset \mathbb{R}^2$ *and calculate*

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

It follows that $\Phi_{\pm}$ *are bijections and it is clear from their formulas that they are continuous mappings. We should check if these are compatible patches. Consider the mapping* $\Phi_+^{-1} \circ \Phi_-$. *A typical point in the* $\Phi_-^{-1}(E)$ *should have the form* $(x,y) \in S^1$ *which means* $x^2+y^2=1$, *consider then*

$$(\Phi_+^{-1} \circ \Phi_-)(x,y) = \Phi_+^{-1}(x,y,-\sqrt{1-x^2-y^2}) = (x,y)$$

thus $\Phi_+^{-1} \circ \Phi_-$ *is the identity mapping which is continuous. We find that the two-sphere is a continuous two-dimensional surface.*

Example 3.4.8. *Let* U *be homeomorphic to* $\mathbb{R}^p$. *The image of a continuous mapping* $F : U \to \mathbb{R}^m$ *is a* p*-dimensional continuous surface in* $\mathbb{R}^m$. *In this case compatibility is trivially satisfied.*

Remark 3.4.9.

A p-dimensional surface is locally modeled by $\mathbb{R}^p$. You can imagine pasting p-space over the surface. Bijectivity and continuity insure that the pasting is not pathological as in Cantors' bijective mapping of $[0,1]$ onto $\mathbb{R}^n$ or Peano's continuous mapping of $[0,1]$ onto $[0,1]\times[0,1]$. In a later chapter we'll add the criteria of differentiability of the mapping. This will make the pasting keep from getting crinkled up at a point. For example, a cone is a continuous surface however it is not a smooth surface due to the point of the cone

Chapter 4

geometry of curves

If the curve is assigned a sense of direction then we call it an **oriented curve**. A particular curve can be parametrized by many different paths. You can think of a parametrization of a curve as a process of pasting a flexible numberline onto the curve.

Definition 4.0.10.

Let $C \subseteq \mathbb{R}^n$ be an oriented curve which starts at P and ends at Q. We say that $\gamma : [a, b] \rightarrow \mathbb{R}^n$ is a **smooth non-stop parametrization** of C if $\gamma([a, b]) = C$, $\gamma(a) = P$, $\gamma(b) = Q$, and γ is smooth with $\gamma'(t) \neq 0$ for all $t \in [a, b]$. We will typically call γ a **path** from P to Q which covers the curve C.

I have limited the definition to curves with endpoints however the definition for curves which go on without end is very similar. You can just drop one or both of the endpoint conditions.

4.1 arclength

Let's begin by analyzing the tangent vector to a path in three dimensional space. Denote $\gamma = (x, y, z)$ where $x, y, z \in C^\infty([a, b], \mathbb{R})$ and calculate that

$$\gamma'(t) = \tfrac{d\gamma}{dt} =< \tfrac{dx}{dt}, \tfrac{dy}{dt}, \tfrac{dz}{dt} > .$$

Multiplying by dt yields

$$\gamma'(t)dt = \tfrac{d\gamma}{dt}dt =< \tfrac{dx}{dt}, \tfrac{dy}{dt}, \tfrac{dz}{dt} > dt.$$

The arclength ds subtended from time t to time $t + dt$ is simply the length of the vector $\gamma'(t)dt$ which yields,

$$ds = ||\gamma'(t)dt|| = \sqrt{\tfrac{dx}{dt}^2 + \tfrac{dy}{dt}^2 + \tfrac{dz}{dt}^2}dt$$

You can think of this as the length of a tiny bit of string that is laid out along the curve from the point $\gamma(t)$ to the point $\gamma(t + dt)$. Of course this infinitesimal notation is just shorthand for an explicit limiting processes. If we sum together all the little bits of arclength we will arrive at the total arclength of the curve. In fact, this is how we define the arclength of a curve. The preceding discussion was in 3 dimensions but the formulas stated in terms of the norm generalizes naturally to $\mathbb{R}^n$.

Definition 4.1.1.

Let $\gamma : [a,b] \to \mathbb{R}^n$ be a smooth, non-stop path which covers the oriented curve C. The **arclength function** of γ is a function $s_\gamma : [a,b] \to \mathbb{R}$ where

$$s_\gamma = \int_a^t ||\gamma'(u)||\, du$$

for each $t \in [a,b]$. If $\tilde{\gamma}$ is a smooth non-stop path such that $||\tilde{\gamma}'(t)|| = 1$ then we say that $\tilde{\gamma}$ is a unit-speed curve. Moreover, we say $\tilde{\gamma}$ is parametrized with respect to arclength.

The arclength function has many special properties. Notice that item (1.) below is actually just the statement that the speed is the magnitude of the velocity vector.

Proposition 4.1.2.

Let $\gamma : [a,b] \to \mathbb{R}^n$ be a smooth, non-stop path which covers the oriented curve C. The **arclength function** of γ denoted by $s_\gamma : [a,b] \to \mathbb{R}$ has the following properties:

1. $\frac{d}{dt}(s_\gamma(w)) = ||\gamma'(w)||\frac{dw}{dt}$,
2. $\frac{ds_\gamma}{dt} > 0$ for all $t \in (a,b)$,
3. s_γ is a 1-1 function,
4. s_γ has inverse $s_\gamma^{-1} : s_\gamma([a,b]) \to [a,b]$.

Proof: We begin with (1.). We apply the fundamental theorem of calculus:

$$\frac{d}{dt}(s_\gamma(w)) = \frac{d}{dt}\int_a^w ||\gamma'(u)||\, du = ||\gamma'(w)||\frac{dw}{dt}$$

for all $w \in (a,b)$. For (2.), set $w = t$ and recall that $||\gamma'(t)|| = 0$ iff $\gamma'(t) = 0$ however we were given that γ is non-stop so $\gamma'(t) \neq 0$. We find $\frac{ds_\gamma}{dt} > 0$ for all $t \in (a,b)$ and consequently the arclength function is an increasing function on (a,b). For (3.), suppose (towards a contradiction) that $s_\gamma(x) = s_\gamma(y)$ where $a < x < y < b$. Note that γ smooth implies s_γ is differentiable with continuous derivative on (a,b) therefore the mean value theorem applies and we can deduce that there is some point on $c \in (x,y)$ such that $s_\gamma'(c) = 0$, which is impossible, therefore (3.) follows. If a function is 1-1 then we can construct the inverse pointwise by simply going backwards for each point mapped to in the range; $s_\gamma^{-1}(x) = y$ iff $s_\gamma(y) = x$. The fact that s_γ is single-valued follows from (3.). $\square$

If we are given a curve C covered by a path γ (which is smooth and non-stop but may not be unit-speed) then we can reparametrize the curve C with a unit-speed path $\tilde{\gamma}$ as follows:

$$\boxed{\tilde{\gamma}(s) = \gamma(s_\gamma^{-1}(s))}$$

where s_γ^{-1} is the inverse of the arclength function.

Proposition 4.1.3.

If γ is a smooth non-stop path then the path $\tilde{\gamma}$ defined by $\tilde{\gamma}(s) = \gamma(s_\gamma^{-1}(s))$ is unit-speed.

Proof: Differentiate $\tilde{\gamma}(t)$ with respect to t, we use the chain-rule,

$$\tilde{\gamma}'(t) = \tfrac{d}{dt}(\gamma(s_\gamma^{-1}(t))) = \gamma'(s_\gamma^{-1}(t))\tfrac{d}{dt}(s_\gamma^{-1}(t)).$$

Hence $\tilde{\gamma}'(t) = \gamma'(s_\gamma^{-1}(t))\frac{d}{dt}(s_\gamma^{-1}(t))$. Recall that if a function is increasing on an interval then its inverse is likewise increasing hence, by (2.) of the previous proposition, we can pull the positive constant $\frac{d}{dt}(s_\gamma^{-1}(t))$ out of the norm. We find, using item (1.) in the previous proposition,

$$||\tilde{\gamma}'(t)|| = ||\gamma'(s_\gamma^{-1}(t))||\tfrac{d}{dt}(s_\gamma^{-1}(t)) = \tfrac{d}{dt}(s_\gamma(s_\gamma^{-1}(t))) = \tfrac{d}{dt}(t) = 1.$$

Therefore, the curve $\tilde{\gamma}$ is unit-speed. We have $ds/dt = 1$ when $t = s$ (this last sentence is simply a summary of the careful argument we just concluded). $\square$

Remark 4.1.4.

While there are many paths which cover a particular oriented curve the unit-speed path is unique and we'll see that formulas for unit-speed curves are particularly simple.

Example 4.1.5. .

$\gamma(t) = \vec{r}(t) = \langle R\cos t, 3, R\sin t\rangle$ for $t \geq 0$, $\underbrace{R > 0}_{\text{fixed constant.}}$

$\frac{d\vec{r}}{dt} = \langle -R\sin t, 0, R\cos t\rangle$

$\Rightarrow \frac{d\vec{r}}{dt}\cdot\frac{d\vec{r}}{dt} = R^2$ & $\left\|\frac{d\vec{r}}{dt}\right\| = R$

$s(t) = \int_0^t \left\|\frac{d\vec{r}}{du}\right\| du = \int_0^t R\,du = Ru\Big|_0^t = \underline{Rt = s}$.

For example, $s(2\pi) = 2\pi R$ (make sense?)

Note $t = s/R$ hence we can reparametrize via s,

$\tilde{\gamma}(s) = \vec{r}(s/R) = \langle R\cos(s/R), 3, R\sin(s/R)\rangle$

unit-speed parametrization of curve.

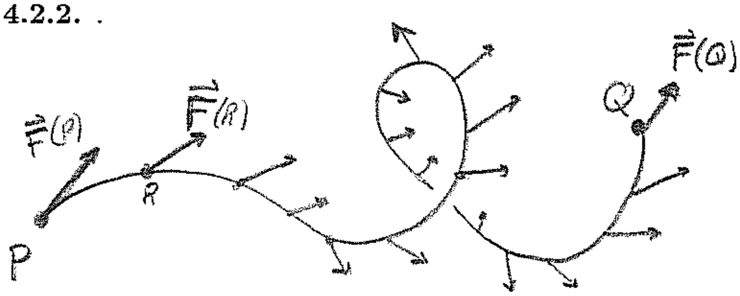
4.2 vector fields along a path

Definition 4.2.1.

Let $C \subseteq \mathbb{R}^3$ be an oriented curve which starts at P and ends at Q. A **vector field along the curve** C is a function from $C \to V^3$. You can visualize this as attaching a vector to each point on C.

The tangent (T), normal(N) and binormal (B) vector fields defined below will allow us to identify when two oriented curves have the same shape.

Example 4.2.2. .



vector field along curve C assigns vector at each point on C.

Definition 4.2.3.

Let $\gamma : [a, b] \to \mathbb{R}^3$ be a path from P to Q in $\mathbb{R}^3$. The **tangent vector field** of γ is a mapping $T : [a, b] \to V^3$ defined by

$$T(t) = \frac{1}{||\gamma'(t)||}\gamma'(t)$$

for each $t \in [a, b]$. Likewise, if $T'(t) \neq 0$ for all $t \in [a, b]$ then the **normal vector field** of γ is a mapping $N : [a, b] \to V^3$ defined by

$$N(t) = \frac{1}{||T'(t)||}T'(t)$$

for each $t \in [a, b]$. Finally, if $T'(t) \neq 0$ for all $t \in [a, b]$ then the **binormal vector field** of γ is defined by $B(t) = T(t) \times N(t)$ for all $t \in [a, b]$

Example 4.2.4. *Let $R > 0$ and suppose $\gamma(t) = (R\cos(t), R\sin(t), 0)$ for $0 \leq t \leq 2\pi$. We can calculate*

$$\gamma'(t) =< -R\sin(t), R\cos(t), 0 > \;\Rightarrow\; ||\gamma'(t)|| = R.$$

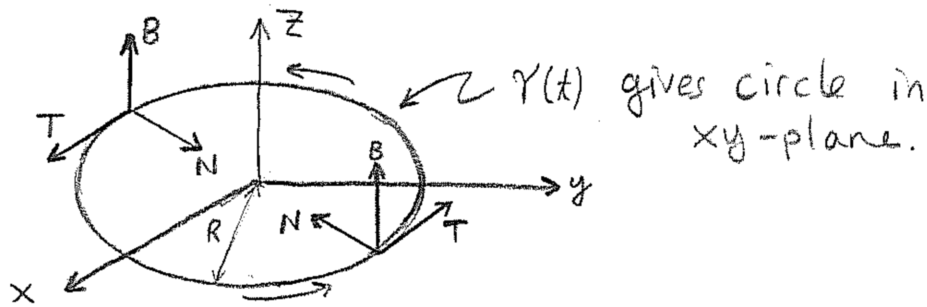
Hence $T(t) =< -\sin(t), \cos(t), 0 >$ and we can calculate,

$$T'(t) =< -\cos(t), -\sin(t), 0 > \;\Rightarrow\; ||T'(t)|| = 1.$$

Thus $N(t) =< -\cos(t), -\sin(t), 0 >$. *Finally we calculate the binormal vector field,*

$$\begin{aligned} B(t) = T(t) \times N(t) &= [-\sin(t)e_1 + \cos(t)e_2] \times [-\cos(t)e_1 - \sin(t)e_2] \\ &= [\sin^2(t)e_1 \times e_2 - \cos^2(t)e_2 \times e_1 \\ &= [\sin^2(t) + \cos^2(t)]e_1 \times e_2 \\ &= e_3 =< 0, 0, 1 > \end{aligned}$$

Notice that $T \cdot N = N \cdot B = T \cdot B = 0$. *For a particular value of* t *the vectors* $\{T(t), N(t), B(t)\}$ *give an orthogonal set of unit vectors, they provide a comoving frame for* γ. *It can be shown that the tangent and normal vectors span the plane in which the path travels for times infintesimally close to* t. *This plane is called the* **osculating plane**. *The binormal vector gives the normal to the osculating plane. The curve considered in this example has a rather boring osculating plane since* B *is constant. This curve is just a circle in the xy-plane which is traversed at constant speed.*



Example 4.2.5. *Notice that* $s_\gamma(t) = Rt$ *in the preceding example. It follows that* $\widetilde{\gamma}(s) = (R\cos(s/R), R\sin(s/R), 0)$ *for* $0 \leq s \leq 2\pi R$ *is the unit-speed path for curve. We can calculate*

$$\widetilde{\gamma}'(s) =< -\sin(s/R), \cos(s/R), 0 > \;\Rightarrow\; ||\widetilde{\gamma}'(s)|| = 1.$$

Hence $\widetilde{T}(s) =< -\sin(s/R), \cos(s/R), 0 >$ *and we can also calculate,*

$$\widetilde{T}'(s) = \tfrac{1}{R} < -\cos(s/R), -\sin(s/R), 0 > \;\Rightarrow\; ||\widetilde{T}'(t)|| = 1/R.$$

Thus $\widetilde{N}(s) =< -\cos(s/R), -\sin(s/R), 0 >$. *Note* $\widetilde{B} = \widetilde{T} \times \widetilde{N} =< 0, 0, 1 >$ *as before.*

Example 4.2.6. *Let* $m, R > 0$ *and suppose* $\gamma(t) = (R\cos(t), R\sin(t), mt)$ *for* $0 \leq t \leq 2\pi$. *We can calculate*

$$\gamma'(t) =< -R\sin(t), R\cos(t), m > \;\Rightarrow\; ||\gamma'(t)|| = \sqrt{R^2 + m^2}.$$

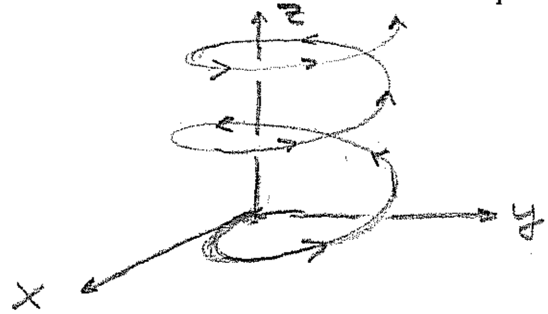
Hence $T(t) = \frac{1}{\sqrt{R^2+m^2}} < -R\sin(t), R\cos(t), m >$ *and we can calculate,*

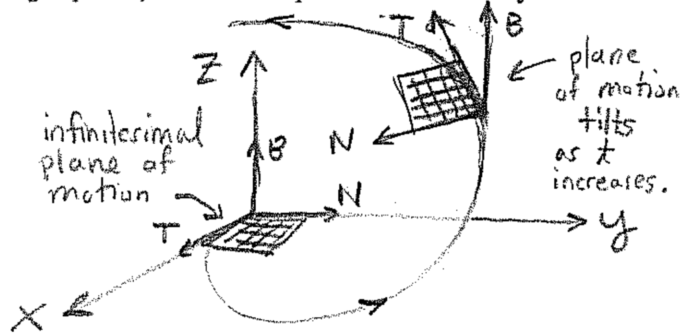
$$T'(t) = \frac{1}{\sqrt{R^2+m^2}} < -R\cos(t), -R\sin(t), 0 > \;\Rightarrow\; ||T'(t)|| = \frac{R}{\sqrt{R^2+m^2}}.$$

Thus $N(t) =< -\cos(t), -\sin(t), 0 >$. *Finally we calculate the binormal vector field,*

$$\begin{aligned} B(t) = T(t) \times N(t) &= \frac{1}{\sqrt{R^2+m^2}}[-R\sin(t)e_1 + R\cos(t)e_2 + me_3] \times [-\cos(t)e_1 - \sin(t)e_2] \\ &= \frac{1}{\sqrt{R^2+m^2}} < m\sin(t), -m\cos(t), R > \end{aligned}$$

We again observe that $T \cdot N = N \cdot B = T \cdot B = 0$. The **osculating plane** *is moving for this curve, note the t-dependence. This curve does not stay in a single plane, it is not a planar curve. In fact this is a circular helix with radius R and slope m. .*





Example 4.2.7. *Lets reparametrize the helix as a unit-speed path. Notice that $s_\gamma(t) = t\sqrt{R^2+m^2}$ thus we should replace t with $s/\sqrt{R^2+m^2}$ to obtain $\tilde{\gamma}(s)$. Let $a = 1/\sqrt{R^2+m^2}$ and $\tilde{\gamma}(s) = (R\cos(as), R\sin(as), ams)$ for $0 \leq s \leq 2\pi\sqrt{R^2+m^2}$. We can calculate*

$$\tilde{\gamma}'(s) = < -Ra\sin(as), Ra\cos(as), am > \;\Rightarrow\; ||\tilde{\gamma}'(s)|| = a\sqrt{R^2+m^2} = 1.$$

Hence $\tilde{T}(s) = a < -R\sin(as), R\cos(as), m >$ and we can calculate,

$$\tilde{T}'(s) = Ra^2 < -\cos(as), -\sin(as), 0 > \;\Rightarrow\; ||\tilde{T}'(s)|| = Ra^2 = \tfrac{R}{R^2+m^2}.$$

Thus $\tilde{N}(s) = < -\cos(as), -\sin(as), 0 >$. Next, calculate the binormal vector field,

$$\begin{aligned}\tilde{B}(s) = \tilde{T}(s) \times \tilde{N}(s) &= a < -R\sin(as), R\cos(as), m > \times < -\cos(as), -\sin(as), 0 > \\ &= \tfrac{1}{\sqrt{R^2+m^2}} < m\sin(as), -m\cos(as), R >\end{aligned}$$

Hopefully you can start to see that the unit-speed path shares the same T, N, B frame at arclength s as the previous example with $t = s/\sqrt{R^2+m^2}$.

4.3 Frenet Serret equations

We now prepare to prove the Frenet Serret formulas for the T, N, B frame fields. It turns out that for nonlinear curves the T, N, B vector fields always provide an orthonormal frame. Moreover, for nonlinear curves, we'll see that the **torsion** and **curvature** capture the geometry of the curve.

Proposition 4.3.1.

If γ is a path with tangent, normal and binormal vector fields T, N and B then $\{T(t), N(t), B(t)\}$ is an orthonormal set of vectors for each $t \in dom(\gamma)$.

Proof: It is clear from $B(t) = T(t) \times N(t)$ that $T(t) \cdot B(t) = N(t) \cdot B(t) = 0$. Furthermore, it is also clear that these vectors have length one due to their construction as unit vectors. In particular this means that $T(t) \cdot T(t) = 1$. We can differentiate this to obtain (by the product rule for dot-products)

$$T'(t) \cdot T(t) + T(t) \cdot T'(t) = 0 \quad \Rightarrow \quad 2T(t) \cdot T'(t) = 0$$

Divide by $||T'(t)||$ to obtain $T(t) \cdot N(t) = 0$. $\square$

We omit the explicit t-dependence for the dicussion to follow here, also you should assume the vector fields are all derived from a particular path γ. Since T, N, B are nonzero and point in three mutually distinct directions it follows that any other vector can be written as a linear combination of T, N, B. This means[1] if $v \in V^3$ then there exist c_1, c_2, c_3 such that $v = c_1T + c_2N + c_3B$. The orthonormality is very nice because it tells us we can calculate the coefficients in terms of dot-products with T, N and B:

$$\boxed{v = c_1T + c_2N + c_3B \quad \Rightarrow \quad c_1 = v \cdot T, \ c_2 = v \cdot N, \ c_3 = v \cdot B}$$

We will make much use of the observations above in the calculations that follow. Suppose that

$$\begin{aligned} T' &= c_{11}T + c_{12}N + c_{13}B \\ N' &= c_{21}T + c_{22}N + c_{23}B \\ B' &= c_{31}T + c_{32}N + c_{33}B. \end{aligned}$$

We observed previously that $T' \cdot T = 0$ thus $c_{11} = 0$. It is easy to show $N' \cdot N = 0$ and $B' \cdot B = 0$ thus $c_{22} = 0$ and c_{33}. Furthermore, we defined $N = \frac{1}{||T'||}T'$ hence $c_{13} = 0$. Note that

$$T' = c_{12}N = \tfrac{c_{12}}{||T'||}T' \quad \Rightarrow \quad c_{12} = ||T'||.$$

To summarize what we've learned so far:

$$\begin{aligned} T' &= c_{12}N \\ N' &= c_{21}T + c_{23}B \\ B' &= c_{31}T + c_{32}N. \end{aligned}$$

We'd like to find some condition on the remaining coefficients. Consider that:

$$\begin{array}{llll} B = T \times N & \Rightarrow & B' = T' \times N + T \times N' & \text{a product rule} \\ & \Rightarrow & B' = [c_{12}N] \times N + T \times [c_{21}T + c_{23}B] & \text{using previous eqn.} \\ & \Rightarrow & B' = c_{23}T \times B & \text{noted } N \times N = T \times T = 0 \\ & \Rightarrow & B' = -c_{23}N & \text{you can show } N = B \times T. \\ & \Rightarrow & c_{31}T + c_{32}N = -c_{23}N & \text{refer to previous eqn.} \\ & \Rightarrow & c_{31} = 0 \text{ and } c_{32} = -c_{23}. & \text{using LI of } \{T, N\} \end{array}$$

[1]You might recognize $[v]_\beta = [c_1, c_2, c_3]^T$ as the coordinate vector with respect to the basis $\beta = \{T, N, B\}$

We have reduced the initial set of equations to the following:

$$\begin{aligned} T' &= c_{12}N \\ N' &= c_{21}T + c_{23}B \\ B' &= -c_{23}N. \end{aligned}$$

The equations above encourage us to define the **curvature** and **torsion** as follows:

Definition 4.3.2.

Let C be a curve which is covered by the unit-speed path $\tilde{\gamma}$ then we define the curvature κ and torsion τ as follows:

$$\boxed{\kappa(s) = \left|\left|\frac{d\widetilde{T}}{ds}\right|\right| \qquad \tau(s) = -\frac{d\widetilde{B}}{ds} \cdot \widetilde{N}(s)}$$

One of your homework questions is to show that $c_{21} = -c_{12}$. Given the result you will prove in the homework we find the famous **Frenet-Serret** equations:

$$\boxed{\frac{d\widetilde{T}}{ds} = \kappa\widetilde{N} \qquad \frac{d\widetilde{N}}{ds} = -\kappa\widetilde{T} + \tau\widetilde{B} \qquad \frac{d\widetilde{B}}{ds} = -\tau\widetilde{N}.}$$

We had to use the arclength parameterization to insure that the formulas above unambiguously define the curvature and the torsion. In fact, if we take a particular (unoriented) curve then there are two choices for orienting the curve. You can show that that the torsion and curvature are independent of the choice of orientation. Naturally the total arclength is also independent of the orientation of a given curve.

Curvature, torsion can also be calculated in terms of a path which is not unit speed. We simply replace s with the arclength function $s_\gamma(t)$ and make use of the chain rule. Notice that $dF/dt = (ds/dt)(d\widetilde{F}/ds)$ hence,

$$\frac{dT}{dt} = \frac{ds}{dt}\frac{d\widetilde{T}}{ds}, \quad \frac{dN}{dt} = \frac{ds}{dt}\frac{d\widetilde{N}}{ds}, \quad \frac{dB}{dt} = \frac{ds}{dt}\frac{d\widetilde{B}}{ds}$$

Or if you prefer, use the dot-notation $ds/dt = \dot{s}$ to write:

$$\frac{1}{\dot{s}}\frac{dT}{dt} = \frac{d\widetilde{T}}{ds}, \quad \frac{1}{\dot{s}}\frac{dN}{dt} = \frac{d\widetilde{N}}{ds}, \quad \frac{1}{\dot{s}}\frac{dB}{dt} = \frac{d\widetilde{B}}{ds}$$

Substituting these into the unit-speed Frenet Serret formulas yield:

$$\boxed{\frac{dT}{dt} = \dot{s}\kappa N \qquad \frac{dN}{dt} = -\dot{s}\kappa T + \dot{s}\tau B \qquad \frac{dB}{dt} = -\dot{s}\tau N.}$$

where $\widetilde{T}(s_\gamma(t)) = T(t)$, $\widetilde{N}(s_\gamma(t)) = N(t)$ and $\widetilde{B}(s_\gamma(t)) = B(t)$. Likewise deduce[2] that

$$\kappa(t) = \frac{1}{\dot{s}}\left|\left|\frac{dT}{dt}\right|\right| \qquad \tau(t) = -\frac{1}{\dot{s}}\left(\frac{dB}{dt} \cdot N(t)\right)$$

[2] I'm using the somewhat ambiguous notation $\kappa(t) = \kappa(s_\gamma(t))$ and $\tau(t) = \tau(s_\gamma(t))$. We do this often in applications of calculus. Ask me if you'd like further clarification on this point.

4.4 curvature, torsion and the osculating plane

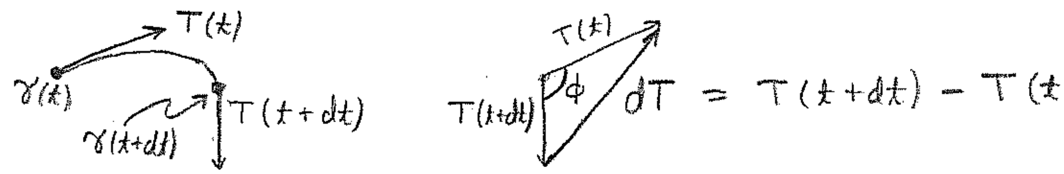
In the preceding section we saw how the calculus and linear algebra suggest we define curvature and torsion. We now stop to analyze the geometric meaning of those definitions.

4.4.1 curvature

Let use begin with the curvature. Assume γ is a non-stop smooth path,

$$\kappa = \frac{1}{\dot{s}}\left|\left|\frac{dT}{dt}\right|\right|$$

Infinitesimally this equation gives $||dT|| = \kappa \dot{s} dt = \kappa \frac{ds}{dt} dt = \kappa\, ds$. But this is a strange equation since $||T|| = 1$. So what does this mean? Perhaps we should add some more detail to resolve this puzzle; let $dT = T(t + dt) - T(t)$.



Notice that

$$\begin{aligned} ||dT||^2 &= [T(t + dt) - T(t)] \cdot [T(t + dt) - T(t)] \\ &= T(t + dt) \cdot T(t + dt) + T(t) \cdot T(t) - 2T(t) \cdot T(t + dt) \\ &= T(t + dt) \cdot T(t + dt) + T(t) \cdot T(t) - 2T(t) \cdot T(t + dt) \\ &= 2(1 - \cos(\phi))) \end{aligned}$$

where we define ϕ to be the angle between $T(t)$ and $T(t + dt)$. This angle measures the change in direction of the tangent vector at t goes to $t + dt$. Since this is a small change in time it is reasonable to expect the angle ϕ is small thus $\cos(\phi) \approx 1 - \frac{1}{2}\phi^2$ and we find that

$$||dT|| = \sqrt{2(1 - \cos(\phi)} = \sqrt{2(1 - 1 + \tfrac{1}{2}\phi^2)} = \sqrt{\phi^2} = |\phi|$$

Therefore, $||dT|| = \kappa\, ds = |\phi|$ and we find $\boxed{\kappa = \pm \frac{ds}{d\phi}}$.

Remark 4.4.1.

The curvature measures the infinitesimal change in the direction of the unit-tangent vector to the curve. We say the the reciprocal of the curvature is the **radius of curvature** $r = \frac{1}{\kappa}$. This makes sense as $ds = |1/\kappa| d\phi$ suggests that a circle of radius $1/\kappa$ fits snuggly against the path at time t. We form the **osculating circle at each point along the path by placing a circle of radius** $1/\kappa$ **tangent to the unit-tangent vector in the plane with normal** $B(t)$. We probably should draw a picture of this.

4.4.2 osculating plane and circle

It was claimed that the "infinitesimal" motion of the path resides in a plane with normal B. Suppose that at some time t_o the path reaches the point $\gamma(t_o) = P_o$. Infinitesimally the tangent line matches the path and we can write the parametric equation for the tangent line as follows:

$$l(t) = \gamma(t_o) + t\gamma'(t_o) = P_o + tv_oT_o$$

where we used that $\gamma'(t) = \dot{s}T(t)$ and we evaluated at $t = t_o$ to define $\dot{s}(t_o) = v_o$ and $T(t_o) = T_o$. The normal line through P_o has parametric equations (using $N_o = N(t_o)$):

$$n(\lambda) = P_o + \lambda N_o$$

We learned in the last section that the path bends away from the tangent line along a circle whose radius is $1/\kappa_o$. We find the infinitesimal motion resides in the plane spanned by T_o and N_o which has normal $T_o \times N_o = B(t_o)$. The tangent line and the normal line are perpendicular and could be thought of as a xy-coordinate axes in the osculating plane. The osculating circle is found with its center on the normal line a distance of $1/\kappa_o$ from P_o. Thus the center of the circle is at:

$$Q_o = P_o - \tfrac{1}{\kappa_o}N_o$$

I'll think of constructing x, y, z coordinates based at P_o with respect to the T_o, N_o, B_o frame. We suppose $\vec{r}$ be a point on the osculating circle and x, y, z to be the coefficients in $\vec{r} = P_o + xT_o + yN_o + zB_o$. Since the circle is in the plane based at P_o with normal B_o we should set $z = 0$ for our circle thus $\vec{r} = xT + yN$.

$$||\vec{r} - Q_o||^2 = \tfrac{1}{\kappa_o^2} \quad \Rightarrow \quad ||xT_o + (y + \tfrac{1}{\kappa_o})N_o]||^2 = \tfrac{1}{\kappa_o^2}.$$

Therefore, by the pythagorean theorem for orthogonal vectors, the x, y, z equations for the osculating circle are simply[3] :

$$\boxed{x^2 + (y + \tfrac{1}{\kappa_o})^2 = \tfrac{1}{\kappa_o^2}, \quad z = 0.}$$

[3]Of course if we already use x, y, z in a different context then we should use other symbols for the equation of the osculating circle.

Finally, notice that if the torsion is zero then the Frenet Serret formulas simplify to:

$$\boxed{\frac{dT}{dt} = \dot{s}\kappa N \qquad \frac{dN}{dt} = -\dot{s}\kappa T \qquad \frac{dB}{dt} = 0.}$$

we see that B is a constant vector field and motion will remain in the osculating plane. The change in the normal vector causes a change in the tangent vector and vice-versa however the binormal vector is not coupled to T or N.

Remark 4.4.2.

> The torsion measures the infinitesimal change in the direction of the binormal vector relative to the normal vector of the curve. Because the normal vector is in the plane of infinitesimal motion and the binormal is perpendicular to that plane we can say that the torsion measures how the path lifts or twists up off the plane of infinitesimal motion. Furthermore, we can expect path which is trapped in a particular plane (these are called **planar** curves) will have torsion which is identically zero. We should also expect that the torsion for something like a helix will be nonzero everywhere since the motion is always twisting up off the plane of infinitesimal motion. It is probable you will examine these questions in your homework.

4.5 acceleration and velocity

Let's see how the preceding section is useful in the analysis of the motion of physical objects. In the study of dynamics or the physics of motion the critical objects of interest are the position, velocity and acceleration vectors. Once a force is supplied we can in principle solve Newton's Second Law $\vec{F} = m\vec{A}$ and find the equation of motion $\vec{r} = \vec{r}(t)$. Moreover, since the map $t \mapsto \vec{r}(t)$ is a path we can analyze the velocity and acceleration in terms of the **Frenet Frame** $\{T, N, B\}$. To keep it interesting we'll assume the motion is non-stop and smooth so that the analysis of the last section applies.

(for now the next two pages are stolen from a course I took from Dr. R.O. Fulp some years back)

Let $\alpha : [a, b] \longrightarrow \mathbb{R}^3$ be any curve. Def$^{\underline{n}}$ $v(t) = \|\alpha'(t)\|$

$$T(t) = \frac{\alpha'(t)}{\|\alpha'(t)\|}$$

$$\alpha'(t) = \|\alpha'(t)\| \, T(t) = v(t)\, T(t)$$

$$\alpha''(t) = v(t)\, T'(t) + v'(t)\, T(t)$$

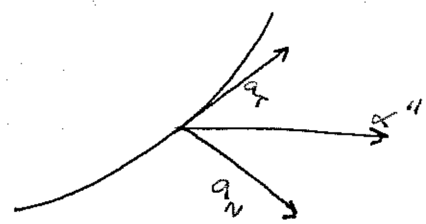
$$= v(t) \|T'(t)\| N(t) + v'(t)\, T(t)$$

<u>Notation</u>

$a_T(t) = v'(t)$ tangential acceleration

$a_N(t) = v(t) \, \|T'(t)\|$ normal or radial acceleration

$$\alpha''(t) = a_T(t)\, T(t) + a_N(t)\, N(t)$$



$$\alpha' \times \alpha'' = (v T) \times (a_T T + a_N N)$$

$$= v a_N (T \times N)$$

$$= \boxed{v a_N B = \alpha' \times \alpha''}$$

$\|\alpha' \times \alpha''\| = \| v a_N B \| = |v a_N| = v a_N$ by def$^{\underline{n}}$'s of v and a_N.

$$\boxed{\frac{\alpha' \times \alpha''}{\|\alpha' \times \alpha''\|} = B}$$

$$\boxed{\frac{\alpha'}{\|\alpha'\|} = T}$$

$$\boxed{B \times T = \frac{1}{\|\alpha' \times \alpha''\|} \frac{1}{\|\alpha'\|} \left[(\alpha' \times \alpha'') \times \alpha' \right] = N}$$

Dot previous relation with B which kills N and T parts as $N, T \perp B$.

$$\alpha''' \cdot B = a_N v \tau$$

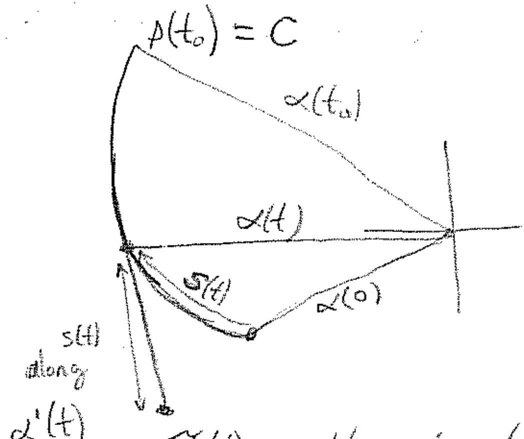
$$a_N \tau = \alpha''' \cdot \frac{\alpha' \times \alpha''}{\|\alpha' \times \alpha''\|}$$

$$\|\alpha' \times \alpha''\| = a_N V$$

$$a_N \tau = \frac{1}{\|\alpha' \times \alpha''\|} \alpha''' \cdot (\alpha' \times \alpha'')$$

$$\tau = \frac{1}{V a_N \|\alpha' \times \alpha''\|} \alpha''' \cdot (\alpha' \times \alpha'')$$

$$\boxed{\tau = \frac{\alpha''' \cdot (\alpha' \times \alpha'')}{\|\alpha' \times \alpha''\|^2}}$$



$0 \leq t \leq t_0$

$I(t)$ = the involute of $\alpha(t)$.

4.6 Keplers' laws of planetary motion

KEPLER'S LAWS OF PLANETARY MOTION (284)

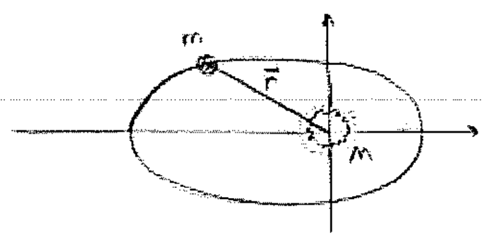
In antiquity there have been radically different views of the universe at large and the motion or lack of motion of the earth through it. At the time of Kepler the heliocentric view of Copernicus (1473–1543) had taken hold, but astronomers insisted that planets traveled in circles, then circles on top of circles on top of circles... This system of "perfect" circles were known as epicycles. Epicycles worked quite well but Kepler (1571–1630) found them unnatural. Kepler instead thought he could explain the motion of planets by a few simple rules. He found these rules empirically by studying the exquisite data taken by Tycho Brahe. These laws were chosen simply to fit the data. Only later were these laws derived from basic physical law. By the way, much of modern physics are still like Kepler's Laws, it is always the dream/goal/aspiration to derive known phenomenological law from basic principles. There is some controversy as to who first derived Kepler's Laws, many credit Newton himself others credit Johann Bernoulli in 1710. The incredible thing is that we can derive the laws in a few short pages. Our notation and understanding of vector calculus is several hundred years in advance, so ordinary folks like myself can grasp the proof.

Set-up

Keplers laws for the Sun and a single planet are:

1.) The orbit of the planet is elliptical with the sun at a focus.
2.) During equal times the planet sweeps out equal areas in the ellipse.
3.) $T^2 \propto a^3$ where T = period of planet's orbit, a = length of semimajor axis of ellipse.

We place the origin at the sun. We expect that



• My proof of Kepler's Laws follows Colley's of §3.1 fairly closely. (285)

Proposition: The motion of the planet lies in a plane which also contains the sun if we assume Newton's Universal Law of Gravitation governs the motion through Newton's Laws.

Proof: our goal is to show that $\vec{r} \times \vec{V} = \vec{C}$ for some constant vector $\vec{C}$. This will show that planet moves in a plane with normal $\vec{C}$. Note,

$$\frac{d}{dt}\left(\vec{r} \times \vec{V}\right) = \underbrace{\frac{d\vec{r}}{dt} \times \vec{V}}_{\vec{V} \times \vec{V} = 0.} + \vec{r} \times \frac{d\vec{V}}{dt} = \vec{r} \times \vec{a}.$$

Recall in our current notation that $\vec{r} = r\hat{r}$ and Newton tells us that,

$$\vec{F} = m\vec{a} = -\frac{GmM}{r^2}\hat{r} = -\frac{GmM}{r^3}\vec{r}$$

m = mass of planet
M = mass of sun
G = Gravitational Constant.

$$\therefore \quad \vec{a} = \frac{-GM}{r^2}\vec{r} \quad \text{thus } \vec{a} \parallel \vec{r}$$

$$\Rightarrow \vec{a} \times \vec{r} = 0 \quad \Rightarrow \frac{d}{dt}\left(\vec{r} \times \vec{V}\right) = \vec{r} \times \vec{a} = 0 \quad \therefore \ \underline{\vec{r} \times \vec{V} = \vec{C}}$$

Th^m/ Kepler's 1^st Law: The planet's orbit is an ellipse with sun at one focus.

Proof: this will take a little work so be patient, let's get a better hold on $\vec{C}$,

$$\vec{V} = \frac{d\vec{r}}{dt} = \frac{d}{dt}\left(r\hat{r}\right) = \frac{dr}{dt}\hat{r} + r\frac{d\hat{r}}{dt} = \dot{r}\hat{r} + r\frac{d\hat{r}}{dt}$$

Apply this to the following,

$$\vec{C} = \vec{r} \times \vec{V} = r\hat{r} \times \left[\cancelto{0}{\dot{r}\hat{r}} + r\frac{d\hat{r}}{dt}\right] = \underbrace{r^2\,\hat{r} \times \frac{d\hat{r}}{dt} = \vec{C}}_{} \quad (I)$$

Calculate then, using (I)

$$\begin{aligned}
\vec{a} \times \vec{C} &= \left(\frac{-GM}{r^2}\hat{r}\right) \times \left(r^2\,\hat{r} \times \frac{d\hat{r}}{dt}\right) \\
&= -GM\left[\hat{r} \times \left(\hat{r} \times \frac{d\hat{r}}{dt}\right)\right] \\
&= GM\left[\left(\hat{r} \times \frac{d\hat{r}}{dt}\right) \times \hat{r}\right] \quad : \text{recall } \overbrace{A \times (B \times C) = (A\cdot C)B - (A \cdot B)C}^{\text{see §9.4 \#30}} \\
&= GM\left[(\hat{r}\cdot\hat{r})\frac{d\hat{r}}{dt} - \underbrace{\left(\hat{r}\cdot\frac{d\hat{r}}{dt}\right)}_{0}\hat{r}\right] \quad : \ \hat{r}\cdot\hat{r} = 1 \ \Rightarrow \ 2\hat{r}\cdot\dot{\hat{r}} = 0 \Rightarrow \hat{r}\cdot\dot{\hat{r}} = 0. \\
&= GM\,\frac{d\hat{r}}{dt} \\
&= \underbrace{\frac{d}{dt}\left(GM\hat{r}\right) = \vec{a} \times \vec{C}}_{} \quad (II)
\end{aligned}$$

286

Proof of Kepler's 1st Law continued

We may derive another identity for $\vec{a}\times\vec{c}$,

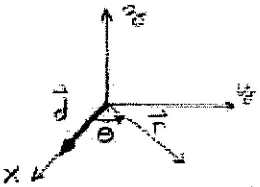
$$\vec{a}\times\vec{c} = \frac{d\vec{V}}{dt}\times\vec{c} + \vec{V}\times\frac{d\vec{c}}{dt} \quad \text{: added zero since } \frac{d\vec{c}}{dt} = 0.$$

$$= \underbrace{\frac{d}{dt}\left[\vec{V}\times\vec{c}\right]}_{\text{III}} \quad \text{: using identity (K) on 265}$$

Thus comparing (II) & (III) we find

$$\frac{d}{dt}\left(GM\hat{r}\right) = \frac{d}{dt}\left(\vec{V}\times\vec{c}\right) \quad \therefore \quad \underbrace{\vec{V}\times\vec{c} = GM\hat{r} + \vec{d}}_{\text{IV}}$$

where $\vec{d}$ is a constant vector, it lies in the orbital plane since $\vec{V}\times\vec{c}$ and $\hat{r}$ do. Now chose coordinates in the orbital plane so that $\vec{d}$ lines up with the x-axis. Let θ be the usual θ in the xy-plane,



$$\hat{r}\cdot\vec{d} = |\hat{r}||\vec{d}|\cos\theta = d\cos\theta$$

where $|\vec{d}| = d$ in our notation here.

Now consider the length of $\vec{c}$ squared,

$$\begin{aligned}
c^2 &= \vec{c}\cdot\vec{c} \\
&= (\vec{r}\times\vec{V})\cdot\vec{c} \\
&= \vec{r}\cdot(\vec{V}\times\vec{c}) \quad \text{: using identity (V) of 248} \\
&= r\hat{r}\cdot\left[GM\hat{r} + \vec{d}\right] \quad \text{: using IV. we found just above.} \\
&= GMr + r\hat{r}\cdot\vec{d} \\
&= GMr + rd\cos\theta \\
&= r(GM + d\cos\theta)
\end{aligned}$$

Therefore we solve for $r = \sqrt{x^2+y^2+z^2} = \sqrt{x^2+y^2}$ (we're in $z=0$) and obtain the eq$^{\underline{n}}$ of an ellipse (or parabola or hyperbola)

$$r = \frac{c^2}{GM + d\cos\theta} = \frac{c^2/GM}{1+(d/GM)\cos\theta} = \boxed{\frac{p}{1+e\cos\theta} = r}$$

where we define $p = c^2/GM$ and the eccentricity $e = d/GM$. This is an ellipse in polar coordinates. Since you've likely not seen that recently (or maybe never) we'll connect to ⟶

Proof of Kepler's 1st Law continued 287

the usual Cartesian eq^{n}'s for the ellipse. The details will be of use to us in proving the 3rd Law of Kepler later on.

$$r = \frac{p}{1+e\cos\theta} \Rightarrow r = p - er\cos\theta$$

Trying to convert the polar coordinates (r,θ) to (x,y) where $x = r\cos\theta$ and $y = r\sin\theta$. We see, using $x = r\cos\theta$

$$r = p - ex$$

$$r^2 = x^2 + y^2 = p^2 - 2epx + e^2x^2$$

$$x^2 - e^2x^2 + y^2 + 2epx = p^2$$

$$x^2(1-e^2) + 2epx + y^2 = p^2$$

$$x^2 + \frac{2ep}{1-e^2}x + \frac{y^2}{1-e^2} = \frac{p^2}{1-e^2} \quad : \quad \underline{\text{assume } e \neq \pm 1}$$

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

$$\therefore \boxed{\frac{\left(x - \frac{ep}{1-e^2}\right)^2}{p^2/(1-e^2)^2} + \frac{y^2}{p^2/(1-e^2)} = 1} \quad \underset{(0<e<1)}{\text{ellipse}} \text{ or } \underset{(e>1)}{\text{hyperbola}}$$

This is an ellipse with center $(ep/(1-e^2), 0)$ and it has semi-major axis length $a = p/(1-e^2)$ and semi-minor axis $b = p/\sqrt{1-e^2}$.
Remark: recall that we defined $p = c^2/GM$ so $p > 0$ and we need not worry about x by p. Now $e = d/GM > 0$ so we can rule out $e = -1$ as a problem. Notice we have divided by $\sqrt{1-e^2}$ as part of our sol^{n}, this only makes sense if $0 < e < 1$. The case $e = 1$ needs seperate treatment. Motion in the case $0 < e < 1$ is that of planets.

$\underline{e = 1)}$ $r = p - r\cos\theta \quad \therefore \quad r^2 = (p-x)^2 = p^2 - 2xp + x^2$

that is $x^2 + y^2 = p^2 - 2xp + x^2 \Rightarrow 2xp = p^2 - y^2$

$$\therefore \boxed{x = p/2 - y^2/2p} \quad \text{parabola}$$

Remark: One nice resource for background on conic-sections and polar coordinates is "Precalculus, Concepts through functions" Sullivan & Sullivan. There is just about all the cases you can imagine, rotated ellipses for example.

288

Th^m/ Kepler's 2nd Law: During equal times a planet sweeps through equal areas.

Proof: Pick a point P_o at angle θ_o. The later in the course we will learn that the area in polar coordinates swept by the region from θ_o to θ is simply

$$A(\theta) = \int_{\theta_o}^{\theta} \frac{1}{2} r^2 \, d\beta$$

We seek to show that $\frac{dA}{dt}$ = constant. Consider then

$$\frac{dA}{d\theta} = \frac{d}{d\theta}\int_{\theta_o}^{\theta} \frac{1}{2} r^2 \, d\beta = \frac{1}{2} r^2 \qquad \text{by F.T.C.}$$

Then the chain rule tells us

$$\frac{dA}{dt} = \frac{dA}{d\theta}\frac{d\theta}{dt} = \frac{1}{2} r^2 \frac{d\theta}{dt}$$

Notice that $\hat{r} = \langle \cos\theta, \sin\theta \rangle$ thus diff. implicitly, remember $\theta = \theta(t)$.

$$\frac{d\hat{r}}{dt} = \langle -\sin\theta, \cos\theta \rangle \frac{d\theta}{dt} = \langle -\sin\theta, \cos\theta, 0 \rangle \frac{d\theta}{dt} \qquad \text{(we've been suppressing the z-comp.)}$$

(E5), (D) $\Rightarrow$ $\vec{C} = r^2\left(\hat{r} \times \frac{d\hat{r}}{dt}\right) = r^2 \frac{d\theta}{dt} \langle \cos\theta, \sin\theta, 0 \rangle \times \langle -\sin\theta, \cos\theta, 0 \rangle$

$$\vec{C} = r^2 \frac{d\theta}{dt} \langle 0, 0, 1 \rangle \qquad \therefore \quad C = r^2 \frac{d\theta}{dt}$$

Hence $\frac{dA}{dt} = \frac{1}{2} r^2 \frac{d\theta}{dt} = \frac{C}{2}$ = constant. //

Th^m/ Kepler's 3rd Law: $T^2 = Ka^3$ where T is the orbital period and a is the length of the semimajor axis, K = some constant

Proof: I proved back on pg. 138 in E7 that the area of an ellipse is $A = \pi ab$. On the other hand we could say that $dA = \frac{dA}{dt} dt$ and integrate over a whole orbit to find

$$\pi ab = \int_0^T \frac{dA}{dt} dt = \int_0^T \frac{C}{2} dt = \frac{CT}{2} \quad \therefore \quad T = \frac{2\pi ab}{C} \quad \therefore \quad T^2 = \frac{4\pi^2 a^2 b^2}{C^2}$$

Notice that $a^2 = p^2/(1-e^2)^2$ and $b^2 = p^2/(1-e^2)$, also $C^2 = GMp$.

$$T^2 = \frac{4\pi^2}{GMp} \frac{p^2}{(1-e^2)^2} \cdot \frac{p^2}{(1-e^2)} = \frac{4\pi^2}{GM}\left(\frac{p}{1-e^2}\right)^3 = \boxed{\frac{4\pi^2 a^3}{GM} = T^2} \quad //.$$

It is interesting that $K = \frac{4\pi^2}{GM}$ is independent of the planets mass. all the planets orbit under the same K-value.

<u>Remark</u>: There is another method of proving Kepler's Laws that begins with the two-body Lagrangian for a central potential (well force really but $\vec{F} = f(r)\hat{r} \Rightarrow V = V(r) \ldots$). In that derivation one need not assume the sun is at the origin. Instead you consider the center of mass to be at the origin and work out how the reduced mass μ orbits. Anyway its very beautiful, take Mechanics at the Junior/Senior level to see the more general derivation. Also they will actually find $r(t)$ explicitly as opposed to the indirect arguments we have offered (or rather stolen from Colley ☺.)

Chapter 5

Euclidean structures and physics

Although much was known about the physical world prior to Newton that knowledge was highly unorganized and formulated in such a way that is was difficult to use and understand[1]. The advent of Newton changed all that. In 1665-1666 Newton transformed the way people thought about the physical world, years later he published his many ideas in "Principia mathematica philosphiae naturalia" (1686). His contribution was to formulate three basic laws or principles which along with his universal law of gravitation would prove sufficient to derive and explain all mechanical systems both on earth and in the heavens known at the time. These basic laws may be stated as follows:

1. **Newton's First Law:** Every particle persists in its state of rest or of uniform motion in a straight line unless it is compelled to change that state by impressed forces.

2. **Newton's Second Law:** The rate of change of motion is proportional to the motive force impressed; and is made in the direction of the straight line in which that force is impressed.

3. **Newton's Third Law:** To every action there is an equal reaction; or the mutual actions of two bodies upon each other are always equal but oppositely directed.

Until the early part of the last century Newton's laws proved adequate. We now know, however that they are only accurate within prescribed limits. They do not apply for things that are very small like an atom or for things that are very fast like cosmic rays or light itself. Nevertheless Newton's laws are valid for the majority of our common macroscopic experiences in everyday life.

[1]What follows is borrowed from Chapter 6 of my *Mathematical Models in Physics* notes which is turn borrowed from my advisor Dr. R.O. Fulp's notes for Math 430 at NCSU. I probably will not cover all of this in lecture but I thought it might be interesting to those of you who are more physically minded. I have repeated some mathematical definitions in this chapter in the interest of making this chapter more readable. This chapter gives you an example of the practice of Mathematical Physics. One common idea in Mathematical Physics is to take known physics and reformulate it in a proper mathematical context. Physicists don't tend to care about domains or existence so if we are to understand their calculations then we need to do some work in most cases.

It is implicitly presumed in the formulation of Newton's laws that we have a concept of a straight line, of uniform motion, of force and the like. Newton realized that Euclidean geometry was a necessity in his model of the physical world. In a more critical formulation of Newtonian mechanics one must address the issues implicit in the above formulation of Newton's laws. This is what we attempt in this chapter, we seek to craft a mathematically rigorous systematic statement of Newtonian mechanics.

5.1 Euclidean geometry

Note: we abandon the more careful notation of the previous chapters in what follows. In a nutshell we are setting $\mathbb{R}^3 = V^3$, this is usually done in physics. We can identify a given point with a vector that eminates from the origin to the point in question. It will be clear from the context if a point or a vector is intended.

Nowadays Euclidean geometry is imposed on a vector space via an inner product structure. Let $x_1, x_2, x_3, y_1, y_2, y_3, c \in \mathbb{R}$. As we discussed $\mathbb{R}^3$ is the set of 3-tuples and it is a vector space with respect to the operations,

$$(x_1, x_2, x_3) + (y_1, y_2, y_3) = (x_1 + y_1, x_2 + y_2, x_3 + y_3)$$

$$c(x_1, x_2, x_3) = (cx_1, cx_2, cx_3)$$

where $x_1, x_2, x_3, y_1, y_2, y_3, c \in \mathbb{R}$. Also we have the dot-product,

$$(x_1, x_2, x_3) \cdot (y_1, y_2, y_3) = x_1y_1 + x_2y_2 + x_3y_3$$

from which the *length* of a vector $x = (x_1, x_2, x_3) \in \mathbb{R}^3$ can be calculated,

$$|x| = \sqrt{x \cdot x} = \sqrt{x_1^2 + x_2^2 + x_3^2}$$

meaning $|x|^2 = x \cdot x$. Also if $x, y \in \mathbb{R}^3$ are nonzero vectors then the angle between them is defined by the formula,

$$\theta = \cos^{-1}\left(\frac{x \cdot y}{|x||y|}\right)$$

In particular nonzero vectors x and y are *perpendicular or orthogonal* iff $\theta = 90^o$ which is so iff $\cos(\theta) = 0$ which is turn true iff $x \cdot y = 0$.

Definition 5.1.1.

A function $L : \mathbb{R}^3 \to \mathbb{R}^3$ is said to be a **linear transformation** if and only if there is a 3×3 matrix A such that $L(x) = Ax$ for all $x \in \mathbb{R}^3$. Here Ax indicates multiplication by the matrix A on the column vector x

Definition 5.1.2.

An **orthogonal transformation** is a linear transformation $L : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ which satisfies

$$L(x) \cdot L(y) = x \cdot y$$

for all $x, y \in \mathbb{R}^3$. Such a transformation is also called an **linear isometry of the Euclidean metric.**

The term *isometry* means the same measure, you can see why that's appropriate from the following,

$$|L(x)|^2 = L(x) \cdot L(x) = x \cdot x = |x|^2$$

for all $x \in \mathbb{R}^3$. Taking the square root of both sides yields $|L(x)| = |x|$; an orthogonal transformation preserves the lengths of vectors in $\mathbb{R}^3$. Using what we just learned its easy to show orthogonal transformations preserve angles as well,

$$\cos(\theta_L) = \frac{L(x) \cdot L(y)}{|L(x)||L(y)|} = \frac{x \cdot y}{|x||y|} = \cos(\theta)$$

Hence taking the inverse cosine of each side reveals that the angle θ_L between $L(x)$ and $L(y)$ is equal to the angle θ between x and y; $\theta_L = \theta$. Orthogonal transformations preserve angles.

Definition 5.1.3.

We say $l \subseteq \mathbb{R}^3$ is a **line** if there exist $a, v \in \mathbb{R}^3$ such that

$$l = \{x \in \mathbb{R}^{\,n \times n} \mid x = a + tv, \quad t \in \mathbb{R}\}.$$

Proposition 5.1.4.

If L is an orthonormal transformation then $L(l)$ is also a line in $\mathbb{R}^3$.

To prove this we simply need to find new a' and v' in $\mathbb{R}^3$ to demonstrate that $L(l)$ is a line. Take a point on the line, $x \in l$

$$\begin{aligned} L(x) &= L(a + tv) \\ &= L(a) + tL(v) \end{aligned} \tag{5.1}$$

thus $L(x)$ is on a line described by $x = L(a) + tL(v)$, so we can choose $a' = L(a)$ and $v' = L(v)$ it turns out; $L(l) = \{x \in \mathbb{R}^3 \mid x = a' + tv'\}$.

If one has a coordinate system with unit vectors $\hat{\imath}, \hat{\jmath}, \hat{k}$ along three mutually orthogonal axes then an orthogonal transformation will create three new mutually orthogonal unit vectors $L(\hat{\imath}) = \hat{\imath}', L(\hat{\jmath}) = \hat{\jmath}', L(\hat{k}) = \hat{k}'$ upon which one could lay out new coordinate axes. In this way orthogonal transformations give us a way of constructing new "rotated" coordinate systems from a given coordinate system. Moreover, it turns out that Newton's laws are preserved (have the same form) under orthogonal transformations. Transformations which are not orthogonal can greatly distort the form of Newton's laws.

Remark 5.1.5.

If we view vectors in $\mathbb{R}^3$ as column vectors then the dot-product of x with y can be written as $x \cdot y = x^T y$ for all $x, y \in \mathbb{R}^3$. Recall that x^T is the *transpose* of x, it changes the column vector x to the corresponding row vector x^T.

Let us consider an orthogonal transformation $L : \mathbb{R}^3 \to \mathbb{R}^3$ where $L(x) = Ax$. What condition on the matrix A follows from the the L being an orthogonal transformation ?

$$\begin{aligned} L(x) \cdot L(y) = x \cdot y \quad &\Longleftrightarrow (Ax)^T(Ay) = x^T y \\ &\Longleftrightarrow x^T(A^T A)y = x^T y \\ &\Longleftrightarrow x^T(A^T A)y = x^T I y \\ &\Longleftrightarrow x^T(A^T A - I)y = 0. \end{aligned} \tag{5.2}$$

But $x^T(A^T A - I)y = 0$ for *all* $x, y \in \mathbb{R}^3$ iff $A^T A - I = 0$ or $A^T A = I$. Thus L is orthogonal iff its matrix A satisfies $A^T A = I$. This is in turn equivalent to A having an inverse and $A^{-1} = A^T$.

Proposition 5.1.6.

The set of **orthogonal transformations** on $\mathbb{R}^3$ is denoted O(3). The operation of function composition on O(3) makes it a group. Likewise we also denote the set of all **orthogonal matrices** by O(3),

$$O(3) = \{A \in \mathbb{R}^{3\times 3} \mid A^T A = I\}$$

it is also a group under matrix multiplication.

Usually we will mean the matrix version, it should be clear from the context, it's really just a question of notation since we know that L and A contain the same information thanks to linear algebra. Recall that every linear transformation L on a finite dimensional vector space can be represented by matrix multiplication of some matrix A.

Proposition 5.1.7.

The set of **special orthogonal matrices** on $\mathbb{R}^3$ is denoted SO(3),

$$SO(3) = \{A \in \mathbb{R}^{3\times 3} \mid A^T A = I \text{ and } det(A) = 1\}$$

it is also a group under matrix multiplication and thus it is a subgroup of O(3). It is shown in standard linear algebra course that every special orthogonal matrix rotates $\mathbb{R}^3$ about some line. Thus, we will often refer to SO(3) as the **group of rotations**.

There are other transformations that do not change the geometry of $\mathbb{R}^3$.

Definition 5.1.8.

A **translation** is a function $T : \mathbb{R}^3 \to \mathbb{R}^3$ defined by $T(x) = x + v$ where v is some fixed vector in $\mathbb{R}^3$ and x is allowed to vary over $\mathbb{R}^3$.

Clearly translations do not change the distance between two points $x, y \in \mathbb{R}^3$,

$$|T(x) - T(y)| = |x + v - (y - v)| = |x - y| = \text{ distance between } x \text{ and } y.$$

Also if x, y, z are points in $\mathbb{R}^3$ and θ is the angle between $y - x$ and $z - x$ then θ is also the angle between $T(y) - T(x)$ and $T(z) - T(x)$. Geometrically this is trivial, if we shift all points by the same vector then the difference vectors between points are unchanged thus the lengths and angles between vectors connecting points in $\mathbb{R}^3$ are unchanged.

Definition 5.1.9.

A function $\phi : \mathbb{R}^3 \to \mathbb{R}^3$ is called a **rigid motion** if there exists a vector $r \in \mathbb{R}^3$ and a rotation matrix $A \in SO(3)$ such that $\phi(x) = Ax + r$.

A rigid motion is the composite of a translation and a rotation therefore it will clearly preserve lengths and angles in $\mathbb{R}^3$. So rigid motions are precisely those transformations which preserve Euclidean geometry and consequently they are the transformations which will preserve Newton's laws. If Newton's laws hold in one coordinate system then we will find Newton's laws are also valid in a new coordinate system iff it is related to the original coordinate system by a rigid motion. We now proceed to provide a careful exposition of the ingredients needed to give a rigorous formulation of Newton's laws.

Definition 5.1.10.

We say that $\mathcal{E}$ is an **Euclidean structure** on a set S iff $\mathcal{E}$ is a family of bijections from S onto $\mathbb{R}^3$ such that,
(1.) $\mathcal{X}, \mathcal{Y} \in \mathcal{E}$ then $\mathcal{X} \circ \mathcal{Y}^{-1}$ is a rigid motion.
(2.) if $\mathcal{X} \in \mathcal{E}$ and ϕ is a rigid motion then $\phi \circ \mathcal{X} \in \mathcal{E}$.
Also a **Newtonian space** is an ordered pair $(S, \mathcal{E})$ where S is a set and $\mathcal{E}$ is an Euclidean structure on S.

Notice that if $\mathcal{X}, \mathcal{Y} \in \mathcal{E}$ then there exists an $A \in SO(3)$ and a vector $r \in \mathbb{R}^3$ so that we have $\mathcal{X}(p) = A\mathcal{Y}(p) + r$ for every $p \in S$. Explicitly in cartesian coordinates on $\mathbb{R}^3$ this means,

$$[\mathcal{X}_1(p), \mathcal{X}_2(p), \mathcal{X}_3(p)]^T = A[\mathcal{Y}_1(p), \mathcal{Y}_2(p), \mathcal{Y}_3(p)]^T + [r_1, r_2, r_3]^T.$$

Newtonian space is the mathematical model of space which is needed in order to properly formulate Newtonian mechanics. The first of Newton's laws states that an object which is subject to no forces must move along a straight line. This means that some observer should be able to show that the object moves along a line in space. We take this to mean that the observer chooses an inertial frame and makes measurements to decide wether or not the object executes straight line motion in the coordinates defined by that frame. If the observations are to be frame independent then the notion of a straight line in space should be independent of which inertial coordinate system is used to make the measurements. We intend to identify inertial coordinate systems as precisely those elements of $\mathcal{E}$. Thus we need to show that if l is a line as measured by $\mathcal{X} \in \mathcal{E}$ then l is also a line as measured by $\mathcal{Y} \in \mathcal{E}$.

Definition 5.1.11.

Let $(S, \mathcal{E})$ be a Newtonian space. A subset l of S is said to be **a line in S** iff $\mathcal{X}(l)$ is a line in $\mathbb{R}^3$ for some choice of $\mathcal{X} \in \mathcal{E}$.

The theorem below shows us that the choice made in the definition above is not special. In fact our definition of a line in S is coordinate independent. Mathematicians almost always work towards formulating geometry in a way which is independent of the coordinates employed, this is known as the coordinate free approach. Physicists in contrast almost always work in coordinates.

Theorem 5.1.12.

If l is a line in a Newtonian space $(S, \mathcal{E})$ then $\mathcal{Y}(l)$ is a line in $\mathbb{R}^3$ for **every** $\mathcal{Y} \in \mathcal{E}$.

Proof: Because l is a line in the S we know there exists $\mathcal{X} \in \mathcal{E}$ and $\mathcal{X}(l)$ is a line in $\mathbb{R}^3$. Let $\mathcal{Y} \in \mathcal{E}$ observe that,

$$\mathcal{Y}(l) = (\mathcal{Y} \circ \mathcal{X}^{-1} \circ \mathcal{X})(l) = (\mathcal{Y} \circ \mathcal{X}^{-1})(\mathcal{X}(l)).$$

Now since $\mathcal{X}, \mathcal{Y} \in \mathcal{E}$ we have that $\mathcal{Y} \circ \mathcal{X}^{-1}$ is a rigid motion on $\mathbb{R}^3$. Thus if we can show that rigid motions take lines to lines in $\mathbb{R}^3$ the proof will be complete. We know that there exist $A \in SO(3)$ and $r \in \mathbb{R}^3$ such that $(\mathcal{Y} \circ \mathcal{X}^{-1})(x) = Ax + r$. Let $x \in \mathcal{X}(l) = \{x \in \mathbb{R}^3 \mid x = p + tq \quad t \in \mathbb{R}$ and p,q are fixed vectors in $\mathbb{R}^3\}$, consider

$$\begin{aligned} (\mathcal{Y} \circ \mathcal{X}^{-1})(x) &= Ax + r \\ &= A(p + tq) + r \\ &= (Ap + r) + tAq \\ &= p' + tq' \qquad \text{letting } p' = Ap + r \text{ and } q' = Aq. \end{aligned} \tag{5.3}$$

The above hold for all $x \in \mathcal{X}(l)$, clearly we can see the line has mapped to a new line $\mathcal{Y}(l) = \{x \in \mathbb{R}^3 \mid x = p' + tq' \quad, t \in \mathbb{R}\}$. Thus we find what we had hoped for, lines are independent of the frame chosen from $\mathcal{E}$ in the sense that a line is always a line no matter which element of $\mathcal{E}$ describes it.

Definition 5.1.13.

An **observer** is a function from an interval $I \subset \mathbb{R}$ into $\mathcal{E}$. We think of such a function $\mathcal{X} : I \to \mathcal{E}$ as being a time-varying coordinate system on S. For each $t \in I$ we denote $\mathcal{X}(t)$ by $\mathcal{X}_t$; thus $\mathcal{X}_t : S \to \mathbb{R}^3$ for each $t \in I$ and $\mathcal{X}_t(p) = [\mathcal{X}_{t1}(p), \mathcal{X}_{t2}(p), \mathcal{X}_{t3}(p)]$ for all $p \in S$.

Assume that a material particle or more generally a "point particle" moves in space S in such a way that at time t the particle is centered at the point $\gamma(t)$. Then the mapping $\gamma : I \to S$ will be called the **trajectory** of the particle.

Definition 5.1.14.

Let us consider a particle with trajectory $\gamma : I \to S$. Further assume we have an observer $\mathcal{X} : I \to \mathcal{E}$ with $t \mapsto \mathcal{X}_t$ then:

(1.) $\mathcal{X}_t(\gamma(t))$ is the **position vector** of the particle at time $t \in I$ relative to the observer $\mathcal{X}$.

(2.) $\frac{d}{dt}\big[\mathcal{X}_t(\gamma(t))\big]\big|_{t=t_o}$ is called the **velocity** of the particle at time $t_o \in I$ relative to the observer $\mathcal{X}$, it is denoted $v_{\mathcal{X}}(t_o)$.

(3.) $\frac{d^2}{dt^2}\big[\mathcal{X}_t(\gamma(t))\big]\big|_{t=t_o}$ is called the **acceleration** of the particle at time $t_o \in I$ relative to the observer $\mathcal{X}$, it is denoted $a_{\mathcal{X}}(t_o)$.

Notice that position, velocity and acceleration are only defined with respect to an observer. We now will calculate how position, velocity and acceleration of a particle with trajectory $\gamma : I \to S$ relative to observer $\mathcal{Y} : I \to \mathcal{E}$ compare to those of another observer $\mathcal{X} : I \to \mathcal{E}$. To begin we note that each particular $t \in I$ we have $\mathcal{X}_t, \mathcal{Y}_t \in \mathcal{E}$ thus there exists a rotation matrix $A(t) \in SO(3)$ and a vector $v(t) \in \mathbb{R}^3$ such that,

$$\mathcal{Y}_t(p) = A(t)\mathcal{X}_t(p) + r(t)$$

for all $p \in S$. As we let t vary we will in general find that $A(t)$ and $r(t)$ vary, in other words we have A a matrix-valued function of time given by $t \mapsto A(t)$ and r a vector-valued function of time given by $t \mapsto r(t)$. Also note that the *origin* of the coordinate coordinate system $\mathcal{X}(p) = 0$ moves to $\mathcal{Y}(p) = r(t)$, this shows that the correct interpretation of $r(t)$ is that it is the position of the old coordinate's origin in the new coordinate system. Consider then $p = \gamma(t)$,

$$\boxed{\mathcal{Y}_t(\gamma(t)) = A(t)\mathcal{X}_t(\gamma(t)) + r(t)} \tag{5.4}$$

this equation shows how the position of the particle in $\mathcal{X}$ coordinates transforms to the new position in $\mathcal{Y}$ coordinates. We should not think that the particle has moved under this transformation, rather we have just changed our viewpoint of where the particle resides. Now move on to the transformation of velocity, (we assume the reader is familiar with differentiating matrix valued functions of a real variable, in short we just differentiate component-wise)

$$\begin{aligned} v_{\mathcal{Y}}(t) &= \tfrac{d}{dt}\big[\mathcal{Y}(\gamma(t))\big] \\ &= \tfrac{d}{dt}\big[A(t)\mathcal{X}_t(\gamma(t)) + r(t)\big] \\ &= \tfrac{d}{dt}[A(t)]\mathcal{X}_t(\gamma(t)) + A(t)\tfrac{d}{dt}[\mathcal{X}_t(\gamma(t))] + \tfrac{d}{dt}[r(t)] \\ &= A'(t)\mathcal{X}_t(\gamma(t)) + A(t)v_{\mathcal{X}}(t) + r'(t). \end{aligned} \tag{5.5}$$

Recalling the dot notation for time derivatives and introducing $\gamma_{\mathcal{X}} = \mathcal{X} \circ \gamma$,

$$\boxed{v_{\mathcal{Y}} = \dot{A}\gamma_{\mathcal{X}} + Av_{\mathcal{X}} + \dot{r}.} \tag{5.6}$$

We observe that the velocity according to various observes depends not only on the trajectory itself, but also the time evolution of the observer itself. The case $A = I$ is more familiar, since $\dot{A} = 0$ we have,

$$v_{\mathcal{Y}} = Iv_{\mathcal{X}} + \dot{r} = v_{\mathcal{X}} + \dot{r}. \tag{5.7}$$

The velocity according to the observer $\mathcal{Y}$ moving with velocity $\dot{r}$ relative to $\mathcal{X}$ is the sum of the velocity according to $\mathcal{X}$ and the velocity of the observer $\mathcal{Y}$. Obviously when $A \neq I$ the story is more complicated, but the case $A = I$ should be familiar from freshman mechanics.
Now calculate how the accelerations are connected,

$$\begin{aligned} a_{\mathcal{Y}}(t) &= \tfrac{d^2}{dt^2}[\mathcal{Y}(\gamma(t))] \\ &= \tfrac{d}{dt}[A'(t)\mathcal{X}_t(\gamma(t)) + A(t)v_{\mathcal{X}}(t) + r'(t)] \\ &= A''(t)\mathcal{X}_t(\gamma(t)) + A'(t)\tfrac{d}{dt}[\mathcal{X}_t(\gamma(t))] + A'(t)v_{\mathcal{X}}(t) + A(t)\tfrac{d}{dt}[v_{\mathcal{X}}(t)] + r''(t) \\ &= A''(t)\mathcal{X}_t(\gamma(t)) + 2A'(t)v_{\mathcal{X}}(t) + +A(t)a_{\mathcal{X}}(t) + r''(t) \end{aligned} \tag{5.8}$$

Therefore we relate acceleration in $\mathcal{X}$ to the acceleration in $\mathcal{Y}$ as follows,

$$\boxed{a_{\mathcal{Y}} = Aa_{\mathcal{X}} + \ddot{r} + \ddot{A}\gamma_{\mathcal{X}} + 2\dot{A}v_{\mathcal{X}}.} \tag{5.9}$$

The equation above explains many things, if you take the junior level classical mechanics course you'll see what those things are. This equation does not look like the one used in mechanics for noninertial frames, it is nevertheless the same and if you're interested I'll show you.

Example 5.1.15. ..

(Examples added at end of Chapter)

Example 5.1.16. ..

Definition 5.1.17.

If $\gamma : I \to S$ is the trajectory of a particle then we say the particle and $\mathcal{X} : I \to \mathcal{E}$ is an observer. We say the particle is in a **state of rest** relative to the observer $\mathcal{X}$ iff $v_{\mathcal{X}} = \frac{d}{dt}[\mathcal{X}_t(\gamma(t))] = 0$. We say the particle experiences **uniform rectilinear motion** relative to the observer $\mathcal{X}$ iff $t \mapsto \mathcal{X}_t(\gamma(t))$ is a straight line in $\mathbb{R}^3$ with velocity vector some nonzero constant vector.

We now give a rigorous definition for the existence of *force*, a little later we'll say how to calculate it.

Definition 5.1.18.

A particle **experiences a force** relative to an observer $\mathcal{X}$ iff the particle is neither in a state of rest nor is it in uniform rectilinear motion relative to $\mathcal{X}$. Otherwise we say the particle experiences no force relative to $\mathcal{X}$.

Definition 5.1.19.

An observer $\mathcal{X} : I \to \mathcal{E}$ is said to be an **inertial observer** iff there exists $\mathcal{X}_o \in \mathcal{E}$, $A \in SO(3)$, $v, w \in \mathbb{R}^3$ such that $\mathcal{X}_t = A\mathcal{X}_o + tv + w$ for all $t \in I$. A particle is called a **free particle** iff it experiences no acceleration relative to an inertial observer.

Observe that a constant mapping into $\mathcal{E}$ is an inertial observer and that general inertial observers are observers which are in motion relative to a "stationary observer" but the motion is "constant velocity" motion. We will refer to a constant mapping $\mathcal{X} : I \to \mathcal{E}$ as a **stationary observer**.

Theorem 5.1.20.

If $\mathcal{X} : I \to \mathcal{E}$ and $\mathcal{Y} : I \to \mathcal{E}$ are inertial observers then there exists $A \in SO(3)$, $v, w \in \mathbb{R}^3$ such that $\mathcal{Y}_t = A\mathcal{X}_t + tv + w$ for all $t \in I$. Moreover if a particle experiences no acceleration relative to $\mathcal{X}$ then it experiences no acceleration relative to $\mathcal{Y}$.

Proof: Since $\mathcal{X}$ and $\mathcal{Y}$ are inertial we have that there exist $\mathcal{X}_o$ and $\mathcal{Y}_o$ in $\mathcal{E}$ and fixed vectors $v_x, w_x, v_y, w_y \in \mathbb{R}^3$ and particular rotation matrices $A_x, A_y \in SO(3)$ such that

$$\mathcal{X}_t = A_x\mathcal{X}_o + tv_x + w_x \qquad\qquad \mathcal{Y}_t = A_y\mathcal{Y}_o + tv_y + w_y.$$

Further note that since $\mathcal{X}_o, \mathcal{Y}_o \in \mathcal{E}$ there exists fixed $Q \in SO(3)$ and $u \in \mathbb{R}^3$ such that $\mathcal{Y}_o = Q\mathcal{X}_o + u$. Thus, noting that $\mathcal{X}_o = A_x^{-1}(\mathcal{X}_t - tv_x - w_x)$ for the fourth line,

$$\begin{aligned} \mathcal{Y}_t &= A_y\mathcal{Y}_o + tv_y + w_y \\ &= A_y(Q\mathcal{X}_o + u) + tv_y + w_y \\ &= A_yQ\mathcal{X}_o + A_yu + tv_y + w_y \\ &= A_yQA_x^{-1}(\mathcal{X}_t - tv_x - w_x) + tv_y + A_yu + w_y \\ &= A_yQA_x^{-1}\mathcal{X}_t + t[v_y - A_yQA_x^{-1}v_x] - A_yQA_x^{-1}w_x + A_yu + w_y \end{aligned} \tag{5.10}$$

Thus define $A = A_yQA_x^{-1} \in SO(3)$, $v = v_y - A_yQA_x^{-1}v_x$, and $w = -A_yQA_x^{-1}w_x + A_yu + w_y$. Clearly $v, w \in \mathbb{R}^3$ and it is a short calculation to show that $A \in SO(3)$, we've left it as an exercise to the reader but it follows immediately if we already know that $SO(3)$ is a group under matrix multiplication (we have not proved this yet). Collecting our thoughts we have established the first half of the theorem, there exist $A \in SO(3)$ and $v, w \in \mathbb{R}^3$ such that,

$$\mathcal{Y}_t = A\mathcal{X}_t + tv + w$$

Now to complete the theorem consider a particle with trajectory $\gamma : I \to S$ such that $a_{\mathcal{X}} = 0$. Then by eqn.[5.9] we find, using our construction of A, v, w above,

$$\begin{aligned} a_{\mathcal{Y}} &= Aa_{\mathcal{X}} + \ddot{r} + \ddot{A}\gamma_{\mathcal{X}} + 2\dot{A}v_{\mathcal{X}} \\ &= A0 + 0 + 0\gamma_{\mathcal{X}} + 2(0)v_{\mathcal{X}} \\ &= 0. \end{aligned} \tag{5.11}$$

Therefore if the acceleration is zero relative to a particular inertial frame then it is zero for **all** inertial frames.

Consider that if a particle is either in a state of rest or uniform rectilinear motion then we can express it's trajectory γ relative to an observer $\mathcal{X} : I \to S$ by

$$\mathcal{X}_t(\gamma(t)) = tv + w$$

for all $t \in I$ and fixed $v, w \in \mathbb{R}^3$. In fact if $v = 0$ the particle is in a state of rest, whereas if $v \neq 0$ the particle is in a state of uniform rectilinear motion. Moreover,

$$\gamma_{\mathcal{X}}(t) = tv + w \iff v_{\mathcal{X}} = v \iff a_{\mathcal{X}} = 0.$$

Therefore we have shown that according to any inertial frame a particle that has zero acceleration necessarily travels in rectilinear motion or stays at rest.

Let us again ponder Newton's laws.

1. **Newton's First Law** Every particle persists in its state of rest or of uniform motion in a straight line unless it is compelled to change that state by impressed forces.

2. **Newton's Second Law** The rate of change of motion is proportional to the motive force impressed; and is made in the direction of the straight line in which that force is impressed.

3. **Newton's Third Law** To every action there is an equal reaction; or the mutual actions of two bodies upon each other are always equal but oppositely directed.

It is easy to see that if the first law holds relative to one observer then it does not hold relative to another observer which is rotating relative to the first observer. So a more precise formulation of the first law would be that it holds relative to *some* observer, or some class of observers, but not relative to all observers. *We have just shown that if $\mathcal{X}$ is an inertial observer then a particle is* *either in a state of rest or uniform rectilinear motion relative to $\mathcal{X}$ iff its acceleration is zero.* If γ is the trajectory of the particle the second law says that the force F acting on the body is proportional to $m(dv_{\mathcal{X}}/dt) = ma_{\mathcal{X}}$. Thus the second law says that a body has zero acceleration iff the force acting on the body is zero (assuming $m \neq 0$). It seems to follow that the first law is a consequence of the second law. What then *does* the first law say that is *not* contained in the second law ?

The answer is that the *first law is not a mathematical axiom but a physical principle.* It says it should be *possible to physically construct, at least in principle, a set* of coordinate systems at each instant of time which may be modeled by the mathematical construct we have been calling an inertial observer. Thus the *first law can be reformulated* to read:

There exists an inertial observer

The second law is also subject to criticism. When one speaks of the force on a body what is it that one is describing? Intuitively we think of a force as something which pushes or pulls the particle off its natural course.

The truth is that a course which seems natural to one observer may not appear natural to another. One usually models forces as vectors. These vectors provide the push or pull. The components of a vector in this context are observer dependent. The second law could almost be relegated to a definition. The force on a particle at time t would be defined to be $ma_{\mathcal{X}}(t)$ relative to the observer $\mathcal{X}$. Generally physicists require that the second law hold **only** for inertial observers. One reason for this is that if $F_{\mathcal{X}}$ is the force on a particle according to an inertial observer $\mathcal{X}$ and $F_{\mathcal{Y}}$ is the force on the same particle measured relative to the inertial observer $\mathcal{Y}$ then we claim $F_{\mathcal{Y}} = AF_{\mathcal{X}}$ where $\mathcal{X}$ and $\mathcal{Y}$ are related by

$$\mathcal{Y}_t = A\mathcal{X}_t + tv + w$$

for $v, w \in \mathbb{R}^3$ and $A \in SO(3)$ and for all t. Consider a particle traveling the trajectory γ we find it's accelerations as measured by $\mathcal{X}$ and $\mathcal{Y}$ are related by,

$$a_{\mathcal{Y}} = Aa_{\mathcal{X}}$$

where we have used eqn.[5.9] for the special case that A is a fixed rotation matrix and $r = tv + w$. Multiply by the mass to obtain that $ma_{\mathcal{Y}} = A(ma_{\mathcal{X}})$ thus $F_{\mathcal{Y}} = AF_{\mathcal{X}}$. Thus the form of Newton's law is maintained under admissible transformations of observer.

Remark 5.1.21.

The invariance of the form of Newton's laws in any inertial frame is known as the Galilean relativity principle. It states that no inertial frame is preferred in the sense that the physical laws are the same no matter which inertial frame you take observations from. This claim is limited to mechanical or electrostatic forces. The force between to moving charges due to a magnetic field does not act along the straight line connecting those charges. This exception was important to Einstein conceptually. Notice that if no frame is preferred then we can never, taking observations solely within an inertial frame, deduce the velocity of that frame. Rather we only can deduce relative velocities by comparing observations from different frames.

In contrast, if one defines the force relative to one observer $\mathcal{Z}$ which is rotating relative to $\mathcal{X}$ by $F_{\mathcal{Z}} = ma_{\mathcal{Z}}$ then one obtains a much more complex relation between $F_{\mathcal{X}}$ and $F_{\mathcal{Z}}$ which involves the force on the particle due to rotation. Such forces are called *fictitious forces* as they arise from the choice of noninertial coordinates, not a genuine force.

Example 5.1.22. ..

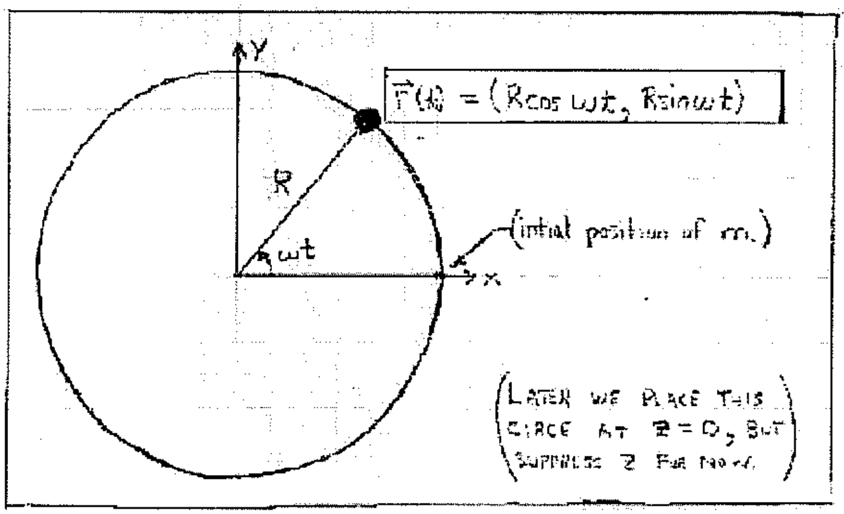
5.2 noninertial frames, a case study of circular motion

Some argue that any force proportional to mass may be viewed as a fictitious force, for example Hooke's law is F=kx, so you can see that the spring force is genuine. On the other hand gravity looks like $F = mg$ near the surface of the earth so some would argue that it is fictitious, however the conclusion of that thought takes us outside the realm of classical mechanics and the mathematics of this course. Anyway, if you are in a noninertial frame then for all intents and purposes fictitious forces are very real. The most familiar of these is probably the centrifugal force. Most introductory physics texts cast aspersion on the concept of centrifugal force (radially outward directed) because it is not a force observed from an inertial frame, rather it is a force due to noninertial motion. They say the centripetal (center seeking) force is really what maintains the motion and that there is no such thing as centrifugal force. I doubt most people are convinced by such arguments because it really feels like there is a force that wants to throw you out of a car when you take a hard turn. If there is no force then how can we feel it ? The desire of some to declare this force to be "fictional" stems from there belief that everything should be understood from the perspective of an inertial frame. Mathematically that is a convenient belief, but it certainly doesn't fit with everday experience. Ok, enough semantics. Lets examine circular motion in some depth.

For notational simplicity let us take $\mathbb{R}^3$ to be physical space and the identity mapping $\mathcal{X} = id$ to give us a stationary coordinate system on $\mathbb{R}^3$. Consider then the motion of a particle moving in a circle of radius R about the origin at a constant angular velocity of ω in the counterclockwise direction in the xy-plane. We will drop the third dimension for the most part throughout since it does not enter the calculations. If we assume that the particle begins at $(R, 0)$ at time zero then it follows that we can parametrize its path via the equations,

$$\begin{aligned} x(t) &= Rcos(\omega t) \\ y(t) &= Rsin(\omega t) \end{aligned} \tag{5.12}$$

this parametrization is geometric in nature and follows from the picture below, remember we took ω constant so that $\theta = \omega t$



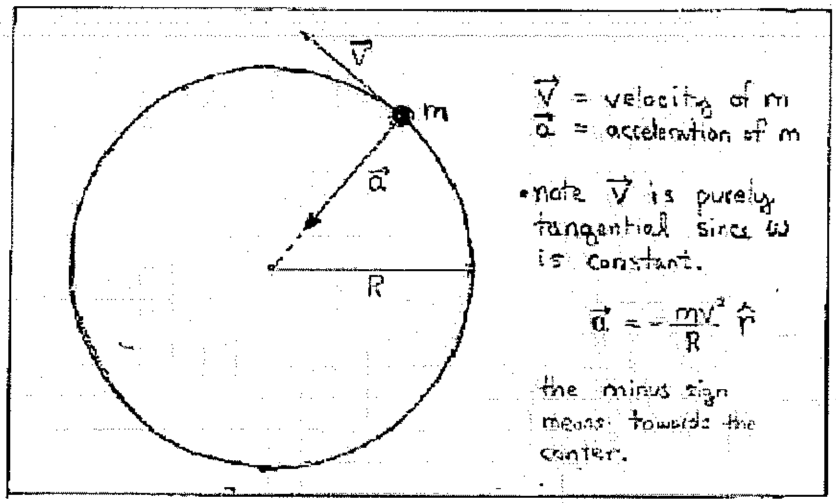
Now it is convenient to write $\vec{r}(t) = (x(t), y(t))$. Let us derive what the acceleration is for the particle, differentiate twice to obtain

$$\begin{aligned} \vec{r}''(t) &= (x''(t), y''(t)) \\ &= (-R\omega^2 cos(\omega t), -R\omega^2 sin(\omega t)) \\ &= -\omega^2 \vec{r}(t) \end{aligned}$$

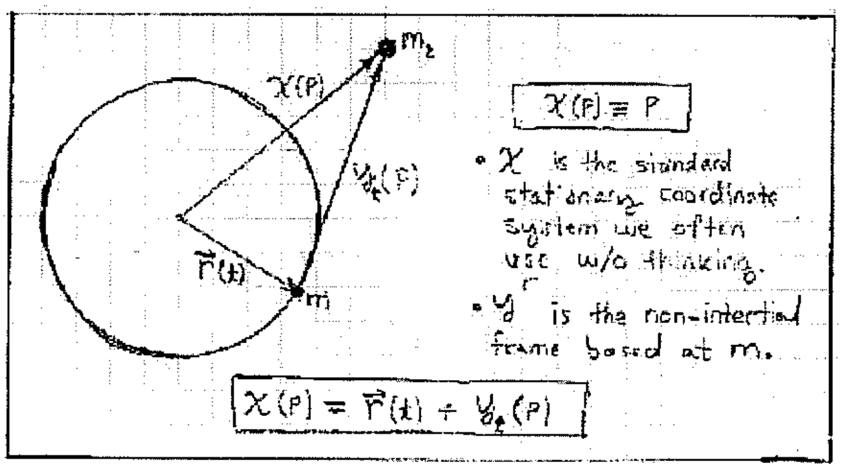
Now for pure circular motion the tangential velocity v is related to the angular velocity ω by $v = \omega R$. In other words $\omega = v/R$, radians per second is given by the length per second divided by the length of a radius. Substituting that into the last equation yields that,

$$\vec{a}(t) = \vec{r}''(t) = -\frac{v^2}{R^2} r(t) \tag{5.13}$$

The picture below summarizes our findings thus far.



Now define a second coordinate system that has its origin based at the rotating particle. We'll call this new frame $\mathcal{Y}$ whereas we have labeled the standard frame $\mathcal{X}$. Let $p \in \mathbb{R}^3$ be an arbitrary point then the following picture reveals how the descriptions of $\mathcal{X}$ and $\mathcal{Y}$ are related.



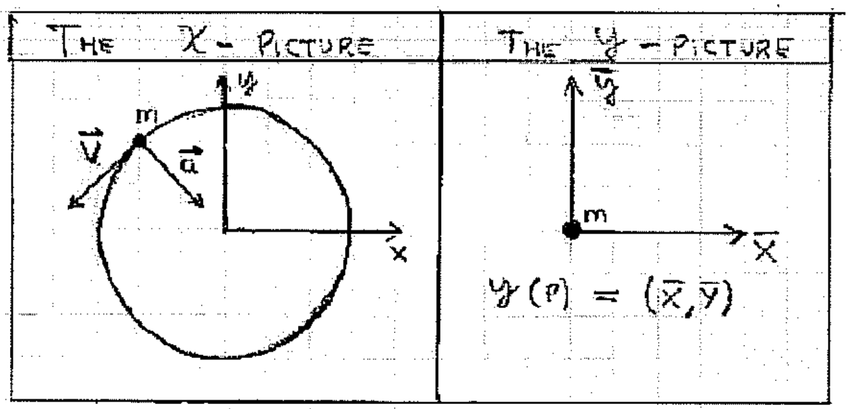
Clearly we find,

$$\mathcal{X}(p) = \mathcal{Y}(p) + \vec{r}(t) \tag{5.14}$$

note that the frames $\mathcal{X}$ and $\mathcal{Y}_t$ are not related by an rigid motion since $\vec{r}$ is not a constant function. Suppose that γ is the trajectory of a particle in $\mathbb{R}^3$, lets compare the acceleration of γ in frame $\mathcal{X}$ to that of it in $\mathcal{Y}_t$.

$$\begin{aligned} &\mathcal{X}(\gamma(t)) = \mathcal{Y}_t(\gamma(t)) + \vec{r}(t) \\ &\implies a_{\mathcal{X}}(t) = \gamma''(t) = a_{\mathcal{Y}_t}(t) + \vec{r}''(t) \end{aligned} \tag{5.15}$$

If we consider the special case of $\gamma(t) = r(t)$ we find the curious but trivial result that $\mathcal{Y}_t(r(t)) = 0$ and consequently $a_{\mathcal{Y}_t}(t) = 0$. Perhaps a picture is helpful,

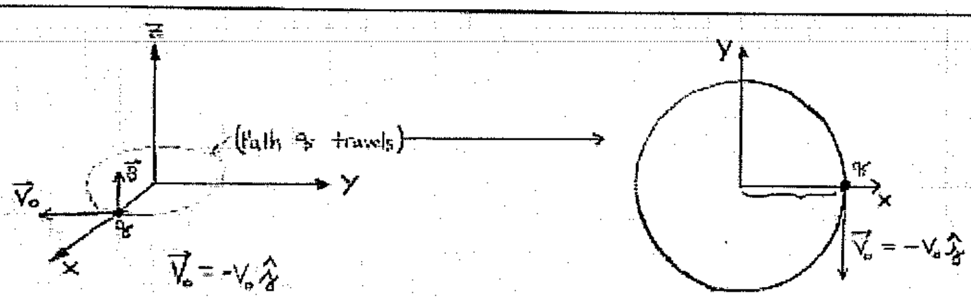


We have radically different pictures of the motion of the rotating particle, in the $\mathcal{X}$ picture the particle is accelerated and using our earlier calculation,

$$\boxed{a_{\mathcal{X}} = \vec{r}''(t) = \frac{-v^2}{R}\hat{r}}$$

on the other hand in the $\mathcal{Y}_t$ frame the mass just sits at the origin with $a_{\mathcal{Y}calt} = 0$. Since $F = ma$ we would conclude (ignoring our restriction to inertial frames for a moment) that the particle has an external force on it in the $\mathcal{X}$ frame but not in the $\mathcal{Y}$ frame. This clearly throws a wrench in the universality of the force concept, it is for this reason that we must restrict to inertial frames if we are to make nice broad sweeping statements as we have been able to in earlier sections. If we allowed noninertial frames in the basic set-up then it would be difficult to ever figure out what if any forces were in fact genuine. Dwelling on these matters actually led Einstein to his theory of general relativity where noninertial frames play a central role in the theory.

Anyway, lets think more about the circle. The relation we found in the $\mathcal{X}$ frame does not tell us *how* the particle is remaining in circular motion, rather only that if it is then it must have an acceleration which points towards the center of the circle with precisely the magnitude mv^2/R. I believe we have all worked problems based on this basic relation. An obvious question remains, which force makes the particle go in a circle? Well, we have not said enough about the particle yet to give a definitive answer to that question. In fact many forces could accomplish the task. You might imagine the particle is tethered by a string to the central point, or perhaps it is stuck in a circular contraption and the contact forces with the walls of the contraption are providing the force. A more interesting possibility for us is that the particle carries a charge and it is subject to a magnetic field in the z-direction. Further let us assume that the initial position of the charge q is $(mv/qB, 0, 0)$ and the initial velocity of the charged particle is v in the negative y-direction. I'll work this one out one paper because I can.



We assume that we have a charge q with mass m that has intial position $(mV_o/qB, 0, 0)$ with intial velocity $\vec{V}_o = (0, -V_o, 0)$ then the charge is subject to a constant background magnetic field of $\vec{B} = (0, 0, B)$ where $B > 0$. We show that the charge travels a circle of radius $R = mV_o/qB$. We begin with Newton's 2nd law paired with the Coulomb force law,

$$m\frac{d^2\vec{r}}{dt^2} = q\vec{V} \times \vec{B} = qB\left(\frac{dx}{dt}, \frac{dy}{dt}, \frac{dz}{dt}\right) \times (0, 0, 1)$$

$$\frac{d^2\vec{r}}{dt^2} = \frac{qB}{m}\left(\frac{dy}{dt}\hat{i} - \frac{dx}{dt}\hat{j}\right) = \frac{d^2x}{dt^2}\hat{i} + \frac{d^2y}{dt^2}\hat{j} + \frac{d^2z}{dt^2}\hat{k}$$

This gives us three 2nd order ODEs to solve, let $\alpha \equiv qB/m$.

$$\frac{d^2z}{dt^2} = 0 \quad \text{has sol}^n: \quad z(t) = c_1 + c_2 t \quad \Rightarrow \quad z(t) = 0$$
(using intial conditions)

$$\left\{\begin{aligned} \frac{d^2y}{dt^2} &= -\alpha\frac{dx}{dt} \\ \frac{d^2x}{dt^2} &= \alpha\frac{dy}{dt} \end{aligned}\right\} \quad \text{Coupled.}$$

Notice that $\frac{dy}{dt} = \frac{1}{\alpha}\frac{d^2x}{dt^2} \Rightarrow \frac{d^2y}{dt^2} = \frac{1}{\alpha}\frac{d^3x}{dt^3} = -\alpha\frac{dx}{dt}$.

Thus we need to solve $x''' = -\alpha^2 x'$, introduce $W \equiv x'$ then our DEq^n becomes $W'' = -\alpha^2 W$ this has well known sol^n,

$$W = \boxed{\frac{dx}{dt} = c_1\cos(\alpha t) + c_2\sin(\alpha t)}$$

Then we can calculate everything else from this

$$\frac{dy}{dt} = \frac{1}{\alpha}\frac{d^2x}{dt^2} = \frac{1}{\alpha}\left(-c_1\alpha\sin(\alpha t) + c_2\alpha\cos(\alpha t)\right)$$

$$\Rightarrow \boxed{\frac{dy}{dt} = -c_1\sin(\alpha t) + c_2\cos(\alpha t)}$$

Continuing,

Continuing, we found for $\alpha = qB/m$ that

$$\boxed{\begin{aligned} \frac{dx}{dt} &= C_1 \cos(\alpha t) + C_2 \sin(\alpha t) \\ \frac{dy}{dt} &= -C_1 \sin(\alpha t) + C_2 \cos(\alpha t) \end{aligned}}$$

Now integrate both to find $x(t)$ & $y(t)$

$$\boxed{\begin{aligned} x(t) &= \frac{1}{\alpha}\left(C_1 \sin(\alpha t) - C_2 \cos(\alpha t)\right) + C_3 \\ y(t) &= \frac{1}{\alpha}\left(C_1 \cos(\alpha t) + C_2 \sin(\alpha t)\right) + C_4 \end{aligned}}$$

Finally apply intial conditions,

$$\begin{aligned} x'(0) &= 0 = C_1 \Rightarrow \boxed{C_1 = 0} \\ y'(0) &= -V_0 = C_2 \Rightarrow \boxed{C_2 = -V_0} \\ x(0) &= mV_0/qB = V_0/\alpha = \frac{1}{\alpha}(V_0) + C_3 \Rightarrow \boxed{C_3 = 0} \\ y(0) &= 0 = \frac{1}{\alpha}(0) + C_4 \Rightarrow \boxed{C_4 = 0} \end{aligned}$$

Thus collecting our results we find, since $C_2/\alpha = -V_0/\alpha = \frac{-mV_0}{qB}$

$$\boxed{\begin{aligned} x(t) &= (mV_0/qB) \cos(qBt/m) \\ y(t) &= -(mV_0/qB) \sin(qBt/m) \\ z(t) &= 0 \end{aligned}}$$

This is the circle $x^2 + y^2 = R^2$ with radius $R = mV_0/qB$, lying in the $z = 0$ plane.

It is curious that magnetic forces cannot be included in the Galilean relativity. For if the velocity of a charge is zero in one frame but not zero in another then does that mean that the particle has a non-zero force or no force? In the rest frame of the constant velocity charge apparently there is no magnetic force, yet in another inertially related frame where the charge is in motion there would be a magnetic force. How can this be? The problem with our thinking is we have not asked how the magnetic field transforms for one thing, but more fundamentally we will find that you cannot separate the magnetic force from the electric force. Later we'll come to a better understanding of this, there is no nice way of addressing it in Newtonian mechanics that I know of. It is an inherently relativistic problem, and Einstein attributes it as one of his motivating factors in dreaming up his special relativity.

> "What led me more or less directly to the special theory of relativity was the conviction that the electromotive force acting on a body in motion in a magnetic field was nothing else but an electric field"

Albert Einstein, 1952.

Rotating Coordinate Systems

The central idea is that different coordinate systems give descriptions of the same point,

$$\vec{r} = x e_1 + y e_2 + z e_3 = \bar{x}\bar{e}_1 + \bar{y}\bar{e}_2 + \bar{z}\bar{e}_3$$

The idea here is that $\bar{e}_1, \bar{e}_2, \bar{e}_3$ is a rotating frame and these coordinate systems share a common origin. Furthermore, we have in mind the coordinates x, y, z of some moving object so generally there are also functions of time. Time is time so differentiate,

$$\frac{d\vec{r}}{dt} = \frac{d}{dt}\left[\bar{x}\bar{e}_1 + \bar{y}\bar{e}_2 + \bar{z}\bar{e}_3\right]$$

$$\frac{d\vec{r}}{dt} = \underbrace{\frac{d\bar{x}}{dt}\bar{e}_1 + \frac{d\bar{y}}{dt}\bar{e}_2 + \frac{d\bar{z}}{dt}\bar{e}_3}_{\text{I}} + \underbrace{\bar{x}\frac{d\bar{e}_1}{dt} + \bar{y}\frac{d\bar{e}_2}{dt} + \bar{z}\frac{d\bar{e}_3}{dt}}_{\text{II}}$$

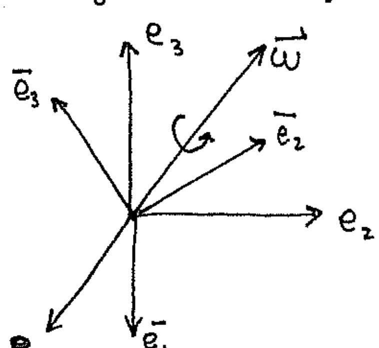
Note that in contrast the fixed, time independent frame e_1, e_2, e_3 has the simple familar form,

$$\frac{d\vec{r}}{dt} = \frac{d}{dt}\left[x e_1 + y e_2 + z e_3\right]$$

$$= \underbrace{\frac{dx}{dt}e_1 + \frac{dy}{dt}e_2 + \frac{dz}{dt}e_3}_{\vec{V}_S}$$

$\vec{V}_S$ the velocity relative to the fixed coordinate system S

In contrast, $\vec{V}_{\bar{S}} = \frac{d\bar{x}}{dt}\bar{e}_1 + \frac{d\bar{y}}{dt}\bar{e}_2 + \frac{d\bar{z}}{dt}\bar{e}_3$ is velocity relative to the rotating coordinate system $\bar{S}$. The terms in II give the velocity of the $\bar{S}$-frame itself. It can be shown that $\frac{d\bar{e}_j}{dt} = \vec{\omega} \times \bar{e}_j$ for $j = 1, 2, 3$ where $\vec{\omega}$ is the <u>angular velocity</u> of the rotating frame



(imagine $\bar{e}_1, \bar{e}_2, \bar{e}_3$ rotate about the axis $\vec{\omega}$)

Here II $\longrightarrow$ $\vec{\omega} \times \underbrace{(\bar{x}\bar{e}_1 + \bar{y}\bar{e}_2 + \bar{z}\bar{e}_3)}_{\vec{r}}$

$\left(\text{using } \frac{d\bar{e}_j}{dt} = \vec{\omega} \times \bar{e}_j \text{ for } j = 1, 2, 3\right)$

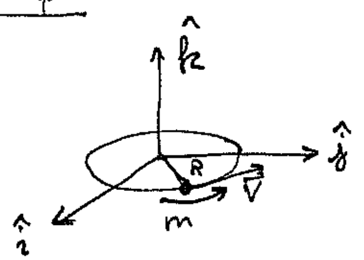
Setting aside the question of <u>why</u> $\frac{d\bar{e}_j}{dt} = \vec{\omega} \times \bar{e}_j$, we find

$$\begin{aligned}
\frac{d\vec{r}}{dt} &= \frac{d\bar{x}}{dt}\bar{e}_1 + \frac{d\bar{y}}{dt}\bar{e}_2 + \frac{d\bar{z}}{dt}\bar{e}_3 + \bar{x}\frac{d\bar{e}_1}{dt} + \bar{y}\frac{d\bar{e}_2}{dt} + \bar{z}\frac{d\bar{e}_3}{dt} \\
&= \frac{d\bar{x}_1}{dt}\bar{e}_1 + \frac{d\bar{y}}{dt}\bar{e}_2 + \frac{d\bar{z}}{dt}\bar{e}_3 + \bar{x}(\vec{\omega}\times\bar{e}_1) + \bar{y}(\vec{\omega}\times\bar{e}_2) + \bar{z}(\vec{\omega}\times\bar{e}_3) \\
&= \underbrace{\frac{d\bar{x}}{dt}\bar{e}_1 + \frac{d\bar{y}}{dt}\bar{e}_2 + \frac{d\bar{z}}{dt}\bar{e}_3}_{\vec{V}_{\bar{S}}} + \vec{\omega}\times(\underbrace{\bar{x}\bar{e}_1 + \bar{y}\bar{e}_2 + \bar{z}\bar{e}_3}_{\vec{r}})
\end{aligned}$$

We find $\boxed{\vec{V}_S = \vec{V}_{\bar{S}} + \vec{\omega}\times\vec{r}}$

This assumed a <u>common</u> origin.

<u>Example:</u>



Let $\bar{S}$ rotate at constant velocity $\vec{\omega}$ about the $\hat{k}$ vector. Consider point m fixed at point in $\bar{S}$ frame; $\vec{r}(t) = \langle R\cos t, R\sin t, 0\rangle$

oops. I just set $\omega = 1$ see *

$$\begin{aligned}
\vec{V}_s &= \omega\hat{k}\times(R\cos t\,\hat{i} + R\sin t\,\hat{j}) \\
&= \omega R\cos t(\hat{k}\times\hat{i}) + \omega R\sin t(\hat{k}\times\hat{j}) \\
&= (\omega R\cos t)\hat{j} - (\omega R\sin t)\hat{i} \\
&= \underbrace{\omega R(\langle -\sin t, \cos t\rangle)}_{\text{the tangential velocity}}
\end{aligned}$$

$\frac{d\vec{r}}{dt} = \langle -R\cos t, R\sin t\rangle$

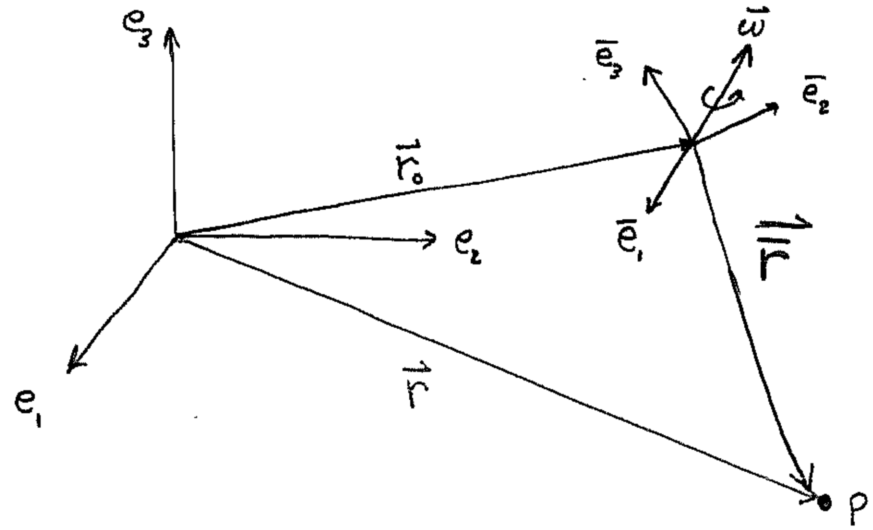
I have shown that $\vec{V} = \vec{\omega}\times\vec{r}$ is true in this special case.

---

* $\vec{r}(t) = \langle R\cos\omega t, R\sin\omega t, 0\rangle$

$$\frac{d\vec{r}}{dt} = R\omega\langle -\sin\omega t, \cos\omega t, 0\rangle = \omega\hat{k}\times\vec{r}$$

(I forgot the ω above)



rotating coordinate system $\bar{S}$ with non-matching origin relative to fixed inertial frame $S: \{e_1, e_2, e_3\}$.

$$\vec{r} = x e_1 + y e_2 + z e_3$$

$$\vec{\bar{r}} = \bar{x}\bar{e}_1 + \bar{y}\bar{e}_2 + \bar{z}\bar{e}_3$$

From picture we see that the variable point P has

$$\vec{r} = \vec{r}_0 + \vec{\bar{r}}$$

Almost same calculation goes through,

$$\frac{d\vec{r}}{dt} = \frac{d\vec{r}_0}{dt} + \frac{d}{dt}\left(\bar{x}\bar{e}_1 + \bar{y}\bar{e}_2 + \bar{z}\bar{e}_3\right)$$

$$= \frac{d\vec{r}_0}{dt} + \underbrace{\frac{d\bar{x}}{dt}\bar{e}_1 + \frac{d\bar{y}}{dt}\bar{e}_2 + \frac{d\bar{z}}{dt}\bar{e}_3}_{} + \bar{x}\frac{d\bar{e}_1}{dt} + \bar{y}\frac{d\bar{e}_2}{dt} + \bar{z}\frac{d\bar{e}_3}{dt}$$

$$\boxed{\vec{V}_S = \frac{d\vec{r}_0}{dt} + \vec{V}_{\bar{S}} + \vec{\omega} \times \vec{\bar{r}}}$$

This formula relates the velocity measured relative to a fixed vs. rotating frame. The first two terms should be familar from freshman mechanics.

$$\frac{d\vec{r}_0}{dt} = \text{velocity of moving frame's origin}$$

$$\vec{V}_{\bar{S}} = \text{velocity relative to moving frame}$$

However, the $\vec{\omega} \times \vec{\bar{r}}$ is probably new to you in this context.

Comparing acceleration in fixed frame S verses rotating $\bar{S}$

Continuing with the notation from the previous page,

$$\frac{d\vec{r}}{dt} = \vec{V}_S = \frac{d\vec{r}_o}{dt} + \frac{d\bar{x}}{dt}\bar{e}_1 + \frac{d\bar{y}}{dt}\bar{e}_2 + \frac{d\bar{z}}{dt}\bar{e}_3 + \vec{\omega}\times\vec{\bar{r}}$$

Now differentiate again, note this calculation is very much like the one we just completed,

$$\frac{d^2\vec{r}}{dt^2} = \frac{d^2\vec{r}_o}{dt^2} + \frac{d^2\bar{x}}{dt^2}\bar{e}_1 + \frac{d^2\bar{y}}{dt^2}\bar{e}_2 + \frac{d^2\bar{z}}{dt^2}\bar{e}_3 + \frac{d\bar{x}}{dt}\frac{d\bar{e}_1}{dt} + \frac{d\bar{y}}{dt}\frac{d\bar{e}_2}{dt} + \frac{d\bar{z}}{dt}\frac{d\bar{e}_3}{dt} + \frac{d}{dt}\left(\vec{\omega}\times\vec{\bar{r}}\right)$$

$$= \vec{a}_o + \vec{a}_{\bar{S}} + \frac{d\bar{x}}{dt}(\vec{\omega}\times\bar{e}_1) + \frac{d\bar{y}}{dt}(\vec{\omega}\times\bar{e}_2) + \frac{d\bar{z}}{dt}(\vec{\omega}\times\bar{e}_3) + \frac{d\vec{\omega}}{dt}\times\vec{\bar{r}} + \vec{\omega}\times\frac{d\vec{\bar{r}}}{dt}$$

$$= \vec{a}_o + \vec{a}_{\bar{S}} + \vec{\omega}\times\left(\frac{d\bar{x}}{dt}\bar{e}_1 + \frac{d\bar{y}}{dt}\bar{e}_2 + \frac{d\bar{z}}{dt}\bar{e}_3\right) + \frac{d\vec{\omega}}{dt}\times\vec{\bar{r}} + \vec{\omega}\times\frac{d\vec{\bar{r}}}{dt}$$

$$\boxed{\vec{a}_S = \vec{a}_o + \vec{a}_{\bar{S}} + 2\vec{\omega}\times\vec{V}_{\bar{S}} + \frac{d\vec{\omega}}{dt}\times\vec{\bar{r}} + \vec{\omega}\times(\vec{\omega}\times\vec{\bar{r}})}$$

The notation $\vec{a}_o = \frac{d^2\vec{r}_o}{dt^2}$ gives acceleration of origin of $\bar{S}$ relative to S.

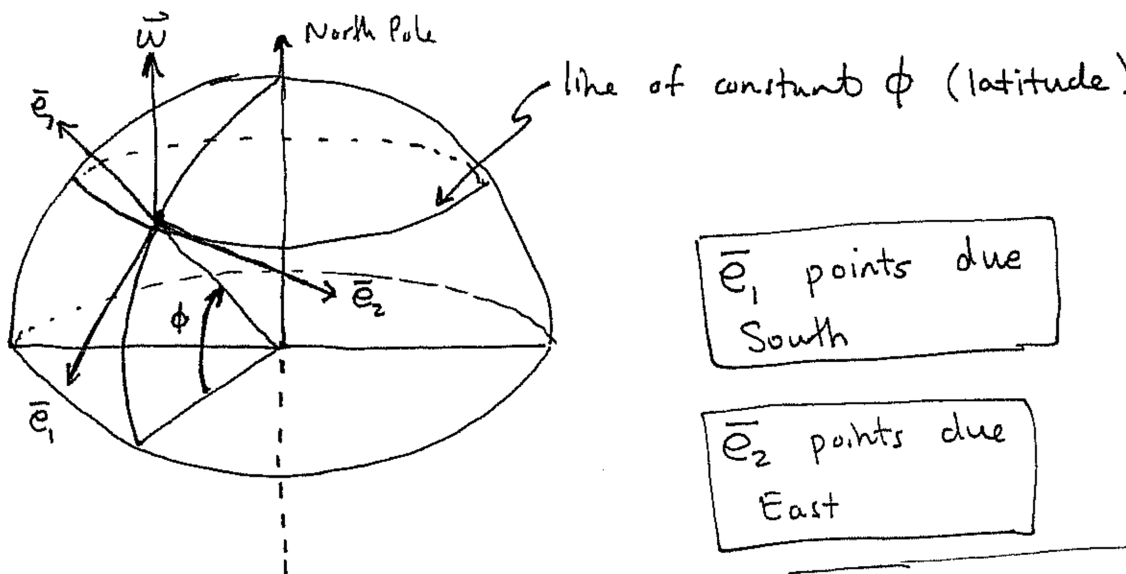
<u>Thought Experiment</u>: Suppose you took measurements relative to $\bar{S}$ and <u>assumed</u> it was an intertial frame. What "fake" forces would you encounter? Assuming Newton's 2nd Law,

$$m\vec{a}_{\bar{S}} = m\vec{a}_S - m\vec{a}_o - 2m\vec{\omega}\times\vec{V}_{\bar{S}} - m\frac{d\vec{\omega}}{dt}\times\vec{\bar{r}} - m\vec{\omega}\times(\vec{\omega}\times\vec{\bar{r}})$$

$$\boxed{m\vec{a}_{\bar{S}} = \underbrace{\vec{F}_{net}}_{\text{real forces like springs or gravity}} - \underbrace{m\vec{a}_o}_{\text{rectilinear acceleration of frame } \bar{S}} - \underbrace{2m\vec{\omega}\times\vec{V}_{\bar{S}}}_{\text{Coriolis Force}} - \underbrace{m\frac{d\vec{\omega}}{dt}\times\vec{\bar{r}}}_{\text{no-name but you've felt this on a tilt-a-whirl as it stops or starts}} - \underbrace{m\vec{\omega}\times(\vec{\omega}\times\vec{\bar{r}})}_{\text{centrifugal force.}}}$$

Rotating Frame of Reference we live in

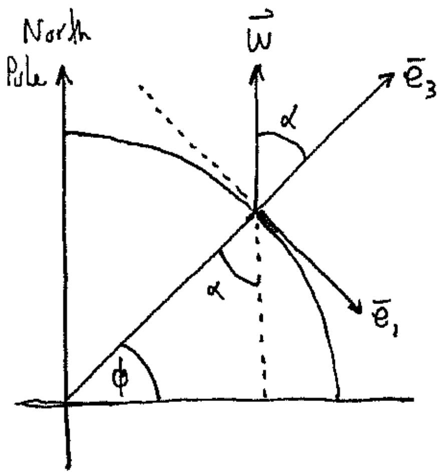
We've done almost all the math. Think about it, the earth is essentially moving at constant velocity relative to solar system over a time of <u>minutes</u>. So we can reasonable put a <u>fixed</u> coordinate system S at the center of the earth. Moreover, modulo earthquakes & tidal waves, the rotation of the earth is nearly constant in magnitude. This means the $\frac{d\vec{\omega}}{dt}$ term vanishes. Let's set up a rotating frame of reference (borrowed from McComb's "Dynamics and Relativity" Oxford Press.)



$\bar{e}_1$ points due South

$\bar{e}_2$ points due East

$\bar{e}_3$ points straight up



Continuing, we find the following Deq$^{\underline{n}}$ of motion near surface of earth at Latitude ϕ, either ignore or lump $\vec{\omega}\times(\vec{\omega}\times\vec{r})$ into the $-mg\bar{e}_3$ term,

$$m\left(\frac{d^2\bar{x}}{dt^2}\bar{e}_1+\frac{d^2\bar{y}}{dt^2}\bar{e}_2+\frac{d^2\bar{z}}{dt^2}\bar{e}_3\right)=\underbrace{-mg\bar{e}_3}_{\text{gravity}}-\underbrace{2m\vec{\omega}\times\vec{V}_{\bar{S}}}_{\text{Corriolis Force}}$$

Work out the Coriolis term,

$$\vec{\omega}=\omega e_3=\omega(e_3\cdot\bar{e}_1)\bar{e}_1+\omega\underbrace{(e_3\cdot\bar{e}_2)}_{0}\bar{e}_2+\omega(e_3\cdot\bar{e}_3)\bar{e}_3$$

$$\Longrightarrow\ \underbrace{\vec{\omega}=\omega\cos(\alpha+\tfrac{\pi}{2})\,\bar{e}_1+\omega\cos\alpha\,\bar{e}_3}_{}$$

See picture on previous page it's clear the angles are $\alpha+\frac{\pi}{2}$ and α between e_3 & $\bar{e}_1$ & e_3 & $\bar{e}_3$ respective.

However, $\alpha=\frac{\pi}{2}-\phi$ thus

$$\cos(\alpha+\tfrac{\pi}{2})=\cos(\pi-\phi)=-\cos\phi$$
$$\cos\alpha=\cos(\tfrac{\pi}{2}-\phi)=+\sin\phi$$

We find that $\underline{\vec{\omega}=-\omega\cos\phi\,\bar{e}_1+\omega\sin\phi\,\bar{e}_3}$.

I did this so we can use $\bar{e}_1\times\bar{e}_2=\bar{e}_3$ etc... in the following,

$$\begin{aligned}\vec{\omega}\times\vec{V}_{\bar{S}}&=(-\omega\cos\phi\,\bar{e}_1+\omega\sin\phi\,\bar{e}_3)\times\left(\frac{d\bar{x}}{dt}\bar{e}_1+\frac{d\bar{y}}{dt}\bar{e}_2+\frac{d\bar{z}}{dt}\bar{e}_3\right)\\&=-\omega\cos\phi\frac{d\bar{y}}{dt}\bar{e}_3+\omega\cos\phi\frac{d\bar{z}}{dt}\bar{e}_2+\omega\sin\phi\frac{d\bar{x}}{dt}\bar{e}_2-\omega\sin\phi\frac{d\bar{y}}{dt}\bar{e}_1\\&=\left(-\omega\sin\phi\frac{d\bar{y}}{dt}\right)\bar{e}_1+\left(\omega\cos\phi\frac{d\bar{z}}{dt}+\omega\sin\phi\frac{d\bar{x}}{dt}\right)\bar{e}_2-\omega\cos\phi\frac{d\bar{y}}{dt}\bar{e}_3\end{aligned}$$

Putting this together with Newton's Law,

$$\boxed{\begin{aligned}\bar{e}_1:&\quad m\frac{d^2\bar{x}}{dt^2}=2m\omega\sin\phi\frac{d\bar{y}}{dt}\\\bar{e}_2:&\quad m\frac{d^2\bar{y}}{dt^2}=-2m\omega\cos\phi\frac{d\bar{z}}{dt}-2m\omega\sin\phi\frac{d\bar{x}}{dt}\\\bar{e}_3:&\quad m\frac{d^2\bar{z}}{dt^2}=2m\omega\cos\phi\frac{d\bar{y}}{dt}-mg\end{aligned}}$$

Remark: since $m,\omega,\cos\phi$ are all constants for a given problem we can solve this by reduction of order to a 6×6 nonhomog. matrix problem!

Approximate Solⁿ for Coriolis Problem

Compared to mg the terms with $2m\omega$ are proportionally smaller. If we consider throwing an object vertically it stands to reason only $\frac{d\bar{z}}{dt}$ is nontrivial, the Coriolis force will create some nonzero $\frac{d\bar{x}}{dt}, \frac{d\bar{y}}{dt}$ as time progresses, but those terms are small. Hence we can solve:

$$m\frac{d^2\bar{x}}{dt^2} = 0$$

$$m\frac{d^2\bar{y}}{dt^2} = -2m\omega\cos\phi\,\frac{d\bar{z}}{dt}$$

$$m\frac{d^2\bar{z}}{dt^2} = -mg$$

We may simply integrate the $\bar{x}$ & $\bar{z}$ eqⁿs to find

$$\boxed{\begin{aligned}\bar{x}(t) &= x_0\\ \bar{z}(t) &= z_0 - \;t - \tfrac{1}{2}gt^2\end{aligned}}$$

Same as usual. The interesting feature is the Eastward drift captured in the $\bar{y}$ - eqⁿ.

Hence $\frac{d\bar{z}}{dt} = \bar{V}_0 - gt$. This gives,

$$m\frac{d^2\bar{y}}{dt^2} = -2m\omega\cos\phi\left[\bar{V}_0 - gt\right]$$

$$\Rightarrow \frac{d^2\bar{y}}{dt^2} = -2\omega\cos\phi\,\bar{V}_0 + (2\omega\cos\phi\, g)\,t$$

$$\Rightarrow \frac{d\bar{y}}{dt} = -2\omega\cos\phi\,\bar{V}_0\,t + \omega\cos\phi\, g t^2 \qquad \left(\text{assumed } \frac{d\bar{y}}{dt}(0) = 0.\right)$$

$$\Rightarrow \boxed{\bar{y}(t) = \left(\tfrac{1}{3}g\omega\cos\phi\right)t^3 - \left(\omega\cos\phi\,\bar{V}_0\right)t}$$

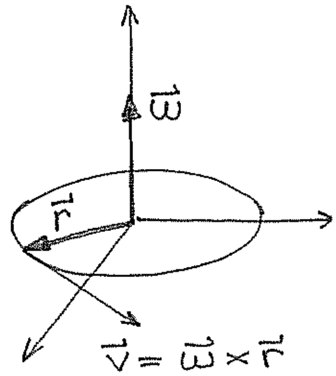
Or, if we drop a mass so $\bar{V}_0 = 0$ we have

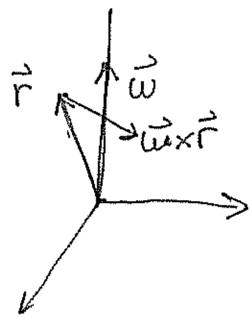
$$\underbrace{\boxed{\bar{y}(t) = \tfrac{1}{3}(g\omega\cos\phi)\,t^3}}$$

Coriolis drift goes east in Northern Hemisphere.

(Notice the $-2m\vec{\omega}\times\vec{V}_{\bar{S}}$ points opposite direction below equator.)

Concerning <u>why</u> $\frac{d\vec{r}}{dt} = \vec{\omega} \times \vec{r}$ (I used $\frac{de_j}{dt} = \vec{\omega} \times e_j$ before) C-8





$A_{\vec{\omega}}$ rotates at <u>constant</u> angular velocity ω about $\frac{\vec{\omega}}{\omega} = \hat{n}$

$$A_{\vec{\omega}}(t) \equiv \begin{bmatrix} \cos\omega t & -\sin\omega t & 0 \\ \sin\omega t & \cos\omega t & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\vec{r}(t) = A_{\vec{\omega}}(t)\,\vec{r}_0 \rightsquigarrow \vec{r}(t) = [\cos\omega t\, x_0 - \sin\omega t\, y_0, \sin\omega t\, x_0 + \cos\omega t\, y_0, z$$

$$\frac{d\vec{r}}{dt} = \frac{dA}{dt}\vec{r}_0 = \begin{bmatrix} -\omega\sin\omega t & -\omega\cos\omega t & 0 \\ \omega\cos\omega t & -\omega\sin\omega t & 0 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix}$$

$$= \begin{bmatrix} -\omega x_0 \sin\omega t - \omega y_0 \cos\omega t \\ \omega x_0 \cos\omega t - \omega y_0 \sin\omega t \\ 0 \end{bmatrix}$$

$$= \underline{\omega[-x_0\sin\omega t - y_0\cos\omega t, \; x_0\cos\omega t - y_0\sin\omega t, \; 0]}$$

$$\vec{\omega} = \omega[0,0,1]^T \rightsquigarrow \vec{\omega}\times\vec{r} = \omega\hat{k}\times(x\hat{i} + y\hat{j} + z_0\hat{k})$$

$$= \omega x\hat{j} - \omega y\hat{i}$$

$$= \underline{[-\omega x_0\sin\omega t - \omega y_0\cos\omega t, \; \omega x_0\cos\omega t - \omega y_0\sin\omega t, 0]}$$

$$\therefore \; \vec{v} = \vec{\omega}\times\vec{r}$$

<u>Remark</u>: this proof is almost general but it needs a little work....

(Just for fun, this is unfinished)

$$A^T A = I$$

$$A = e^{Bt}, \quad \frac{dA}{dt} = Be^{Bt}$$

$$\frac{dA^T}{dt} A + A^T \frac{dA}{dt} = 0$$

Assume $\gamma(t) = e^{Bt}$ then $\gamma(0) = I$ & $\gamma'(0) = B$.
If $\gamma^T\gamma = I$ then $\frac{d\gamma^T}{dt}(t)\,\gamma(t) + \gamma^T(t)\frac{d\gamma}{dt}(t) = 0$

$$\therefore \quad B^T + B = 0 \longrightarrow B^T = -B.$$

$$\Rightarrow \quad B = \begin{bmatrix} 0 & b_3 & b_2 \\ -b_3 & 0 & b_1 \\ -b_2 & -b_1 & 0 \end{bmatrix} = b_3 \underbrace{\begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}}_{\epsilon_{ij3}} + b_2 \underbrace{\begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}}_{\epsilon_{ij2}} + b_1 \underbrace{\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix}}_{\epsilon_{ij1}}$$

Let $(J_k)_{ij} = \epsilon_{ijk}$.

<u>Claim</u>: $R_{\vec{\omega}} = \exp(\vec{J}\cdot\vec{\omega})$

$$R_{\omega\hat{k}} = \exp(\omega J_3)$$

Umm....

$$\begin{aligned} \frac{dR_{\vec{\omega}}}{dt} &= \frac{d}{dt}\exp\left(t(\vec{J}\cdot\vec{\omega})\right) \\ &= \exp(t\vec{J}\cdot\vec{\omega})\,\frac{d}{dt}\left(t(\vec{J}\cdot\vec{\omega})\right) \\ &= \underline{\vec{J}\cdot\vec{\omega}}\,\exp(t\vec{J}\cdot\vec{\omega}) \\ &= \epsilon_{ijk}\omega_k\, R_{\vec{\omega}} \end{aligned}$$

$$\vec{r}(t) = R_{\vec{\omega}}(t)\,\vec{r}_0$$

$$\begin{aligned} \frac{d\vec{r}}{dt} &= \frac{dR_{\vec{\omega}}}{dt}(t)\,\vec{r}_0 \\ &= \left(\omega_1 J_1 + \omega_2 J_2 + \omega_3 J_3\right) R_{\vec{\omega}}(t)\,\vec{r}_0 \\ &= \left(\omega_1 J_1 + \omega_2 J_2 + \omega_3 J_3\right)\vec{r}(t). \\ &= (\omega_1, \omega_2, \omega_3) \times \vec{r}(t) \end{aligned}$$

$$J_1\vec{r} = \epsilon_{ij1} r_j = \epsilon_{231} r_3 + \epsilon_{321} r_2 = r_3 - r_2 = z - y.$$

Chapter 6

differentiation

In this chapter we define differentiation of mappings. I follow Edwards fairly closely, his approach is efficient and his langauge clarifies concepts which are easily confused. Susan Colley's text *Vector Calculus* is another good introductory text which describes much of the mathematics in this chapter. When I teach calculus III I do touch on the main thrust of this chapter but I shy away from proofs and real use of linear algebra. That is not the case here.

6.1 derivatives and differentials

In this section we motivate the general definition of the derivative for mappings from $\mathbb{R}^n$ to $\mathbb{R}^m$. Naturally this definition must somehow encompass the differentiation concepts we've already discussed in the calculus sequence: let's recall a few examples to set the stage,

1. derivatives of functions of $\mathbb{R}$, for example $f(x) = x^2$ has $f'(x) = 2x$
2. derivatives of mappings of $\mathbb{R}$, for example $f(t) = (t, t^2, t^3)$ has $f'(t) =< 1, 2t, 3t^2 >$.
3. $f : dom(f) \subseteq \mathbb{R}^2 \to \mathbb{R}$ has directional derivative $(D_u f)(p) = (\nabla f)(p) \cdot u$ where $\nabla f = grad(f) =< \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} >$
4. $X : U \subset \mathbb{R}^2_{uv} \to \mathbb{R}^3_{xyz}$ parametrizes a surface $X(U)$ and $N(u, v) = \frac{\partial X}{\partial u} \times \frac{\partial X}{\partial v}$ gives the normal vector field to the surface.

We'd like to understand how these derivatives may be connected in some larger context. If we could find such a setting then that gives us a way to state theorems about derivatives in an efficient and general manner. We also should hope to gain a deeper insight into the geometry of differentiation.

6.1.1 derivatives of functions of a real variable

Let's revisit the start of Calculus I. We begin by defining the change in a function f between the point a and $a+h$:

$$\Delta f = f(a+h) - f(a).$$

We can approximate this change for small values of h by replacing the function with a line. Recall that the line closest to the function at that point is the **tangent line** which has slope $f'(a)$ which we define below.

Definition 6.1.1.

Suppose $f : U \subseteq \mathbb{R} \to \mathbb{R}$ then we say that f has **derivative** $f'(a)$ defined by the limit below (if the limit exists, otherwise we say f is not differentiable at a)

$$f'(a) = \lim_{h \to 0} \frac{f(a+h) - f(a)}{h}$$

If f has a derivative at a then it also has a **differential** $df_a : \mathbb{R} \to \mathbb{R}$ at a which is a function defined by $df_a(h) = hf'(a)$. Finally, if f has derivative $f'(a)$ at a then the tangent line to the curve has equation $y = f(a) + f'(a)(x-a)$.

Notice that the derivative at a point is a number whereas the differential at a point is a linear map[1]. Also, the tangent line is a "*paralell translate*" of the line through the origin with slope $f'(a)$.

Example 6.1.2. . .

Definition 6.1.3.

Suppose $f : U \subseteq \mathbb{R} \to \mathbb{R}$ and suppose $f'(v)$ exists for each $v \in V \subset U$. We say that f has **derivative** $f' : V \subseteq \mathbb{R} \to \mathbb{R}$ defined by

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

for each $x \in V$.

In words, the derivative function is defined pointwise by the derivative at a point.

[1]We will maintain a similar distinction in the higher dimensional cases so I want to draw your attention to the distinction in terminology from the outset.

Proposition 6.1.4.

Suppose $a \in dom(f)$ where $f : dom(f) \subseteq \mathbb{R} \to \mathbb{R}$ and $a \in dom(f')$ then df_a is a linear transformation from $\mathbb{R}$ to $\mathbb{R}$.

Proof: Let $c, h, k \in \mathbb{R}$ and $a \in dom(f')$ which simply means $f'(a)$ is well-defined. Note that:

$$df_a(ch + k) = (ch + k)f'(a) = chf'(a) + kf'(a) = cdf_a(h) + df_a(k)$$

for all c, h, k thus df_a is linear transformation. $\square$

The **differential** is likewise defined to be the **differential form** $df : dom(f) \to L(\mathbb{R}, \mathbb{R}) = \mathbb{R}^*$ where $df(a) = df_a$ and df_a is a linear function from $\mathbb{R}$ to $\mathbb{R}$. We'll study differential forms in more depth in a later section.

6.1.2 derivatives of vector-valued functions of a real variable

A **vector-valued function of a real variable** is a mapping from a subset of $\mathbb{R}$ to some subset $\mathbb{R}^n$. In this section we discuss how to differentiate such functions as well as a few interesting theorems which are known for the various vector products.

We can revisit the start of Calculus III. We begin by defining the change in a vector-valued function f between the inputs a and $a + h$:

$$\Delta f = f(a + h) - f(a).$$

This is a vector. We can approximate this change for small values of h by replacing the space curve $a \mapsto f(a)$ with a line $t \mapsto f(a) + tf'(a)$ in $\mathbb{R}^n$. The direction vector of the tangent line is $f'(a)$ which we define below.

Definition 6.1.5.

Suppose $f : U \subseteq \mathbb{R} \to \mathbb{R}$ then we say that f has **derivative** $f'(a)$ defined by the limit below (if the limit exists, otherwise we say f is not differentiable at a)

$$f'(a) = \lim_{h \to 0} \frac{f(a + h) - f(a)}{h}$$

We define f' to be the function defined pointwise by the limit above for all such values as the limit converges. If f has a derivative at a then it also has a **differential** $df_a : \mathbb{R} \to \mathbb{R}^n$ at a which is a mapping defined by $df_a(h) = hf'(a)$. The vector-valued-differential form df is defined pointwise by $df(a) = df_a$ for all $a \in dom(f')$.

The tangent line is a "*paralell translate*" of the line through the origin with direction-vector $f'(a)$. In particular, if f has a derivative of $f'(a)$ at a then the tangent line to the curve has parametric equation $\vec{r}(t) = f(a) + tf'(a)$.

Proposition 6.1.6.

Suppose $a \in dom(f)$ where $f : dom(f) \subseteq \mathbb{R} \to \mathbb{R}^n$ and $a \in dom(f')$ then the differential df_a is a linear transformation from $\mathbb{R}$ to $\mathbb{R}^n$.

The proof is almost identical to the proof for real-valued functions of a real variable. Note:

$$df_a(ch + k) = (ch + k)f'(a) = chf'(a) + kf'(a) = cdf_a(h) + df_a(k)$$

for all $h, k, c \in \mathbb{R}$ hence df_a is a linear transformation.

6.1.3 directional derivatives

Let $m \geq n$, the image of a injective continuous mapping $F : dom(F) \subseteq \mathbb{R}^n \to \mathbb{R}^m$ gives an n-dimensional continuous surface in $\mathbb{R}^m$ provided the mapping F satisfy the topological requirement $dom(F) \approx \mathbb{R}^n$. This topological fine print is just a way to avoid certain pathological cases like space filling curves. We proved in Example 3.4.7 that the unit-sphere is a continuous surface. The proof that the sphere of radius 2 is a continuous surface is similar. In the example that follows we'll see how curves on the surface provide a definition for the tangent plane.

Example 6.1.7. *The sphere of radius 2 centered at the origin has equation $x^2 + y^2 + z^2 = 4$. We can view the top-half of the sphere as the image of the mapping $F : \mathbb{R}^2 \to \mathbb{R}^3$ where*

$$F(x, y) = \left(\ x, \ \ y, \ \ \sqrt{4 - x^2 - y^2} \ \right).$$

The tangent plane to the sphere at some point on the sphere can be defined as the set of all tangent vectors to curves on the sphere which pass through the point: let S be the sphere and $p \in S$ then the tangent space to p is intuitively defines as follows:

$$\boxed{T_pS = \{\gamma'(0) \mid \gamma : \mathbb{R} \to S, \ \text{a smooth curve with} \ \ \gamma(0) = p\}}$$

A line in the direction of $< a, b >$ through $(1, 1)$ in $\mathbb{R}^2$ has parametric representation $\vec{r}(t) = (1 + at, 1 + bt)$. We can construct curves on the sphere that pass through $F(1, 1) = (1, 1, \sqrt{2})$ by simply mapping the lines in the plane to curves on the sphere; $\gamma(t) = F(\vec{r}(t))$ which gives

$$\gamma(t) = \left(\ 1 + at, \ \ 1 + bt, \ \ \sqrt{4 - (1 + at)^2 - (1 + bt)^2} \ \right)$$

Now, not all curves through p have the same form as $\gamma(t)$ above but it is fairly clear that if we allow (a, b) to trace out all possible directions in $\mathbb{R}^2$ then we should cover T_pS. A short calculation reveals that

$$\gamma'(0) =< a, b, \tfrac{-1}{\sqrt{2}}(a + b) >$$

These are vectors we should envision as attached to the point $(1, 1, \sqrt{2})$. A generic point in the tangent plane to the point should have the form $p + \gamma'(0)$. This gives equations:

$$x = 1 + a, \qquad y = 1 + b, \qquad z = \sqrt{2} - \tfrac{1}{\sqrt{2}}(a + b)$$

we can find the Cartesian equation for the plane by eliminating a, b

$$a = x - 1, \qquad b = y - 1 \;\Rightarrow\; z = \sqrt{2} - \tfrac{1}{\sqrt{2}}(x + y - 2) \;\Rightarrow\; x + y + \sqrt{2}z = 4.$$

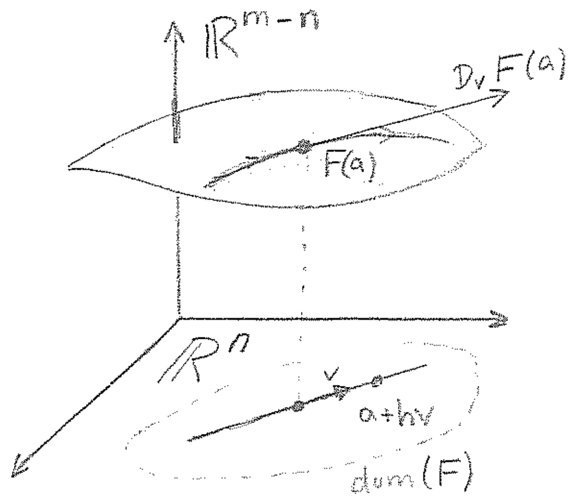
We find the tangent plane to the sphere $x^2 + y^2 + z^2 = 4$ has normal $< 1, 1, \sqrt{2} >$ at the point $(1, 1, \sqrt{2})$.

Of course there are easier ways to calculate the equation for a tangent plane. 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. In the case that $m = 1$ then $F : dom(F) \subseteq \mathbb{R}^n \to \mathbb{R}$ and the directional derivative gives the instantaneous rate of change of the function F at the point a along v. You probably insisted that $||v|| = 1$ in calculus III but we make no such demand here. We define the directional derivative for mappings and vectors of non-unit length.

Definition 6.1.8.

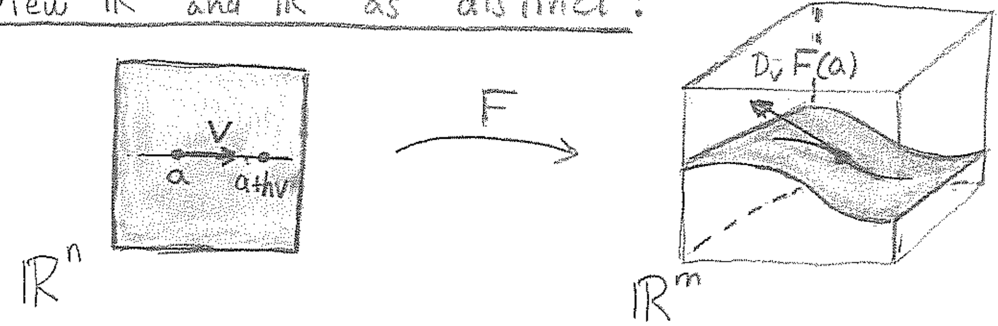
Let $F : dom(F) \subseteq \mathbb{R}^n \to \mathbb{R}^m$ and suppose the limit below exists for $a \in dom(F)$ and $v \in \mathbb{R}^n$ then we define the **directional derivative of F at a along v** to be $D_vF(a) \in \mathbb{R}^m$ where

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



you can picture $\mathbb{R}^n \subset \mathbb{R}^m$ when $m > n$. Typically we identify $\mathbb{R}^n$ with $\mathbb{R}^n \times \{\underbrace{(0,0,\dots,0)}_{m-n \text{ zeros.}}\} \subset \mathbb{R}^m$

Or, view $\mathbb{R}^n$ and $\mathbb{R}^m$ as distinct:



The directional derivative $D_vF(a)$ is homogenous in v.

Proposition 6.1.9.

Let $F : dom(F) \subseteq \mathbb{R}^n \to \mathbb{R}^m$ then if $D_vF(a)$ exists in $\mathbb{R}^m$ then $D_{cv}F(a) = cD_vF(a)$

See Edwards pg. 66 the proof is not hard. Let $F : U \to \mathbb{R}^m$ define a continuous surface S with dimension n. The tangent space of S at $p \in S$ should be the paralell translate of a n-dimensional subspace of $\mathbb{R}^m$. Moreover, we would like for the tangent space at a point $p \in S$ to be very close to the surface near that point. The change of F near $p = F(a)$ along the curve $\gamma(t) = F(a + tv)$ is given by

$$\Delta F = F(a + hv) - F(a).$$

It follows that $F(a + hv) \approxeq F(a) + hD_vF(a)$ for $h \approxeq 0$. We'd like for the the set of all directional derivatives at p to form a subspace of $\mathbb{R}^m$. Recall(or learn) that in linear algebra we learn that every subspaces of $\mathbb{R}^m$ is the range of some linear operator[2] This means that if $D_vF(a)$ was a linear operator with respect to v then we would know the set of all directional derivatives formed a subspace of $\mathbb{R}^m$. Note that directional derivative almost gives us linearity since its homogeneous but we also need the condition of additivity:

$$\boxed{D_{v+w}F(a) = D_vF(a) + D_wF(a) \qquad \text{additivity of directional derivative}}$$

This condition is familar. Recall that Propositions 6.1.4 and 6.1.6 showed the *differential* df_a was linear for $f : dom(f) \subseteq \mathbb{R} \to \mathbb{R}^m$. In fact the differential is the directional derivative in these special cases if we let $v = 1$; $D_1F(a) = dF_a(1)$ for $F : dom(F) \subseteq \mathbb{R} \to \mathbb{R}^m$ where $a \in dom(F')$. So we have already proved the directional derivative is linear in those special cases. Fortunately it's not so simple for a general mapping. We have to make an additional assumption if we wish for the tangent space to be well-defined.

Definition 6.1.10.

Suppose that U is open and $F : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ is a mapping the we say that F is **differentiable** at $a \in U$ iff there exists a linear mapping $L : \mathbb{R}^n \to \mathbb{R}^m$ such that

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

In such a case we call the linear mapping L the **differential at** a and we denote $L = dF_a$. The matrix of the differential is called the **derivative 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 $v \in \mathbb{R}^n$.

[2]don't believe it? Let $W \leq \mathbb{R}^m$ and choose a basis $\beta = \{f_1, \dots, f_n\}$ for W. You can verify that $L(v) = [f_1|f_2|\cdots|f_n|f_n|\cdots|f_n]v$ defines a linear transformation with $range(L) = Col[\beta] = W$.

The preceding definition goes hand in hand with the definition of the tangent space given below.

Definition 6.1.11.

Suppose that $U \approx \mathbb{R}^n$ is open and $F : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ is a mapping which is differentiable on U. If $rank(F'(a)) = n$ at each $a \in U$ then we say that $F(U)$ is a **differentiable surface of dimension** n. Also, a map such as F is said to be **regular**. Moreover, we define the tangent space to $S = F(U)$ at $p \in S$ to be the paralell translate of the subspace $Col(F'(a)) \leq \mathbb{R}^m$. A typical point in the tangent space at $p \in S$ has the form $p + F'(a)v$ for some $v \in \mathbb{R}^n$.

The condition that $rank(F'(a)) = n$ is the higher-dimensional analogue of the condition that the direction vector of a line must be nonzero for a line. If we want a genuine n-dimensional surface then there must be n-linearly independent vectors in the columns in the derivative matrix. If there were two columns which were linearly dependent then the subspace $W = \{F'(a)v \mid v \in \mathbb{R}^n\}$ would not be n-dimensional.

Remark 6.1.12.

If this all seems a little abstract, relax, the examples are in the next section. I want to wrap up the mostly theoretical aspects in this section then turn to more calculational ideas such as partial derivatives and the Jacobian matrix in the next section. We'll see that partial differentiation gives us an easy straight-forward method to calculate all the theoretical constructs of this section. Edwards has the calculations mixed with the theory, I've ripped them apart for better or worse. Also, we will discuss surfaces and manifolds independently in the next chapter. I wouldn't expect you to entirely understand them from the discussion in this chapter.

Example 6.1.13. *Let $F : \mathbb{R}^n \to \mathbb{R}^m$ be defined by $F(v) = p + Av$ for all $v \in \mathbb{R}^n$ where the matrix $A \in \mathbb{R}^{\ m\times n}$ such that $rank(A) = n$ and $p \in \mathbb{R}^m$. We can calculate that $[dF_a] = A$. Observe that for $x \in \mathbb{R}^n$,*

$$\lim_{h\to 0} \frac{F(x+h) - F(x) - A(h)}{||h||} = \lim_{h\to 0} \frac{Ax + Ah - Ax - Ah}{||h||} = 0.$$

Therefore, $dF_x(h) = Ah$ for each $x \in \mathbb{R}^n$ and we find $F(\mathbb{R}^n)$ is an differentiable surface of dimensional n. Moreover, we find that $F(\mathbb{R}^n)$ is its own tangent space, the tangent space is the paralell translate of $Col(A)$ to the point $p \in \mathbb{R}^m$. This is the higher dimensional analogue of finding the tangent line to a line, it's just the line again.

The directional derivative helped us connect the definition of the derivative of mapping with the derivative of a function of $\mathbb{R}$. We now turn it around. If we're given the derivative of a mapping then the directional derivative exists. The converse is not true, see Example 4 on page 69 of Edwards.

Proposition 6.1.14.

If $F : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at $a \in U$ then the directional derivative $D_vF(a)$ exists for each $v \in \mathbb{R}^n$ 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 $\mathbb{R}^n$, when it exists it follows that the limits that approach the origin along particular paths also exist and are zero. In particular we can 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 $v \in \mathbb{R}^n$ since the $v = 0$ case follows trivially. $\square$

6.2 partial derivatives and the existence of the derivative

Definition 6.2.1.

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 we may use the notation $D_{e_i}F(a) = D_iF(a)$ or $\partial_i F = \frac{\partial F}{\partial x_i}$ when convenient. We also construct the partial derivative mapping $\partial_i F : V \subseteq \mathbb{R}^n \to \mathbb{R}^m$ as the mapping 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

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

Proposition 6.2.2.

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. $\square$

Proposition 6.2.3.

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),

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

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 6.2.4. *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. 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. Incidentally, those notes may well be better than these in certain respects.

6.2.1 examples of derivatives

Our goal here is simply to exhbit 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 indepedence is important. We'll study the case the variables are not independent in a later section.

Remark 6.2.5.

> I have put remarks about the rank of the derivative in the examples below. Of course this has nothing to do with the process of calculating Jacobians. It's something to think about once we master the process of calculating the Jacobian. Ignore the red comments for now if you wish

Example 6.2.6. *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 6.2.7. *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 6.2.8. *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 6.2.9. *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, $rank(F'(x,y,z)) = 3$ 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 6.2.10. *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,

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

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.

$$M_1 = \begin{bmatrix} 2x & 0 \\ 0 & z \end{bmatrix} \qquad M_2 = \begin{bmatrix} 2x & 2z \\ 0 & y \end{bmatrix} \qquad M_3 = \begin{bmatrix} 0 & 2z \\ z & y \end{bmatrix}$$

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 6.2.11. *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,

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

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:

$$M_1 = \begin{bmatrix} 2x & 2y \\ y & x \end{bmatrix} \qquad M_2 = \begin{bmatrix} 2x & 2y \\ 1 & 1 \end{bmatrix} \qquad M_3 = \begin{bmatrix} y & x \\ 1 & 1 \end{bmatrix}$$

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 6.2.12. *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,

$$P'(x, v, m) = \begin{bmatrix} kx & mv & \frac{1}{2}v^2 \\ 0 & m & v \end{bmatrix}$$

Example 6.2.13. *Let $F(r, \theta) = (r\cos\theta, r\sin\theta)$. We calculate,*

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

Hence,

$$F'(r, \theta) = \begin{bmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{bmatrix}$$

We calculate $det(F'(r, \theta)) = r$ thus this mapping has full rank everywhere except the origin.

Example 6.2.14. *Let $G(x, y) = (\sqrt{x^2 + y^2}, \tan^{-1}(y/x))$. We calculate,*

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

Hence,

$$G'(x, y) = \begin{bmatrix} \frac{x}{\sqrt{x^2+y^2}} & \frac{y}{\sqrt{x^2+y^2}} \\ \frac{-y}{x^2+y^2} & \frac{x}{x^2+y^2} \end{bmatrix} = \begin{bmatrix} \frac{x}{r} & \frac{y}{r} \\ \frac{-y}{r^2} & \frac{x}{r^2} \end{bmatrix} \qquad \left(\text{using } r = \sqrt{x^2 + y^2} \right)$$

We calculate $det(G'(x, y)) = 1/r$ thus this mapping has full rank everywhere except the origin.

Example 6.2.15. *Let $F(x,y) = (x, y, \sqrt{R^2 - x^2 - y^2})$ for a constant R. We calculate,*

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

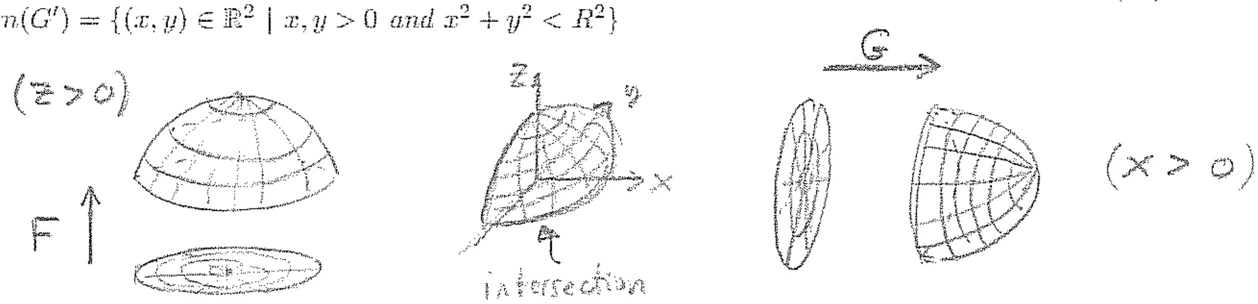
Also, $\nabla x = (1,0)$ and $\nabla y = (0,1)$ thus

$$F'(x,y) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ \frac{-x}{\sqrt{R^2-x^2-y^2}} & \frac{-y}{\sqrt{R^2-x^2-y^2}} \end{bmatrix}$$

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 6.2.16. *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 6.2.17. *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 6.2.18. *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 6.2.19. *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 $\text{rank}(f') = 3$.

Example 6.2.20. *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.

Example 6.2.21. *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 6.2.22. *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.*

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

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,

$$det\begin{bmatrix} x_u & x_v \\ y_u & y_v \end{bmatrix} \quad det\begin{bmatrix} x_u & x_v \\ z_u & z_v \end{bmatrix} \quad det\begin{bmatrix} y_u & y_v \\ z_u & z_v \end{bmatrix}$$

Example 6.2.23. . .

Example 6.2.24. . .

6.2.2 sick examples and continuously differentiable mappings

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[3] examples.

Example 6.2.25. *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 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 6.2.26. *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

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

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.

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

Definition 6.2.27.

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

The defintion above is interesting because of the proposition below. The import of the proposition is that we can build the tangent plane from the Jacobian matrix provided the partial derivatives are all continuous. This is a very nice result because the concept of the linear mapping is quite abstract but partial differentiation of a given mapping is easy.

Proposition 6.2.28.

If F is continuously differentiable at a then F is differentiable at a

We'll follow the proof in Edwards on pages 72-73.

Let $f = F \cdot e_p$ for some $p \in \{1, 2, \dots, m\}$. We seek to show continuous diff. of F at $a \Rightarrow f$ is diff. at a. Lemma 2.3 of Edwards then says $dF = (dF_1, \dots, dF_m)$ provided we have $dF_p = df$. (we need to show all components of F are differentiable at a)

Following Edwards (and other texts on this subject), let $h = (h_1, h_2, \dots, h_n)$ and note $h = \sum_{k=0}^{n} h_k e_k$ and we could break up h as follows,

$$h = \vec{h}_0 + h_1 e_1 + \cdots + h_n e_n = \vec{h}_1 + h_2 e_2 + \cdots + h_n e_n = \vec{h}_k + h_{k+1} e_{k+1} + \cdots + h_n e_n .$$

Anyway, since $\vec{h}_0 = 0$,

$$\begin{aligned} f(a+h) - f(a) &= \sum_{k=1}^{n} \left[f(a + \vec{h}_k) - f(a + h_{k-1}) \right] \\ &= \cancel{f(a+\vec{h}_1)} - f(a + \vec{h}_0) + \cancel{f(a+\vec{h}_1)} - \cancel{f(a+\vec{h}_2)} + \cdots + f(a+\vec{h}_n) - \cancel{f(a+\vec{h}_{n-1})} \\ &= f(a + \vec{h}_n) - f(a). \end{aligned}$$

Notice that,

$$f(a + \vec{h}_i) - f(a + \vec{h}_{i-1}) = \underbrace{f(a_1 + h_1, \dots, a_i + h_i, a_{i+1}, \dots, a_n) - f(a_1 + h_1, \dots, a_{i-1} + h_{i-1}, a_i, \dots, a_n)}$$

entries match except at i^{th} position where we have $a_i + h_i$ verses a_i. Recognize this is $\Delta g\,(a_1 + h_1, \dots, a_{i-1} + h_{i-1}, a_i, \dots, a_n)$

Let $g(x) = f(a_1, \dots, a_{i-1}, x, a_{i+1}, \dots, a_n)$ then $g'(x) = D_i f(a_1, \dots, x, \dots, a_n)$ which exists by assumption of continuous differentiability of F. Apply the mean value $Th^{\underline{m}}$ to conclude $\exists\, b_i$ such that $f(a + \vec{h}_i) - f(a + \vec{h}_{i-1}) = h_i D_i f(b_i)$

Thus,

$$\lim_{h \to 0} \frac{\| f(a+h) - f(a) - \sum_{i=1}^{n} D_i f(a) h_i \|}{\|h\|} = \lim_{h \to 0} \frac{\| \sum_{i=1}^{n} (D_i f(b_i) - D_i f(a)) h_i \|}{\|h\|}$$

$$\leq \lim_{h \to 0} \sum_{i=1}^{n} | D_i f(b_i) - D_i f(a) | \frac{|h_i|}{\|h\|}$$

Uses continuity of partials again $\longrightarrow$ $\displaystyle \leq \lim_{h \to 0} \sum_{i=1}^{n} \| D_i f(b_i) - D_i f(a) \| = 0.$

6.3 properties of the derivative

Of course much of what we discover in this section should be old news to you if you understood differentiation in calculus III. However, in our current context we have efficient methods of proof and the langauge of linear algebra allows us to summarize pages of calculations in a single line.

6.3.1 additivity and homogeneity of the derivative

Suppose $F_1 : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ and $F_2 : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$. Furthermore, suppose both of these are differentiable at $a \in \mathbb{R}^n$. It follows that $(dF_1)_a = L_1$ and $(dF_2)_a = L_2$ are linear operators from $\mathbb{R}^n$ to $\mathbb{R}^m$ which approximate the change in F_1 and F_2 near a, in particular,

$$\lim_{h\to 0} \frac{F_1(a+h) - F_1(a) - L_1(h)}{||h||} = 0 \qquad \lim_{h\to 0} \frac{F_2(a+h) - F_2(a) - L_2(h)}{||h||} = 0$$

To prove that $H = F_1 + F_2$ is differentiable at $a \in \mathbb{R}^n$ we need to find a differential at a for H. Naturally, we expect $dH_a = d(F_1 + F_2)_a = (dF_1)_a + (dF_2)_a$. Let $L = (dF_1)_a + (dF_2)_a$ and consider,

$$\begin{aligned}
\lim_{h\to 0} \tfrac{H(a+h)-H(a)-L(h)}{||h||} &= \lim_{h\to 0} \tfrac{F_1(a+h)+F_2(a+h)-F_1(a)-F_2(a)-L_1(h)-L_2(h)}{||h||} \\
&= \lim_{h\to 0} \tfrac{F_1(a+h)-F_1(a)-L_1(h)}{||h||} + \lim_{h\to 0} \tfrac{F_2(a+h)-F_2(a)-L_2(h)}{||h||} \\
&= 0 + 0 \\
&= 0
\end{aligned}$$

Note that breaking up the limit was legal because we knew the subsequent limits existed and were zero by the assumption of differentiability of F_1 and F_2 at a. Finally, since $L = L_1 + L_2$ we know L is a linear transformation since the sum of linear transformations is a linear transformation. Moreover, the matrix of L is the sum of the matrices for L_1 and L_2. Let $c \in \mathbb{R}$ and suppose $G = cF_1$ then we can also show that $dG_a = d(cF_1)_a = c(dF_1)_a$, the calculation is very similar except we just pull the constant c out of the limit. I'll let you write it out. Collecting our observations:

Proposition 6.3.1.

Suppose $F_1 : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ and $F_2 : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ are differentiable at $a \in U$ then $F_1 + F_2$ is differentiable at a and

$$\boxed{d(F_1 + F_2)_a = (dF_1)_a + (dF_2)_a} \quad \text{or} \quad \boxed{(F_1 + F_2)'(a) = F_1'(a) + F_2'(a)}$$

Likewise, if $c \in \mathbb{R}$ then

$$\boxed{d(cF_1)_a = c(dF_1)_a} \quad \text{or} \quad \boxed{(cF_1)'(a) = c(F_1'(a))}$$

6.3.2 product rules?

What sort of product can we expect to find among mappings? Remember two mappings have vector outputs and there is no way to multiply vectors in general. Of course, in the case we have two mappings that have equal-dimensional outputs we could take their dot-product. There is a product rule for that case: if $\vec{A}, \vec{B} : \mathbb{R}^n \to \mathbb{R}^m$ then

$$\boxed{\partial_j(\vec{A} \cdot \vec{B}) = (\partial_j \vec{A}) \cdot \vec{B}) + \vec{A} \cdot (\partial_j \vec{B})}$$

Or in the special case of $m = 3$ we could even take their cross-product and there is another product rule in that case:

$$\boxed{\partial_j(\vec{A} \times \vec{B}) = (\partial_j \vec{A}) \times \vec{B} + \vec{A} \times (\partial_j \vec{B})}$$

What other case can we "multiply" vectors? One very important case is $\mathbb{R}^2 = \mathbb{C}$ where is is customary to use the notation $(x, y) = x + iy$ and $f = u + iv$. If our range is complex numbers then we again have a product rule: if $f : \mathbb{R}^n \to \mathbb{C}$ and $g : \mathbb{R}^n \to \mathbb{C}$ then

$$\boxed{\partial_j(fg) = (\partial_j f)g + f(\partial_j g)}$$

I have relegated the proof of these product rules to the end of this chapter. One other object worth differentiating is a matrix-valued function of $\mathbb{R}^n$. If we **define** the partial derivative of a matrix to be the matrix of partial derivatives then partial differentiation will respect the sum and product of matrices (we may return to this in depth if need be later on):

$$\boxed{\partial_j(A + B) = \partial_j B + \partial_j B} \qquad \boxed{\partial_j(AB) = (\partial_j A)B + A(\partial_j B)}$$

Moral of this story? If you have a pair mappings whose ranges allow some sort of product then it is entirely likely that there is a corresponding product rule [4]. There is one product rule which we can state for arbitrary mappings, note that we can always sensibly multiply a mapping by a function. Suppose then that $G : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ and $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$ are differentiable at $a \in U$. It follows that there exist linear transformations $L_G : \mathbb{R}^n \to \mathbb{R}^m$ and $L_f : \mathbb{R}^n \to \mathbb{R}$ where

$$\lim_{h \to 0} \frac{G(a+h) - G(a) - L_G(h)}{||h||} = 0 \qquad \lim_{h \to 0} \frac{f(a+h) - f(a) - L_f(h)}{h} = 0$$

Since $G(a+h) \approx G(a) + L_G(h)$ and $f(a+h) \approx f(a) + L_f(h)$ we expect

$$\begin{aligned} fG(a+h) &\approx (f(a) + L_f(h))(G(a) + L_G(h)) \\ &\approx (fG)(a) + \underbrace{G(a)L_f(h) + f(a)L_G(h)}_{\text{linear in } h} + \underbrace{L_f(h)L_G(h)}_{2^{nd} \text{ order in } h} \end{aligned}$$

[4] In my research I consider functions on supernumbers, these also can be multiplied. Naturally there is a product rule for super functions, the catch is that super numbers z, w do not necessarily commute. However, if they're homogeneneous $zw = (-1)^{\epsilon_w \epsilon_z} wz$. Because of this the super product rule is $\partial_M(fg) = (\partial_M f)g + (-1)^{\epsilon_f \epsilon_M} f(\partial_M g)$

Thus we propose: $L(h) = G(a)L_f(h) + f(a)L_G(h)$ is the best linear approximation of fG.

$$\begin{aligned}
\lim_{h\to 0}&\frac{(fG)(a+h) - (fG)(a) - L(h)}{||h||} = \\
&= \lim_{h\to 0}\frac{f(a+h)G(a+h) - f(a)G(a) - G(a)L_f(h) - f(a)L_G(h)}{||h||} \\
&= \lim_{h\to 0}\frac{f(a+h)G(a+h) - f(a)G(a) - G(a)L_f(h) - f(a)L_G(h)}{||h||} + \\
&\quad + \lim_{h\to 0}\frac{f(a)G(a+h) - G(a+h)f(a)}{||h||} \\
&\quad + \lim_{h\to 0}\frac{f(a+h)G(a) - G(a)f(a+h)}{||h||} \\
&\quad + \lim_{h\to 0}\frac{f(a)G(a) - G(a)f(a)}{||h||} \\
&= \lim_{h\to 0}\bigg[f(a)\frac{G(a+h) - G(a) - L_G(h)}{||h||} + \frac{f(a+h) - f(a) - L_f(h)}{||h||}G(a) + \\
&\qquad + \Big(f(a+h) - f(a)\Big)\frac{G(a+h) - G(a)}{||h||}\bigg] \\
&= f(a)\left[\lim_{h\to 0}\frac{G(a+h) - G(a) - L_G(h)}{||h||}\right] + \left[\lim_{h\to 0}\frac{f(a+h) - f(a) - L_f(h)}{||h||}\right]G(a) \\
&= 0
\end{aligned}$$

Where we have made use of the differentiability and the consequent continuity of both f and G at a. Furthermore, note

$$\begin{aligned}
L(h + ck) &= G(a)L_f(h + ck) + f(a)L_G(h + ck) \\
&= G(a)(L_f(h) + cL_f(k)) + f(a)(L_G(h) + cL_G(k)) \\
&= G(a)L_f(h) + f(a)(L_G(h) + c(G(a)L_f(k) + f(a)L_G(k)) \\
&= L(h) + cL(k)
\end{aligned}$$

for all $h, k \in \mathbb{R}^n$ and $c \in \mathbb{R}$ hence $L = G(a)L_f + f(a)L_G$ is a linear transformation. We have proved (most of) the following proposition:

Proposition 6.3.2.

If $G : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ and $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$ are differentiable at $a \in U$ then fG is differentiable at a and

$G(a)\,f'(a)$

$$\boxed{d(fG)_a = (df)_a G(a) + f(a)dG_a} \qquad \boxed{(fG)'(a) = \cancel{f'(a)G(a)} + f(a)G'(a)}$$

The argument above covers the ordinary product rule and a host of other less common rules. Note again that $G(a)$ and $G'(a)$ are vectors.

$$\underbrace{G(a)}_{m\times 1}\ \underbrace{f'(a)}_{1\times n} = m\times n$$

probably should write

$$d(fG)_a = G(a)\,(df)_a + f(a)(dG)_a$$

but, it's ok as is as well.

6.4 chain rule

The proof in Edwards is on 77-78. I'll give a heuristic proof here which captures the essence of the argument. The simplicity of this rule continues to amaze me.

Proposition 6.4.1.

If $F : U \subseteq \mathbb{R}^n \to \mathbb{R}^p$ is differentiable at a and $G : V \subseteq \mathbb{R}^p \to \mathbb{R}^m$ is differentiable at $F(a) \in V$ then $G \circ F$ is differentiable at a and

$\boxed{d(G \circ F)_a = (dG)_{F(a)} \circ dF_a}$ or, in matrix notation, $\boxed{(G \circ F)'(a) = G'(F(a))F'(a)}$

Proof Sketch:

$$F(a+h) \approx F(a) + dF_a(h) = F(a) + F'(a)h$$
$$G(z+k) \approx G(z) + dG_z(k) = G(z) + G'(z)k$$

Consider then,

$$\begin{aligned}(G \circ F)(a+h) &= G(F(a+h)) \\ &\approx G(F(a) + F'(a)h) \quad : \text{let } z = F(a) \text{ and } k = F'(a)h, \text{ note } h \text{ small} \Rightarrow k \text{ small.} \\ &= G(z+k) \\ &\approx G(z) + G'(z)k \quad \leftarrow \text{hence this reasonable} \\ &= G(F(a)) + G'(F(a))F'(a)h\end{aligned}$$

Thus $\Delta(G \circ F) \approx G'(F(a))F'(a)h \Rightarrow \underline{d(G \circ F) = dG \circ dF}$.

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

We have $f : \mathbb{R}^2 \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}^2$ 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!

Remark: I omit some point dependence to reduce clutter

Example 6.4.3. . . . $\underbrace{f : \mathbb{R} \to \mathbb{R}^n}$ and $g : \mathbb{R}^n \to \mathbb{R}$

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

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

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

Example 6.4.4. . . .

Recall, $x = x(u,v,w)$ $y = y(u,v,w)$, $f = f(x,y)$ then,

$$\left.\begin{aligned} \frac{\partial f}{\partial u} &= \frac{\partial x}{\partial u}\frac{\partial f}{\partial x} + \frac{\partial y}{\partial u}\frac{\partial f}{\partial y} \\ \frac{\partial f}{\partial v} &= \frac{\partial x}{\partial v}\frac{\partial f}{\partial x} + \frac{\partial y}{\partial v}\frac{\partial f}{\partial y} \\ \frac{\partial f}{\partial w} &= \frac{\partial x}{\partial w}\frac{\partial f}{\partial x} + \frac{\partial y}{\partial w}\frac{\partial f}{\partial y} \end{aligned}\right\}$$

note we derive these via matrix multiplication in example below

Example 6.4.5. . . .

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

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

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

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

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

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

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

Example 6.4.6. . .

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

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

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

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

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

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

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

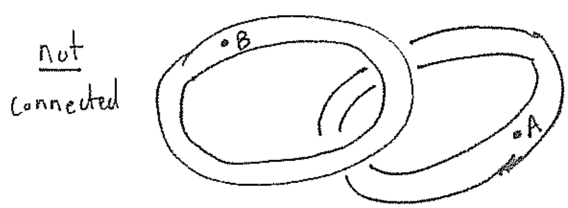
6.4.1 theorems

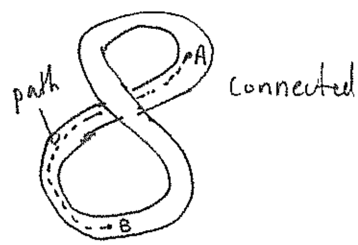
The goal of this section is to prove the partial derivatives commute for nice functions. Of course some of the results we discuss on the way to that goal are interesting in their own right as well.

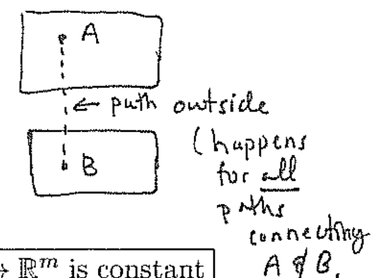
Definition 6.4.8.

We say $U \subseteq \mathbb{R}^n$ is **path connected** iff any two points in U can be connected by a path which is contained within the set.

For example, $\mathbb{R}^n$ is connected since given any two points $a, b \in \mathbb{R}^n$ we can construct the path $\phi(t) = a + t(b-a)$ from a to b and naturally the path is within the set. You can easily verify that open and closed balls are also path connected. Even a donut is path connected. However, a pair donuts is not path connected unless it's one of those artsy figure-8 deals.







Proposition 6.4.9.

If U is a connected open subset of $\mathbb{R}^n$ then a differenitable mapping $F : U \to \mathbb{R}^m$ is constant iff $F'(u) = 0$ for all $u \in U$.

(?) ☹

Let $A, B \in U$ and connect these points with a smooth path $\gamma : [a,b] \longrightarrow U$.
(perhaps we should insist path connected $\Rightarrow$ smooth paths exist inside U...)
Consider, $F(\gamma(t)) = g(t)$. If F is constant then $g(t)$ is constant
hence $F'(\gamma(t))\gamma'(t) = 0$ but γ smooth $\Rightarrow \gamma'(t) \neq 0$ hence
$F'(\gamma(a)) = F'(a) = 0 \Rightarrow F'(a) = 0 \ \forall a \in U$. Conversely
if $F'(u) = 0 \ \forall u \in U \Rightarrow \frac{d}{dt}[g(t)] = \underbrace{F'(\gamma(t))}_{\text{zero}}\gamma'(t) = 0$
Thus $g(a) = g(b) \Rightarrow F(\gamma(a)) = F(\gamma(b))$
$\Rightarrow F(A) = F(B) \quad \forall A, B \in U \therefore F$ constant on U.

Proposition 6.4.10.

☹

If U is a connected open subset of $\mathbb{R}^n$ and the differenitable mappings $F, G : U \to \mathbb{R}$ such that $F'(x) = G'(x)$ for all $x \in U$ then there exists a constant vector $c \in \mathbb{R}^m$ such that $F(x) = G(x) + c$ for all $x \in U$.

Construct $H(x) = F(x) - G(x)$ and note $H'(x) = F'(x) - G'(x) = 0 \quad \forall x \in U$.
By the previous prop. $H(x) = c \quad \forall x \in U$ hence
$$F(x) = G(x) + c \quad \forall x \in U.$$

Remark: if U was not connected we could only conclude that the F was constant on connected subsets of U.
In topology one discusses breaking down a space into its path components, these are maximal connected subsets.

There is no mean value theorem for mappings since counter-examples exist. For example, Exercise 1.12 on page 63 shows the mean value theorem fails for the helix. In particular, you can find average velocity vector over a particular time interval such that the velocity vector never matches the average velocity over that time period. Fortunately, if we restrict our attention to mappings with one-dimensional codomains we still have a nice theorem:

Proposition 6.4.11. *(Mean Value Theorem)*

Suppose that $f : U \to \mathbb{R}$ is a differentiable function and U is an open set. Furthermore, suppose U contains the line segment from a to b in U;

$$L_{a,b} = \{a + t(b-a) \mid t \in [0,1]\} \subset U.$$

It follows that there exists some point $c \in L_{a,b}$ such that

$$f(b) - f(a) = f'(c)(b-a).$$

The proof follows from the construction of a function on $\mathbb{R}$ to which the elementary mean value Thm is applied. Let

$$g(t) = f(a + t(b-a)) \quad \text{for } 0 \leq t \leq 1.$$

Or if you prefer, construct $\varphi(t) = a + t(b-a)$ which parametrizes the line segment from a to b. Clearly $\varphi'(t) = b - a$ and by the chain-rule,

$$g'(t) = f'(\varphi(t))\varphi'(t)$$

Note $g : [0,1] \xrightarrow{\varphi} U \xrightarrow{f} \mathbb{R}$ is differentiable on $[0,1]$ thus the MVT gives $c_0 \in [0,1]$ such that $g(1) - g(0) = g'(c_0)$.

Thus, $$f(b) - f(a) = g(1) - g(0) = g'(c_0) = f'(\varphi(c_0))\varphi'(c_0) = f'(c)\cdot(b-a).$$

$$c = \varphi(c_0).$$

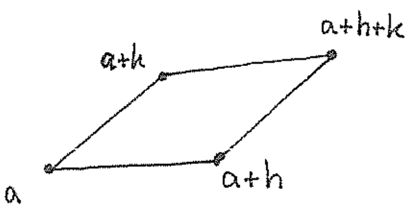
Definition 6.4.12. *(higher derivatives)*

We define nested directional derivatives in the natural way:

$$D_k D_h f(x) = D_k(D_h f(x)) = \lim_{t \to 0} \frac{D_h f(x+tk) - D_h f(x)}{t}$$

Furthermore, the **second difference** is defined by

$$\Delta^2 f_a(h,k) = f(a+h+k) - f(a+h) - f(a+k) + f(a)$$



note by MVT: $f(a+h) - f(a) = f'(a+\alpha h)h$

$f(a+h+k) - f(a+k) = f'(a+k+\beta h)h$

$$\Delta^2 f_a(h,k) = \Delta f_a(h) - \Delta f_{a+k}(h)$$

This is Lemma 3.5 on page 86 of Edwards.

Proposition 6.4.13.

Suppose U us an open set and $f : U \to \mathbb{R}$ which is differentiable on U with likewise differentiable directional derivative function on U. Suppose that $a, a+h, a+k, a+h+k$ are all in U then there exist $\alpha, \beta \in (0,1)$ such that

$$\Delta^2 f_a(h,k) = D_k D_h f(a + \alpha h + \beta k).$$

The proof is rather neat. The α and β stem from two applications of the MVT, once for the function then once for its directional derivative.

Let $g(x) = f(x+k) - f(x)$ then $dg_x = df_{x+k} - df_x$.

Furthermore, using $\Delta^2 f_a(h,h) = f(a+h+k) - f(a+h) - f(a+h) + f(a)$ note,

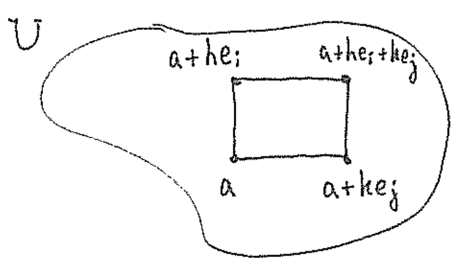
following Edwards pg. 87.

$$
\begin{aligned}
\Delta^2 f_a(h,k) &= g(a+h) - g(a) \\
&= g'(a+\alpha h)h \quad : \text{by MVT } \exists\, \alpha \in (0,1). \\
&= (D_h g)(a+\alpha h) \quad : \text{diff} \Rightarrow \text{directional derivative exist.} \\
&= dg_{a+\alpha h}(h) \\
&= df_{a+\alpha h+k}(h) - df_{a+\alpha h}(h) \\
&= D_h f(a+\alpha h+k) - D_h f(a+\alpha h) \\
&= (D_h f)'(a+\alpha h+\beta k)(k) \quad \text{for some } \beta \in (0,1) \text{ by MVT again} \\
&= (D_k D_h f)(a+\alpha h+\beta k) \quad : \text{unraveling notation.}
\end{aligned}
$$

Proposition 6.4.14.

Let U be an open subset of $\mathbb{R}^n$. If $f : U \to \mathbb{R}$ is a function with continuous first and second partial derivatives on U then for all $i, j = 1, 2, \dots, n$ we have $D_i D_j f = D_j D_i f$ on U;

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}.$$



$$\Delta^2 f_a(he_i, ke_j) = D_{ke_j} D_{he_i} f(a + \alpha_1 he_i + \beta_1 ke_j)$$

$$\Delta^2 f_a(ke_j, he_i) = D_{he_i} D_{ke_j} f(a + \alpha_2 ke_j + \beta_2 he_i)$$

Then $\Delta^2 f_a(he_i, ke_j) = \Delta^2 f_a(ke_j, he_i)$

$$\Rightarrow D_j D_i f(a) = D_i D_j f(a)$$

(pull out h, k using homogeneity of $D_{cv} f = c D_v f$)
take limit $h, k \to 0$ to drop the h, k inside the $Df(\_\_)$ terms. Can do this by continuity of partial derivatives near a.

6.5 differential forms and differentials

Definition 6.5.1.

A **form field** on $\mathbb{R}$ is a function from $\mathbb{R}$ to the space of all linear maps from $\mathbb{R}$ to $\mathbb{R}$. In other words, a form field assigns a dual vector at each point in $\mathbb{R}$. Remember that $\mathbb{R}^* = \{f : \mathbb{R} \to \mathbb{R} \mid f \text{ is a linear function}\}$. We call α a **differential one-form** or **differential form** if α can be written as $\alpha = \alpha_1 dx$ for some smooth function α_1.

The definition above is probably unecessary for this section. I give it primarily for the sake of making a larger trend easier to grasp later on. Feel free to ignore it for now.

6.5.1 differential form notation

Let $g(x) = x$ for all $x \in \mathbb{R}$. Note that $g'(x) = 1$ and it follows that $dg_a(x) = 1 \cdot x = x$ for all $x \in \mathbb{R}$. Therefore, $dg = g$. If we denote $g = x$ so that $dx = x$ in this notation. Note then we can write the differential in terms of the derivative function:

$$df(a)(h) = df_a(h) = f'(a)h = f'(a)dx_a(h) \qquad \text{for all } h \in \mathbb{R}$$

Hence $df(a) = f'(a)dx_a$ for all $a \in \mathbb{R}$ hence $df = f'dx$ or we could denote this by the deceptively simple formula $df = \frac{df}{dx}dx$. Thus the differential notation introduced in this section is in fact consistent with our usual notation for the derivative from calculus I. However, df and dx are actually differential forms in this viewpoint so I'm not so sure that df/dx really makes sense anymore. In retrospect, the main place we shift differentials around as if they are tiny real numbers is in the calculations of u-substitution or separation of variables. In both of those cases the differential notation serves as a shorthand for the application of a particular theorem. Just as in calculus III the differentials dx, dy, dz in the line integral $\int_C pdx + qdy + rdz$ provide a notational shortand for the rigorous definition in terms of a path covering the curve C.

Differentials are notational devices in calculus, one should be careful not to make more of them then is appropriate for a given context. That said, if you adopt the view point that dx, dy, dz are differential forms and their product is properly defined via a wedge product then the wedge product together with the total differential (to be discussed in the next section) will generate the formulas for coordinate change. Let me give you a taste:

$$\begin{aligned} dx \wedge dy &= d(r\cos(\theta)) \wedge d(r\sin(\theta)) \\ &= [\cos(\theta)dr - r\sin(\theta)d\theta] \wedge [\sin(\theta)dr + r\cos(\theta)d\theta] \\ &= r\cos^2(\theta)dr \wedge d\theta - r\sin^2(\theta)d\theta \wedge dr \\ &= rdr \wedge d\theta \end{aligned}$$

where I used that $dr \wedge d\theta = -d\theta \wedge dr$, $dr \wedge dr = 0$ and $d\theta \wedge d\theta = 0$ because of the antisymmety of the **wedge product** $\wedge$. In calculus III we say for polar coordinates the Jacobian is $\frac{\partial(x,y)}{\partial(r,\theta)} = r$. The determinant in the Jacobian is implicitly contained in the algebra of the wedge product. If you want to change coordinates in differential form notation you just substitute in the coordinate change formulas and take a few total differentials then the wedge product does the rest. In other words, the Jacobian change of coordinates formula is naturally encoded in the langauge of differential forms.

6.5.2 linearity properties of the derivative

Proposition 6.5.2.

Suppose that f, g are functions such that their derivative functions f' and g' share the same domain U then $(f+g)' = f' + g'$ and $(cf)' = cf'$. Moreover, the differentials of those functions have

$$d(f+g) = df + dg \qquad \text{and} \qquad d(cf) = cdf$$

Proof: The proof that $(f+g)' = f' + g'$ and $(cf)' = cf'$ follows from earlier general arguments in this chapter. Consider that,

$$\begin{aligned} d(f+g)_a(h) &= h(f+g)'(a) && \text{def. of differential for } f+g \\ &= h(f'(a)+g'(a)) && \text{using linearity of derivative.} \\ &= df_a(h) + dg_a(h) && \text{algebra and def. of differential for } f \text{ and } g. \\ &= (df+dg)_a(h) && \text{def. of sum of functions.} \end{aligned}$$

thus $d(f+g) = df + dg$ and the proof that $d(cf) = cdf$ is similar. $\square$.
We see that properties of the derivative transfer over to corresponding properties for the differential. Problem 1.7 on pg 62-63 of Edwards asks you to work out the product and chain rule for differentials.

6.5.3 the chain rule revisited

Proposition 6.5.3.

Suppose that $f : dom(f) \to range(f)$ and $g : dom(g) \to range(g)$ are functions such that g is differentiable on U and f differentiable on $g(U)$ then

$$(f \circ g)'(a) = g'(a)f'(g(a))$$

for each $a \in U$ and it follows $d(f \circ g)_a = df_{g(a)} \circ dg_a$.

An intuitive proof is this: the derivative of a composite is the slope of the tangent line to the composite. However, if f_1 and f_2 are linear functions with slopes m_1 and m_2 then $f_1 \circ f_2$ is a linear function with slope $m_1 m_2$. Therefore, the derivative of a composite is the product of the derivatives of the inside and outside function and we are forced to evaluate the outside function at $g(a)$ since that's the only thing that makes sense[5]. Finally,

$$d(f \circ g)(a)(h) = h(f \circ g)'(a) = hg'(a)f'(g(a)) = df_{g(a)}(hg'(a)) = df_{g(a)}(dg_a(h)) = (df_{g(a)} \circ dg_a)(h)$$

Therefore we find $d(f \circ g)_a = df_{g(a)} \circ dg_a$.

Proof: Let $a \in U$ then $g'(a) = \lim_{h\to 0} \frac{g(a+h)-g(a)}{h}$ thus $\lim_{h\to 0} g(a+h) = \lim_{h\to 0} g(a) + hg'(a)$. In other words, the function g and it's tangent line are equal in the limit you approach the point

[5]this is argument by inevitability, see Agent Smith for how this turns out as a pattern of deduction.

of tangency. Likewise, $f'(g(a)) = \lim_{\delta\to 0} \frac{f(g(a)+\delta)-f(g(a))}{\delta}$ hence $\lim_{\delta\to 0} f(g(a)+\delta) = f(g(a)) + \delta f'(g(a))$. Calculate then,

$$\begin{aligned}
(f\circ g)'(a) &= \lim_{h\to 0} \tfrac{(f\circ g)(a+h)-(f\circ g)(a)}{h} && \text{defn. of derivative} \\
&= \lim_{h\to 0} \tfrac{f(g(a+h))-f(g(a))}{h} && \text{defn. of } f\circ g \\
&= \lim_{h\to 0} \tfrac{f(g(a)+hg'(a))-f(g(a))}{h} && \text{since } g(a+h)\approx g(a)+hg'(a) \\
&= g'(a)\lim_{\delta\to 0} \tfrac{f(g(a)+\delta)-f(g(a))}{\delta} && \text{made subst. } \delta = g'(a)h \\
&= g'(a)\lim_{\delta\to 0} \tfrac{f(g(a))+\delta f'(g(a))-f(g(a))}{\delta} && \text{as } f(g(a)+\delta)\approx f(g(a))+\delta f'(g(a)) \\
&= g'(a)f'(g(a)) && \text{limit of constant is just the constant.}
\end{aligned}$$

I have used the notation $\approx$ to indicate that those equations were not precisely true. However, the error is small when h or δ are close to zero and that is precisely the case which we were faced with in those calculations. Admittably we could give a more rigoruous proof in terms of ϵ and δ but this proof suffices for our purposes here. The main thing I wanted you to take from this is that the chain rule is a consequence of the tangent line approximation. $\square$

Notice that most of the work I am doing here is to prove the result for the derivative. The same was true in the last subsection. In your homework I say you can assume the product and quotient rules for functions so that problem shouldn't be too hard. You just have to pay attention to how I defined the differential and how it is related to the derivative.

6.6 special product rules

In this section I gather together a few results which are commonly needed in applications of calculus.

6.6.1 calculus of paths in $\mathbb{R}^3$

A **path** is a mapping from $\mathbb{R}$ to $\mathbb{R}^m$. We use such mappings to model position, velocity and acceleration of particles in the case $m = 3$. Some of these things were proved in previous sections of this chapter but I intend for this section to be self-contained so that you can read it without digging through the rest of this chapter.

Proposition 6.6.1.

If $F, G : U \subseteq \mathbb{R} \to \mathbb{R}^m$ are differentiable vector-valued functions and $\phi : U \subseteq \mathbb{R} \to \mathbb{R}$ is a differentiable real-valued function then for each $t \in U$,

1. $(F + G)'(t) = F'(t) + G'(t)$.
2. $(cF)'(t) = cF'(t)$.
3. $(\phi F)'(t) = \phi'(t)F(t) + \phi(t)F'(t)$.
4. $(F \cdot G)'(t) = F'(t) \cdot G(t) + F(t) \cdot G'(t)$.
5. provided $m = 3$, $(F \times G)'(t) = F'(t) \times G(t) + F(t) \times G'(t)$.
6. provided $\phi(U) \subset dom(F')$, $(F \circ \phi)'(t) = \phi'(t)F(\phi(t))$.

We have to insist that $m = 3$ for the statement with cross-products since we only have a standard cross-product in $\mathbb{R}^3$. We prepare for the proof of the proposition with a useful lemma. Notice this lemma tells us how to actually calculate the derivative of paths in examples. The derivative of component functions is nothing more than calculus I and one of our goals is to reduce things to those sort of calculations whenever possible.

Lemma 6.6.2.

If $F : U \subseteq \mathbb{R} \to \mathbb{R}^m$ is differentiable vector-valued function then for all $t \in U$,

$$F'(t) = (F_1'(t), F_2'(t), \dots, F_m'(t))$$

We are given that the following vector limit exists and is equal to $F'(t)$,

$$F'(t) = \lim_{h \to 0} \frac{F(t + h) - F(t)}{h}$$

then by Proposition 3.2.10 the limit of a vector is related to the limits of its components as follows:

$$F'(t) \cdot e_j = \lim_{h \to 0} \frac{F_j(t + h) - F_j(t)}{h}.$$

Thus $(F'(t))_j = F_j'(t)$ and the lemma follows[6]. ∇

Proof of proposition: We use the notation $F = \sum F_j e_j = (F_1, \dots, F_m)$ and $G = \sum_i G_i e_i = (G_1, \dots, G_m)$ throughout the proofs below. The $\sum$ is understood to range over $1, 2, \dots m$. Begin with (1.),

$$\begin{aligned}
[(F+G)']_j &= \tfrac{d}{dt}[(F+G)_j] && \text{using the lemma} \\
&= \tfrac{d}{dt}[F_j + G_j] && \text{using def. } (F+G)_j = F_j + G_j \\
&= \tfrac{d}{dt}[F_j] + \tfrac{d}{dt}[G_j] && \text{by calculus I, } (f+g)' = f' + g'. \\
&= [F' + G']_j && \text{def. of vector addition for } F' \text{ and } G'
\end{aligned}$$

Hence $(F \times G)' = F' \times G + F \times G'$.The proofs of 2,3,5 and 6 are similar. I'll prove (5.),

$$\begin{aligned}
[(F \times G)']_k &= \tfrac{d}{dt}[(F \times G)_k] && \text{using the lemma} \\
&= \tfrac{d}{dt}[\textstyle\sum \epsilon_{ijk} F_i G_j] && \text{using def. } F \times G \\
&= \textstyle\sum \epsilon_{ijk} \tfrac{d}{dt}[F_i G_j] && \text{repeatedly using, } (f+g)' = f' + g' \\
&= \textstyle\sum \epsilon_{ijk} [\tfrac{dF_i}{dt} G_j + F_i \tfrac{dG_j}{dt}] && \text{repeatedly using, } (fg)' = f'g + fg' \\
&= \textstyle\sum \epsilon_{ijk} \tfrac{dF_i}{dt} G_j \sum \epsilon_{ijk} F_i \tfrac{dG_j}{dt}] && \text{property of finite sum } \textstyle\sum \\
&= (\tfrac{dF}{dt} \times G)_k + (F \times \tfrac{dG}{dt})_k) && \text{def. of cross product} \\
&= \left(\tfrac{dF}{dt} \times G + F \times \tfrac{dG}{dt} \right)_k && \text{def. of vector addition}
\end{aligned}$$

Notice that the calculus step really just involves calculus I applied to the components. The ordinary product rule was the crucial factor to prove the product rule for cross-products. We'll see the same for the dot product of mappings. Prove (4.)

$$\begin{aligned}
(F \cdot G)'(t) &= \tfrac{d}{dt}[\textstyle\sum F_k G_k] && \text{using def. } F \cdot G \\
&= \textstyle\sum \tfrac{d}{dt}[F_k G_k] && \text{repeatedly using, } (f+g)' = f' + g' \\
&= \textstyle\sum [\tfrac{dF_k}{dt} G_k + F_k \tfrac{dG_k}{dt}] && \text{repeatedly using, } (fg)' = f'g + fg' \\
&= \tfrac{dF}{dt} \cdot G + F \cdot \tfrac{dG}{dt}. && \text{def. of dot product}
\end{aligned}$$

The proof of (3.) follows from applying the product rule to each component of $\phi(t)F(t)$. The proof of (2.) follow from (3.) in the case that $phi(t) = c$ so $\phi'(t) = 0$. Finally the proof of (6.) follows from applying the chain-rule to each component. $\square$

---

[6]this notation I first saw in a text by Marsden, it means the proof is partially completed but you should read on to finish the proof

6.6.2 calculus of matrix-valued functions of a real variable

Definition 6.6.3.

A matrix-valued function of a real variable is a function from $I \subseteq \mathbb{R}$ to $\mathbb{R}^{\,m\times n}$. Suppose $A : I \subseteq \mathbb{R} \rightarrow \mathbb{R}^{\,m\times n}$ is such that $A_{ij} : I \subseteq \mathbb{R} \rightarrow \mathbb{R}$ is differentiable for each i, j then we define

$$\frac{dA}{dt} = \left[\frac{dA_{ij}}{dt}\right]$$

which can also be denoted $(A')_{ij} = A'_{ij}$. We likewise define $\int A dt = [\int A_{ij} dt]$ for A with integrable components. Definite integrals and higher derivatives are also defined component-wise.

Example 6.6.4. *Suppose* $A(t) = \begin{bmatrix} 2t & 3t^2 \\ 4t^3 & 5t^4 \end{bmatrix}$. *I'll calculate a few items just to illustrate the definition above. calculate; to differentiate a matrix we differentiate each component one at a time:*

$$A'(t) = \begin{bmatrix} 2 & 6t \\ 12t^2 & 20t^3 \end{bmatrix} \qquad A''(t) = \begin{bmatrix} 0 & 6 \\ 24t & 60t^2 \end{bmatrix} \qquad A'(0) = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}$$

Integrate by integrating each component:

$$\int A(t)dt = \begin{bmatrix} t^2 + c_1 & t^3 + c_2 \\ t^4 + c_3 & t^5 + c_4 \end{bmatrix} \qquad \int_0^2 A(t)dt = \begin{bmatrix} t^2\big|_0^2 & t^3\big|_0^2 \\ t^4\big|_0^2 & t^5\big|_0^2 \end{bmatrix} = \begin{bmatrix} 4 & 8 \\ 16 & 32 \end{bmatrix}$$

Proposition 6.6.5.

Suppose A, B are matrix-valued functions of a real variable, f is a function of a real variable, c is a constant, and C is a constant matrix then

1. $(AB)' = A'B + AB'$ (product rule for matrices)
2. $(AC)' = A'C$
3. $(CA)' = CA'$
4. $(fA)' = f'A + fA'$
5. $(cA)' = cA'$
6. $(A + B)' = A' + B'$

where each of the functions is evaluated at the same time t and I assume that the functions and matrices are differentiable at that value of t and of course the matrices A, B, C are such that the multiplications are well-defined.

Proof: Suppose $A(t) \in \mathbb{R}^{\ m\times n}$ and $B(t) \in \mathbb{R}^{\ n\times p}$ consider,

$$\begin{aligned}
(AB)'_{ij} &= \tfrac{d}{dt}((AB)_{ij}) && \text{defn. derivative of matrix} \\
&= \tfrac{d}{dt}(\textstyle\sum_k A_{ik}B_{kj}) && \text{defn. of matrix multiplication} \\
&= \textstyle\sum_k \tfrac{d}{dt}(A_{ik}B_{kj}) && \text{linearity of derivative} \\
&= \textstyle\sum_k \left[\tfrac{dA_{ik}}{dt}B_{kj} + A_{ik}\tfrac{dB_{kj}}{dt}\right] && \text{ordinary product rules} \\
&= \textstyle\sum_k \tfrac{dA_{ik}}{dt}B_{kj} + \sum_k A_{ik}\tfrac{dB_{kj}}{dt} && \text{algebra} \\
&= (A'B)_{ij} + (AB')_{ij} && \text{defn. of matrix multiplication} \\
&= (A'B + AB')_{ij} && \text{defn. matrix addition}
\end{aligned}$$

this proves (1.) as i, j were arbitrary in the calculation above. The proof of (2.) and (3.) follow quickly from (1.) since C constant means $C' = 0$. Proof of (4.) is similar to (1.):

$$\begin{aligned}
(fA)'_{ij} &= \tfrac{d}{dt}((fA)_{ij}) && \text{defn. derivative of matrix} \\
&= \tfrac{d}{dt}(fA_{ij}) && \text{defn. of scalar multiplication} \\
&= \tfrac{df}{dt}A_{ij} + f\tfrac{dA_{ij}}{dt} && \text{ordinary product rule} \\
&= (\tfrac{df}{dt}A + f\tfrac{dA}{dt})_{ij} && \text{defn. matrix addition} \\
&= (\tfrac{df}{dt}A + f\tfrac{dA}{dt})_{ij} && \text{defn. scalar multiplication.}
\end{aligned}$$

The proof of (5.) follows from taking $f(t) = c$ which has $f' = 0$. I leave the proof of (6.) as an exercise for the reader. $\square$.

To summarize: the calculus of matrices is the same as the calculus of functions with the small qualifier that we must respect the rules of matrix algebra. The noncommutativity of matrix multiplication is the main distinguishing feature.

6.6.3 calculus of complex-valued functions of a real variable

Differentiation of functions from $\mathbb{R}$ to $\mathbb{C}$ is defined by splitting a given function into its real and imaginary parts then we just differentiate with respect to the real variable one component at a time. For example:

$$\begin{aligned}
\frac{d}{dt}(e^{2t}\cos(t) + ie^{2t}\sin(t)) &= \frac{d}{dt}(e^{2t}\cos(t)) + i\frac{d}{dt}(e^{2t}\sin(t)) \\
&= (2e^{2t}\cos(t) - e^{2t}\sin(t)) + i(2e^{2t}\sin(t) + e^{2t}\cos(t)) \qquad (6.1) \\
&= e^{2t}(2+i)(\cos(t) + i\sin(t)) \\
&= (2+i)e^{(2+i)t}
\end{aligned}$$

where I have made use of the identity[7] $e^{x+iy} = e^x(\cos(y) + i\sin(y))$. We just saw that $\frac{d}{dt}e^{\lambda t} = \lambda e^{\lambda t}$ which seems obvious enough until you appreciate that we just proved it for $\lambda = 2 + i$.

[7] or definition, depending on how you choose to set-up the complex exponential, I take this as the definition in calculus II

Chapter 7

local extrema for multivariate functions

In this chapter I show how the multivariate Taylor series and the theory of quadratic forms give a general form of the second derivative test. In particular we recover the second derivative tests of calculus I and III as special cases. There are technical concerns about remainders and convergence that I set aside for this chapter. The techniques developed here are not entirely general, there are exceptional cases but that is not surprising, we had the same trouble in calculus I. If you read the fine print you'll find we really only have nice theorems for continuously differentiable functions. When functions have holes or finite jump discontinuities we have to treat those separately.

7.1 Taylor series for functions of two variables

Our goal here is to find an analogue for Taylor's Theorem for function from $\mathbb{R}^n$ to $\mathbb{R}$. Recall that if $g : U \subseteq \mathbb{R} \to \mathbb{R}$ is *smooth* at $a \in \mathbb{R}$ then we can compute as many derivatives as we wish, moreover we can generate the Taylor's series for g centered at a:

$$\boxed{g(a+h) = g(a) + g'(a)h + \frac{1}{2}g''(a)h^2 + \frac{1}{3!}g''(a)h^3 + \cdots = \sum_{n=0}^{\infty} \frac{g^{(n)}(a)}{n!}h^n}$$

The equation above assumes that g is analytic at a. In other words, the function actually matches it's Taylor series near a. This concept can be made rigorous by discussing the remainder. If one can show the remainder goes to zero then that proves the function is analytic. (*read p117-127 of Edwards for more on these concepts, I did cover some of that in class this semester, Theorem 6.3 is particularly interesting*).

7.1.1 deriving the two-dimensional Taylor formula

The idea is fairly simple: create a function on $\mathbb{R}$ with which we can apply the ordinary Taylor series result. Much like our discussion of directional derivatives we compose a function of two variables with linear path in the domain. Let $f : U \subseteq \mathbb{R}^2 \to \mathbb{R}$ be 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 differentiate, note we use the chain rule for functions of several variables in what follows:

$$\begin{aligned} g'(t) = (f \circ \phi)'(t) &= f'(\phi(t))\phi'(t) \\ &= \nabla f(\phi(t)) \cdot (h_1, h_2) \\ &= h_1 f_x(a + th_1, b + th_2) + h_2 f_y(a + th_1, b + th_2) \end{aligned}$$

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),

$$\begin{aligned} g''(t) &= h_1 f_x'(a + th_1, b + th_2) + h_2 f_y'(a + th_1, b + th_2) \\ &= h_1 \nabla f_x(\phi(t)) \cdot (h_1, h_2) + h_2 \nabla f_y(\phi(t)) \cdot (h_1, h_2) \\ &= h_1^2 f_{xx} + h_1 h_2 f_{yx} + h_2 h_1 f_{xy} + h_2^2 f_{yy} \\ &= h_1^2 f_{xx} + 2h_1 h_2 f_{xy} + h_2^2 f_{yy} \end{aligned}$$

Thus, making explicit the point dependence, $g''(0) = h_1^2 f_{xx}(a, b) + 2h_1 h_2 f_{xy}(a, b) + h_2^2 f_{yy}(a, b)$. We may construct the Taylor series for g up to quadratic terms:

$$\begin{aligned} g(0 + t) &= g(0) + tg'(0) + \frac{1}{2}g''(0) + \cdots \\ &= f(a, b) + t[h_1 f_x(a, b) + h_2 f_y(a, b)] + \frac{t^2}{2}\left[h_1^2 f_{xx}(a, b) + 2h_1 h_2 f_{xy}(a, b) + h_2^2 f_{yy}(a, b)\right] + \cdots \end{aligned}$$

Note that $g(t) = f(a + th_1, b + th_2)$ hence $g(1) = f(a + h_1, b + h_2)$ and consequently,

$$\begin{aligned} f(a + h_1, b + h_2) = f(a, b) &+ h_1 f_x(a, b) + h_2 f_y(a, b) + \\ &+ \frac{1}{2}\Big[h_1^2 f_{xx}(a, b) + 2h_1 h_2 f_{xy}(a, b) + h_2^2 f_{yy}(a, b)\Big] + \cdots \end{aligned}$$

Omitting point dependence on the 2^{nd} derivatives,

$$\boxed{f(a + h_1, b + h_2) = f(a, b) + h_1 f_x(a, b) + h_2 f_y(a, b) + \tfrac{1}{2}\left[h_1^2 f_{xx} + 2h_1 h_2 f_{xy} + h_2^2 f_{yy}\right] + \cdots}$$

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,

$$\begin{aligned} f(x, y) = f(a, b) &+ (x - a)f_x(a, b) + (y - b)f_y(a, b) + \\ &+ \frac{1}{2}\Big[(x - a)^2 f_{xx}(a, b) + 2(x - a)(y - b)f_{xy}(a, b) + (y - b)^2 f_{yy}(a, b)\Big] + \cdots \end{aligned}$$

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'd find that, using $a = a_1$ and $b = a_2$ as well as $x = x_1$ and $y = x_2$,

$$\boxed{f(x,y) = \sum_{n=0}^{\infty}\sum_{i_1=0}^{n}\sum_{i_2=0}^{n}\cdots\sum_{i_n=0}^{n}\frac{1}{n!}\frac{\partial^{(n)} f(a_1,a_2)}{\partial x_{i_1}\partial x_{i_2}\cdots\partial x_{i_n}}(x_{i_1}-a_{i_1})(x_{i_2}-a_{i_2})\cdots(x_{i_n}-a_{i_n})}$$

Let me expand the third order case just for fun:

$$\sum_{i_1,i_2,i_3=0}^{3} \xrightarrow[\text{terms}]{\text{order three}} \sum_{\substack{i_1+i_2+i_3=3\\ i_1,i_2,i_3=0}}^{3} \frac{1}{3!}\left(\frac{\partial^{(3)} f(a,b)}{\partial x_{i_1}\partial x_{i_2}\partial x_{i_3}}\right)(x_{i_1}-a_{i_1})(x_{i_2}-a_{i_2})(x-a_{i_3}) =$$

$$\begin{aligned}
= \frac{\partial^3 f}{\partial x\partial x\partial x}(a,b)(x-a)^3 &+ \frac{\partial^3 f}{\partial x\partial x\partial y}(x-a)^2(y-b)^2 && + f_{yyx}(x-a)(y-b)^2 \\
&+ f_{xyx}(x-a)^2(y-b)^2 && + f_{yxy}(x-a)(y-b)^2 \\
&+ f_{yxx}(x-a)^2(y-b)^2 && + f_{xyy}(x-a)(y-b)^2 \\
& && + f_{yyy}(y-b)^3
\end{aligned}$$

Thus,

$$\begin{aligned}
f(x,y) = {} & f(a,b) + f_x(x-a) + f_y(y-b) \\
& + \frac{1}{2}\left(f_{xx}(x-a)^2 + 2f_{xy}(x-a)(y-b) + f_{yy}(y-b)^2\right) \\
& + \frac{1}{3!}\left(f_{xxx}(x-a)^3 + 3f_{xxy}(x-a)^2(y-b) + 3f_{xyy}(x-a)(y-b)^2 + f_{yyy}(y-b)^3\right) \\
& + \cdots
\end{aligned}$$

Fortunately we're only really interested in the $n = 0, 1, 2$ order terms. Conceptually, $n = 0$ tells us where to base the tangent plane, $n = 1$ tell us how to build the tangent plane. We will soon discuss how $n = 2$ show us if the tangent plane is at the top or bottom of a hill if we're at a critical point. We pause to play with multivariate series:

Example 7.1.1.

Example: $f(x,y) = \sin(x)\cos(y)$

$$= \left(x - \frac{1}{3!}x^3 + \frac{1}{5!}x^5 + \cdots\right)\left(1 - \frac{1}{2!}y^2 + \frac{1}{4!}y^4 - \frac{1}{6!}y^6 + \cdots\right)$$

$$= \underbrace{x - \frac{1}{2}xy^2 + \frac{1}{4!}xy^4 - \frac{1}{3!}x^3 + \frac{1}{3!2!}x^3y^2 + \frac{1}{5!}x^5}_{\text{terms upto } 5^{\text{th}} \text{ order}} + \cdots$$

$$= x + \frac{1}{3!}\left(\frac{-3!}{2}xy^2 - x^3\right) + \frac{1}{5!}\left(x^5 + \frac{5!}{4!}xy^4 + \frac{5!}{3!2!}x^3y^2\right)$$

$$= f(0,0) + f_x(0,0)x + f_y(0,0)y + \frac{1}{2}\left[f_{xx}(0,0)x^2 + 2f_{xy}(0,0)xy + f_{yy}(0,0)y^2\right]$$

$$+ \frac{1}{3!}\left(f_{xxx}(0,0)x^3 + 3f_{xxy}(0,0)x^2y + 3f_{xyy}(0,0)xy^2 + f_{yyy}(0,0)y^3\right) + \cdots$$

Neat way to calculate partial derivatives of high orders w/o actually calculating them.

Note:

$$f_y(0,0) = 0$$

$$f_{yyy}(0,0) = 0$$

$$f_{xxx}(0,0) = -1$$

$$f_{xyy}(0,0) = -1 \quad \left(\text{since } \frac{3!}{2} = \frac{3\cdot 2}{2} = 3\right)$$

$$f_{xxxyy}(0,0) = \frac{5!}{3!2!5} = \frac{\cancel{5}\cdot 4\cdot \cancel{3}\cdot \cancel{2}\cdot 1}{\cancel{3}\cdot 2\cdot \cancel{2}\cdot \cancel{5}} = 2.$$

- You can use the Cauchy-Product for series to calculate much higher orders w/o doing the whole series multiplication.

Of course many functions of two variables cannot be separated into a product of a function of x and a function of y. In those cases we'd have to calculate the Taylor series directly.

Example 7.1.2.

Example: choose your path carefully.

$$f(x,y) = \sin(x+y)$$

$$= \sum_{n=0}^{\infty} (-1)^n \frac{(x+y)^{2n+1}}{(2n+1)!}$$

$$= x+y - \frac{1}{3!}(x+y)^3 + \frac{1}{5!}(x+y)^5 + \dots$$

Verses

$$f(x,y) = \sin(x+y)$$

$$= \sin x \cos y + \sin y \cos x$$

$$= \left(x - \frac{1}{3!}x^3 + \cdots\right)\left(1 - \frac{1}{2}y^2 + \cdots\right) + \left(y - \frac{1}{3!}y^3 + \cdots\right)\left(1 - \frac{1}{2}x^2 + \cdots\right)$$

which is better? are they the same?

(YES! thank you absolute convergence $\Rightarrow$ rearrangements ok)

generally you can't just shift terms w/o changing

I.O.C. ... there be dragons.

Example 7.1.3.

Example: center $f(x,y) = xy$ about $(1,1)$.

$f(1,1) = 1$

$f_x(1,1) = y|_{(1,1)} = 1$

$f_y(1,1) = x|_{(1,1)} = 1$

$f_{xx}(1,1) = 0$

$f_{yy}(1,1) = 0$

$f_{xy}(1,1) = 1$

Higher derivatives are all dead.

Idea: use Taylor's Th$^{\underline{m}}$ to do sneaky algebra.

$$\boxed{f(x,y) = 1 + 1(x-1) + 1(y-1) + \tfrac{1}{2}(x-1)(y-1)}$$

centered at $(1,1)$.

7.2 Taylor series for functions of many variables

Consider $f : V \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ and denote $x \longmapsto f(x)$ where $x = (x_1, x_2, \dots, x_n)$ then consider point $p = (a_1, a_2, \dots, a_n)$ and construct path $\varphi : \mathbb{R} \to V$ via the rule:

$$\varphi(t) = (a_1 + th_1,\ a_2 + th_2, \dots, a_n + th_n)$$

Note $\varphi(0) = (a_1, \dots, a_n) = p$ and $\varphi'(0) = (h_1, \dots, h_n) = h$. Form the composite $g = f \circ \varphi : \mathbb{R} \to V \longrightarrow \mathbb{R}$ and <u>assume</u> f has continuous partial derivatives of <u>all</u> orders. The Taylor series for g centered at zero has form $g(0+\lambda) = g(0) + \lambda g'(0) + \frac{1}{2}\lambda^2 g''(0) + \dots$ we only intend to calculate upto this order. Focus on $g'(0)$ since $g(0) = f(\varphi(0)) = f(p)$ is easy.

$$\begin{aligned} g'(t) &= \frac{d}{dt}\left[f(x_1, x_2, \dots, x_n)\right] \qquad \text{where } x_j = a_j + th_j \text{ for } j = 1, 2, \dots, n. \\ &= \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(a+th)\frac{d}{dt}(a_j + th_j) \\ &= \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(a+th) \cdot h_j \end{aligned}$$

$$\boxed{g'(0) = \sum_{j=1}^{n} h_j \frac{\partial f}{\partial x_j}(a)}$$

Likewise,

$$\begin{aligned} g''(t) &= \frac{d}{dt}\left(\sum_{j=1}^{n} h_j \frac{\partial f}{\partial x_j}(a+th)\right) \\ &= \sum_{j=1}^{n}\left(h_j \sum_{k=1}^{n} \frac{\partial^2 f}{\partial x_k \partial x_j}(a+th)\frac{d}{dt}(a_k + th_k)\right) \\ &= \underline{\sum_{j,k=1}^{n} h_j h_k \frac{\partial^2 f}{\partial x_k \partial x_j}(a+th)}. \qquad \text{using continuous diff to commute partials!} \end{aligned}$$

Continuing:

$$g''(0) = \sum_{j,k=1}^{n} h_j h_k \frac{\partial^2 f}{\partial x_k \partial x_j}(a)$$

Put it all together,

$$g(0+\lambda) = g(0) + \lambda g'(0) + \tfrac{1}{2}\lambda^2 g''(0) + \cdots$$
$$= f(a) + \lambda \sum_{j=1}^{n} h_j \frac{\partial f}{\partial x_j}(a) + \tfrac{1}{2}\lambda^2 \sum_{j,k=1}^{n} h_j h_k \frac{\partial^2 f}{\partial x_k \partial x_j}(a) + \cdots$$

Finally, $g(0+1) = f(\varphi(1)) = f(a+h) = f(a_1+h_1, a_2+h_2, \dots, a_n+h_n)$

$$\boxed{f(a+h) = f(a) + \sum_{j=1}^{n} h_j \frac{\partial f}{\partial x_j}(a) + \sum_{j,k=1}^{n} \frac{h_j h_k}{2!} \frac{\partial^2 f}{\partial x_k \partial x_j}(a) + \cdots}$$

We can substitute $x = a+h \longrightarrow h = x-a$ which means $h_1 = x_1 - a_1, \dots, h_n = x_n - a_n$ and then

$$\boxed{f(x) = f(a) + \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(a)(x_j - a_j) + \sum_{j,k=1}^{n} \frac{\partial^2 f}{\partial x_k \partial x_j} \frac{1}{2!}(x_k - a_k)(x_j - a_j) + \cdots}$$

Summary: We can approximate $f(x)$ locally by a polynomial in x, y with terms upto order two by stringing together partial derivatives of orders 1 and 2 appropriately. Generally a 2^{nd} order polynomial in $\mathbb{R}[x_1, x_2, \dots, x_n]$ is called a quadratic form on $\mathbb{R}^n$.

7.3 quadratic forms, conic sections and quadric surfaces

Conic sections and quadratic surfaces are common examples in calculus. For example:

$$x^2 + y^2 = 4 \qquad \text{level curve; generally has form } f(x, y) = k$$

$$x^2 + 4y^2 + z^2 = 1 \qquad \text{level surface; generally has form } F(x, y, z) = k$$

Our goal is to see what linear algebra and multivariate calculus have to say about conic sections and quadric surfaces. (these notes borrowed from my linear algebra notes)

7.3.1 quadratic forms and their matrix

We are primarily interested in the application of this discussion to $\mathbb{R}^2$ and $\mathbb{R}^3$, however, these concepts equally well apply to arbitrarily high finite dimensional problems where the geometry is not easily pictured.

Definition 7.3.1.

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 $\vec{x} = [x, y]^T$ and $A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$ then

$$\vec{x}^T A\vec{x} = ax^2 + bxy + byx + cy^2 = ax^2 + 2bxy + y^2.$$

The $n = 3$ case is similar,denote $A = [A_{ij}]$ and $\vec{x} = [x, y, z]^T$ so that

$$\vec{x}^T A\vec{x} = A_{11}x^2 + 2A_{12}xy + 2A_{13}xz + A_{22}y^2 + 2A_{23}yz + A_{33}z^2.$$

Generally, if $[A_{ij}] \in \mathbb{R}^{\ n\times n}$ and $\vec{x} = [x_i]^T$ then the quadratic form

$$\vec{x}^T A\vec{x} = \sum_{i,j} A_{ij}x_ix_j = \sum_{i=1}^{n} A_{ii}x_i^2 + \sum_{i<j} 2A_{ij}x_ix_j.$$

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.

You should notice can write a given quadratic form in terms of a dot-product:

$$\vec{x}^T A\vec{x} = \vec{x} \cdot (A\vec{x}) = (A\vec{x}) \cdot \vec{x} = \vec{x}^T A^T\vec{x}$$

Some texts actually use the middle equality above to define a symmetric matrix.

Example 7.3.2.

$$2x^2 + 2xy + 2y^2 = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

Example 7.3.3.

$$2x^2 + 2xy + 3xz - 2y^2 - z^2 = \begin{bmatrix} x & y & z \end{bmatrix} \begin{bmatrix} 2 & 1 & 3/2 \\ 1 & -2 & 0 \\ 3/2 & 0 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

Proposition 7.3.4.

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 7.3.5.

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.

7.3.2 almost an introduction to eigenvectors

Eigenvectors and eigenvalues play an important role in theory and application. In particular, eigenvalues and eigenvectors allow us to (if possible) *diagonalize* a matrix. This essentially is the problem of choosing coordinates for a particular system which most clearly reveals the true nature of the system. For example, the fact that $2xy = 1$ is a hyperbola is clearly seen once we change to coordinates whose axes point along the eigenvectors for the quadratic form $Q(x, y) = 2xy$.

Likewise, in the study of rotating rigid bodies the eigenvectors of the inertia tensor give the so-called principle axes of inertia. When a body is set to spin about such an axes through its center of mass the motion is natural, smooth and does not wobble. The inertia tensor gives a quadratic form in the angular velocity which represents the rotational kinetic energy. I've probably assigned a homework problem so you can understand this paragraph. In any event, there are many motivations for studying eigenvalues and vectors. I explain much more theory for e-vectors in the linear course.

Definition 7.3.6.

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 7.3.7.

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. $\square$

Example 7.3.8. *Let* $A = \left[\begin{array}{cc} 3 & 1 \\ 3 & 1 \end{array}\right]$ *find the e-values and e-vectors of* A.

$$det(A - \lambda I) = det \left[\begin{array}{cc} 3-\lambda & 1 \\ 3 & 1-\lambda \end{array}\right] = (3-\lambda)(1-\lambda) - 3 = \lambda^2 - 4\lambda = \lambda(\lambda - 4) = 0$$

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,*

$$(A - 0I)u_1 = \left[\begin{array}{cc} 3 & 1 \\ 3 & 1 \end{array}\right]\left[\begin{array}{c} u \\ v \end{array}\right] = \left[\begin{array}{c} 3u+v \\ 3u+v \end{array}\right] = \left[\begin{array}{c} 0 \\ 0 \end{array}\right]$$

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,* $u_1 = \left[\begin{array}{c} u \\ -3u \end{array}\right] = u\left[\begin{array}{c} 1 \\ -3 \end{array}\right]$. *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 7.3.9.

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[1] (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 7.3.10. *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 $\begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

7.3.3 quadratic form examples

Example 7.3.11. *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* $Q(v) = ||v||^2$

[1]think about it, there is a 1-1 correspondance between symmetric matrices and quadratic forms

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.*

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 7.3.12. *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.*

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 7.3.13. *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.*
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 7.3.14. *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.*
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* $\lambda_3 = 3$.

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.

Proposition 7.3.15.

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.

Example 7.3.16. *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[2]:

$$\tilde{Q}(\bar{x}, \bar{y}) = \bar{x}^2 + 3\bar{y}^2$$

[2] technically $\tilde{Q}(\bar{x}, \bar{y})$ is $Q(x(\bar{x}, \bar{y}), y(\bar{x}, \bar{y}))$

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.

Example 7.3.17. *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

$$P = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \quad \Rightarrow \quad \begin{matrix} x & = \frac{1}{2}(\bar{x}+\bar{y}) \\ y & = \frac{1}{2}(-\bar{x}+\bar{y}) \end{matrix} \quad or \quad \begin{matrix} \bar{x} & = \frac{1}{2}(x-y) \\ \bar{y} & = \frac{1}{2}(x+y) \end{matrix}$$

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 $y = \sqrt{k} - x$.

Example 7.3.18. *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 $Q(x,y) = [x,y]\begin{bmatrix} 0 & 2 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$. Denote the matrix of the form by A and calculate the e-values/vectors:*

$$det(A - \lambda I) = det\begin{bmatrix} -\lambda & 2 \\ 2 & -\lambda \end{bmatrix} = \lambda^2 - 4 = (\lambda+2)(\lambda-2) = 0$$

Therefore, the e-values are $\lambda_1 = -2$ and $\lambda_2 = 2$.

$$(A+2I)\vec{u}_1 = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \Rightarrow \quad \vec{u}_1 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

I just solved $u+v=0$ to give $v=-u$ choose $u=1$ then normalize to get the vector above. Next,

$$(A-2I)\vec{u}_2 = \begin{bmatrix} -2 & 2 \\ 2 & -2 \end{bmatrix}\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \Rightarrow \quad \vec{u}_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

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:

$$\tilde{Q}(\bar{x},\bar{y}) = -2\bar{x}^2 + 2\bar{y}^2$$

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.

Remark 7.3.19.

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. [3]

We ought to do at least one 3-dimensional example.

Example 7.3.20. *Consider the quadric form 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*

$$P = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & \sqrt{2} \end{bmatrix} \quad \Rightarrow \quad \begin{array}{rcl} x & = & \frac{1}{2}(\bar{x} + \bar{y}) \\ y & = & \frac{1}{2}(-\bar{x} + \bar{y}) \\ z & = & \bar{z} \end{array} \quad \text{or} \quad \begin{array}{rcl} \bar{x} & = & \frac{1}{2}(x - y) \\ \bar{y} & = & \frac{1}{2}(x + y) \\ \bar{z} & = & z \end{array}$$

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* $z - axis$*.*

Remark 7.3.21.

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 $Q(h, k) = \frac{1}{2}h^2 f_{xx}(a, b) + hk f_{xy}(a, b) + \frac{1}{2}h^2 f_{yy}(a, b) = [h, k][Q](h, k)$. 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.

7.4 local extrema from eigenvalues and quadratic forms

We have all the tools we need, let's put them to use now.

Example 7.4.1.

$f(x,y) = x^2 - 2xy + y^2$

$\nabla f = \langle 2x - 2y,\ 2y - 2x \rangle \quad \therefore \quad \nabla f = 0 \Rightarrow \underline{y = x}.$

Infinitely many critical points, have the form (a,a).

$f_{xx}(a,a) = 2 \quad , \quad f_{xy}(a,a) = -2, \quad f_{yy}(a,a) = 2$

Hence, expanding about (a,a),

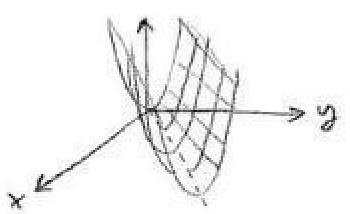
$$\begin{aligned} f(x,y) &= f(a,a) + Q_{(a,a)}(x,y) \\ &= a^2 - 2a^2 + a^2 + \tfrac{1}{2}\left(2(x-a)^2 - 4(x-a)(y-a) + 2(y-a)^2\right) \\ &= \underline{(x-a)^2 - 2(x-a)(y-a) + (y-a)^2}. \end{aligned}$$

Note that

$$[Q_{(a,a)}] = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$$

$$\begin{aligned} \det\left[[Q_{(a,a)}] - \lambda I \right] &= \det\begin{bmatrix} 1-\lambda & -1 \\ -1 & 1-\lambda \end{bmatrix} \\ &= (\lambda - 1)^2 - 1 \\ &= \lambda^2 - 2\lambda \\ &= \lambda(\lambda - 2) \quad \therefore \quad \underline{\lambda_1 = 0,\ \lambda_2 = 2}. \end{aligned}$$

This is a semi-definite form. Each point along $y = x$ gives local min. Note that $f(x,y) = (x-y)^2$ thus $f(x,y) \geq 0$ hence $f(a,a) = 0$ is the global min. for f as well.



(Zero e-values for Q indicate this type of behaviour.)

Example 7.4.2.

$$f(x,y) = \exp(-x^2-y^2)$$
$$f_x(x,y) = -2x\exp(-x^2-y^2)$$
$$f_y(x,y) = -2y\exp(-x^2-y^2)$$
$$f_{xx}(x,y) = -2\exp(-x^2-y^2) + 4x^2\exp(-x^2-y^2)$$
$$f_{xy}(x,y) = 4xy\exp(-x^2-y^2)$$
$$f_{yy}(x,y) = (4y^2-2)\exp(-x^2-y^2)$$

Note $\nabla f(x,y) = 0 = \langle -2x, -2y\rangle e^{-x^2-y^2} \Rightarrow \underline{x = y = 0}$

Only critical pt. is $(0,0)$. We find

$$\begin{aligned} Q(x,y) &= \tfrac{1}{2}f_{xx}(0,0)x^2 + f_{xy}(0,0)xy + \tfrac{1}{2}f_{yy}(0,0)y^2 \\ &= \tfrac{1}{2}(-2)x^2 + 0 + \tfrac{1}{2}(-2)y^2 \\ &= -x^2-y^2 \end{aligned}$$

Thus, we find $(0,0)$ gives local max since Q is negative definite.

$$\boxed{f(x,y) = f(0,0) + Q(x,y) = 1 - x^2 - y^2 + \cdots}$$

Note,

$$[Q] = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \quad \text{thus } \lambda_1 = \lambda_2 = -1 < 0$$

$\therefore$ <u>negative definite</u>

$\therefore$ <u>local maximum</u>.

Remark: this is actually a <u>global</u> max. since $-x^2-y^2 \le 0$ and $\exp((-\infty,0]) = (0,1]$.

Example 7.4.3.

Example: Let $f(x,y) = 2x^2 - xy - 3y^2 - 3x + 7y$ find all critical points and analyze those pts. Find local extrema.

$$\nabla f = \langle 4x - y - 3, -x - 6y + 7 \rangle = \langle 0, 0 \rangle$$

$$4x - y - 3 = 0 \longrightarrow y = 4x - 3$$
$$-x - 6y + 7 = 0 \longrightarrow -x - 6(4x-3) + 7 = 0$$
$$-25x + 18 + 7 = 0$$
$$\Rightarrow -25x = -25$$
$$\Rightarrow \underline{x = 1} \Rightarrow \underline{y = 4 - 3 = 1}.$$

We find a single critical point; $(1,1)$.

$$f_x(x,y) = 4x - y - 3 \quad : \text{note} \quad f_x(1,1) = 4 - 1 - 3 = 0$$
$$f_y(x,y) = -x - 6y + 7 \quad : \text{and} \quad f_y(1,1) = -1 - 6 + 7 = 0$$
$$f_{xx}(x,y) = 4$$
$$f_{xy}(x,y) = -1$$
$$f_{yy}(x,y) = -6$$

good check on earlier calculation.

We find $Q(x,y) = 2(x-1)^2 - 2(x-1)(y-1) - 3(y-1)^2$ hence,

$$f(x,y) = f(1,1) + Q(x,y) \qquad f(1,1) = 2 - 1 - 3 - 3 + 7 = 2$$

$$\therefore \boxed{f(x,y) = 2 + 2(x-1)^2 - 2(x-1)(y-1) - 3(y-1)^2}$$

Note this is not an approximation since higher terms all vanish. It seems likely this is a saddle point, but we need not guess. Note the matrix of Q is

$$[Q] = \begin{bmatrix} 2 & -1 \\ -1 & -3 \end{bmatrix}$$

$$0 = \det([Q] - \lambda I) = \det\begin{pmatrix} 2-\lambda & -1 \\ -1 & -3-\lambda \end{pmatrix}$$
$$= (\lambda+3)(\lambda-2) - 1$$
$$= \lambda^2 + \lambda - 6$$
$$= (\lambda+3)(\lambda-2) \quad \therefore \underline{\lambda_1 = -3 \;\; \& \;\; \lambda_2 = 2}.$$

Thus $f(1,1)$ is neither max nor min near $(1,1)$. It is at the saddle point $(1,1,2)$.

Example 7.4.4.

Example: $f(x,y) = \sin x \cosh y$

$\nabla f = \langle \cos x \cosh y, \sin x \sinh y \rangle$

Thus critical pts must have

① $\cos(x)\cosh y = 0$

② $\sin x \sinh y = 0$

Note $\cosh y = \frac{1}{2}(e^y + e^{-y}) \neq 0$ for all $y \in \mathbb{R}$ hence we need $\cos(x) = 0$. But then $\cos(x) = 0 \Rightarrow \sin x \neq 0$ hence we also need $\sinh y = \frac{1}{2}(e^y - e^{-y}) = 0$. Note

$$\sinh y = \frac{1}{2}(e^y - e^{-y}) = 0 \Rightarrow e^y = e^{-y} \Rightarrow y = -y \Rightarrow \underline{y = 0}.$$

We find a whole family of critical points. Namely

$(x,y) \in \mathbb{R}^2$ such that $\cos(x) = 0$ AND $y = 0$

That gives critical pts. $= \{(n\pi + \frac{\pi}{2}, 0) \mid n \in \mathbb{Z}\}$.

$f_{xx}(x,y) = -\sin x \cosh y$

$f_{xy}(x,y) = \cos x \sinh y$

$f_{yy}(x,y) = \sin x \cosh y$

Thus $f_{xx}(x,0) = -\sin x$, $f_{xy}(x,0) = 0$ and $f_{yy}(x,0) = \sin x$.

If $x = 2k\pi + \frac{\pi}{2}$ then $\sin(x) = 1$ whereas if $x = (2k+1)\pi + \frac{\pi}{2}$ then $\sin(x) = -1$ for all $k \in \mathbb{Z}$.

To summarize, $f_{xx}(n\pi + \frac{\pi}{2}, 0) = (-1)^n$ and $f_{yy}(n\pi + \frac{\pi}{2}, 0) = (-1)^{n+1}$.

We can approximate $f(x,y) = \sin x \cosh y$ by

$$\boxed{f(x,y) \cong (-1)^n + \frac{1}{2}(-1)^n\left(x - n\pi - \frac{\pi}{2}\right)^2 - \frac{1}{2}(-1)^n y^2}$$

This is again a saddle-type for each $n \in \mathbb{Z}$.

$$[Q] = \begin{bmatrix} \frac{1}{2}(-1)^n & 0 \\ 0 & -\frac{1}{2}(-1)^n \end{bmatrix} \quad \therefore \quad \underline{\lambda_1 = \pm \frac{1}{2} \quad \lambda_2 = \mp \frac{1}{2}}$$

Chapter 8

on manifolds and multipliers

In this chapter we show the application of the most difficult results in this course, namely the implicit and inverse mapping theorems. Our first application is in the construction of manifolds as graphs or level sets. Then once we have a convenient concept of a manifold we discuss the idea of Lagrange multipliers. The heart of the method combines orthogonal complements from linear algebra along side the construction of tangent spaces in this course. Hopefully this chapter will help you understand why the implicit and inverse mapping theorems are so useful and also why we need manifolds to make sense of our problems. The patching definition for a manifold is not of much use in this chapter although we will mention how it connects to the other two formulations of a manifold in $\mathbb{R}^m$ in the context of a special case.

8.1 surfaces in $\mathbb{R}^3$

Manifolds or surfaces play a role similar to functions in this course. Our goal is not the study of manifolds alone but it's hard to give a complete account of differentiation unless we have some idea of what is a tangent plane. This subsection does break from the larger pattern of thought in this chapter. I include it here to try to remind how surfaces and tangent planes are described in $\mathbb{R}^3$. We need some amount of generalization beyond this section because the solution of max/min problems with constraints will take us into higher dimensional surfaces even for problems that only involve two or three spatial dimensions. We treat those questions in the next chapter.

There are three main methods to describe surfaces:

1. As a **graph**: $S = \{(x, y, z) \mid z = f(x, y) \text{ where } (x, y) \in dom(f)\}$.
2. As a **level surface**: $S = \{(x, y, z) \mid F(x, y, z) = 0\}$
3. As a **parametrized surface**: $S = \{X(u, v) \mid (u, v) \in dom(X)\}$

Let me remind you we found the tangent plane at $(x_o, y_o, z_o) \in S$ for each of these formalisms as follows (continuing to use the same notation as above):

1. For the **graph**: $z = z_o + f(x_o, y_o) + f_x(x_o, y_o)(x - x_o) + f_y(x_o, y_o)(y - y_o)$.
2. For the **level surface**: plane through (x_o, y_o, z_o) with normal $(\nabla F)(x_o, y_o, z_o)$
3. For the **parametrized surface**: find (u_o, v_o) with $X(u_o, v_o) = (x_o, y_o, z_o)$, the tangent plane goes through $X(u_o, v_o)$ and has normal $N(u_o, v_o) = X_u(u_o, v_o) \times X_v(u_o, v_o)$.

Perhaps you recall that the normal vector field to the surface S was important in the formulation of surface integrals to calculate the flux of vector fields.

Example 8.1.1. *The plane through the point $\vec{r}_o$ with normal $\vec{n} =< a, b, c >$ can be described as:*

1. *all $\vec{r} \in \mathbb{R}^3$ such that $(\vec{r} - \vec{r}_o) \cdot \vec{n} = 0$.*
2. *all $(x, y, z) \in \mathbb{R}^3$ such that $a(x - x_o) + b(y - y_o) + c(z - z_o) = 0$*
3. *if $c \neq 0$, the graph $z = f_3(x, y)$ where $f_3(x, y) = z_o + \frac{a}{c}(x - x_o) + \frac{b}{c}(y - y_o)$*
4. *if $b \neq 0$, the graph $y = f_3(x, z)$ where $f_2(x, z) = y_o + \frac{a}{b}(x - x_o) + \frac{c}{b}(z - z_o)$*
5. *if $a \neq 0$, the graph $x = f_1(y, z)$ where $f_1(y, z) = x_o + \frac{b}{a}(y - y_o) + \frac{c}{a}(z - z_o)$*
6. *given any two linearly independent vectors $\vec{a}, \vec{b}$ in the plane, the plane is the image of the mapping $X : \mathbb{R}^2 \to \mathbb{R}^3$ defined by $X(u, v) = \vec{r}_o + u\vec{a} + v\vec{b}$*

Example 8.1.2. *The sphere of radius R centered about the origin can be described as:*

1. *all $(x, y, z) \in \mathbb{R}^3$ such that $F(x, y, z) = x^2 + y^2 + z^2 = R^2$*
2. *the graphs of $z = f_\pm(x, y)$ where $f_\pm(x, y) = \pm\sqrt{R^2 - x^2 - y^2}$*
3. *for $(u, v) \in [0, 2\pi] \times [0, \pi]$, $X(u, v) = (R\cos u \sin v, R\sin u \sin v, R\cos v)$*

You may recall that the level surface concept allowed by far the easiest computation of the normal of the tangent plane for a particular point. For example, $\nabla F =< 2x, 2y, 2z >$ in the preceding example. Contrast that to calculation of $X_u \times X_v$ where the $\times$ denotes the dreaded cross-product. Of course each formalism has its place in calculus III.

Remark 8.1.3.

In this warm-up section we have hopefully observed this much about surfaces in $\mathbb{R}^3$:

1. the tangent plane is always 2-dimensional, it is really a plane in the traditional sense of the term.

2. the normal to the tangent plane is always 1-dimensional, the normal through a particular point on the surface is just a line which is orthogonal to all possible tangents through the point.

3. the dimension of the tangent plane and normal give the total dimension of the ambient space; $2 + 1 = 3$.

8.2 manifolds as level sets

We will focus almost exclusively on the level surface formulation of a manifold in the remainder of this chapter. We say $M \subseteq \mathbb{R}^n$ is a **manifold** of dimension $p \leq n$ if M has a p-dimensional tangent plane for each point on M. In other words, M is a p-dimensional manifold if it can be locally approximated by $\mathbb{R}^p$ at each point on M. Moreover, the set of all vectors normal to the tangent space will be $n - p$ dimensional.

These are general concepts which encompasses lines, planes volumes and much much more. Let me illustrate by example:

Example 8.2.1. *Let $g : \mathbb{R}^2 \to \mathbb{R}$ be defined by $g(x, y) = y - x - 1$ note that $g(x, y) = 0$ gives the line $y - x - 1 = 0$ commonly written as $y = x + 1$; note that the line has direction vector $< -1, 1 >$. Furthermore, $\nabla g =< 1, -1 >$ which is orthogonal to $< -1, 1 >$.*

Example 8.2.2. *Let $g : \mathbb{R}^3 \to \mathbb{R}$ be defined by $g(x, y, z) = y - x - 1$ note that $g(x, y, z) = 0$ gives the plane $y - x - 1 = 0$. Furthermore, $\nabla g =< 1, -1, 0 >$ which gives the normal to the plane $g = 0$.*

Example 8.2.3. *Let $g : \mathbb{R}^4 \to \mathbb{R}$ be defined by $g(x, y, z, t) = y - x - 1$ note that $g(x, y, z, t) = 0$ gives the hyperplane $y - x - 1 = 0$. Furthermore, $\nabla g =< 1, -1, 0, 0 >$ which gives the normal to the hyperplane $g = 0$. What does that mean? It means that if I take any vector in the hyperplane it is orthogonal to $< 1, -1, 0, 0 >$. Let $\vec{r}_1, \vec{r}_2$ be points in the solution set of $g(x, y, z, t) = 0$. Denote $\vec{r}_1 = (x_1, y_1, z_1, t_1)$ and $\vec{r}_1 = (x_2, y_2, z_2, t_2)$, we have $y_1 = x_1 + 1$ and $y_2 = x_2 + 1$. The vector in the hyperplane is found from the difference of these points:*

$$\vec{v} = \vec{r}_2 - \vec{r}_1 = (x_2, x_2 + 1, z_2, t_2) - (x_1, x_1 + 1, z_1, t_1) = (x_2 - x_1, x_2 - x_1, z_2 - z_1, t_2 - t_1).$$

It's easy to see that $\vec{v} \cdot \nabla g = 0$ hence ∇g is perpendicular to an arbitrary vector in the hyperplane

If you've begun to develop an intuition for the story we're telling this last example ought to bug you a bit. Why is the difference of points a tangent vector? What happened to the set of all tangent vectors pasted together or the differential or the column space of the derivative? All those concepts still apply but since we were looking at a linear space the space itself matched the tangent hyperplane. The point of the triple of examples above is just to constrast the nature of the equation $g = 0$ in various contexts. We find the dimension of the ambient space changes the dimension of the level set. Basically, we have one equation $g = 0$ and n-unknowns then the inverse image of zero gives us a $(n-1)$-dimensional manifold. If we wanted to obtain a $n-2$ dimensional manifold then we would need two equations which were independent. Before we get to that perhaps I should give a curvy example.

Example 8.2.4. *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 $\nabla g(2, 2, 1, 0)$.

One more example before we dive into the theory of Lagrange multipliers. (which is little more than this section applied to word problems plus the powerful orthogonal complement theorem from linear algebra)

Example 8.2.5. *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, $rref \begin{bmatrix} 2 & 2 & 1 & 0 \\ 0 & 2 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & -1/2 \\ 0 & 1 & 1/2 & 1/2 \end{bmatrix}$. 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

$$< a, b, c, d >=< d/2, -c/2 - d/2, c, d >= \tfrac{c}{2} < 0, -1, 2, 0 > + \tfrac{d}{2} < 1, -1, 0, 2 >$$

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.*

Remark 8.2.6.

The reason I am bothering with these seemingly bizarre examples is that the method of Lagrange multipliers comes down to the observation that both the constraint and objective function's gradient vectors should be normal to the tangent plane of the constraint surface. This means they must both reside in the normal to the tangent plane and hence they will either be colinear or for several constraints they will be linearly dependent. The geometry we consider here justifies the method. Linear algebra supplies the harder part which is that if two vectors are both orthogonal to the tangent plane then they must both be in the orthogonal complement to the tangent plane. The heart of the method of Lagrange multipliers is the orthogonal complement theory from linear algebra. Of course, you can be heartless and still sucessfully apply the method of Lagrange.

8.3 Lagrange multiplier method for one constraint

8.4 Lagrange multiplier method for several constraints

Chapter 8

on manifolds and multipliers

In this chapter we show the application of the most difficult results in this course, namely the implicit and inverse mapping theorems. Our first application is in the construction of manifolds as graphs or level sets. Then once we have a convenient concept of a manifold we discuss the idea of Lagrange multipliers. The heart of the method combines orthogonal complements from linear algebra along side the construction of tangent spaces in this course. Hopefully this chapter will help you understand why the implicit and inverse mapping theorems are so useful and also why we need manifolds to make sense of our problems. The patching definition for a manifold is not of much use in this chapter although we will mention how it connects to the other two formulations of a manifold in $\mathbb{R}^m$ in the context of a special case.

8.1 surfaces in $\mathbb{R}^3$

Manifolds or surfaces play a role similar to functions in this course. Our goal is not the study of manifolds alone but it's hard to give a complete account of differentiation unless we have some idea of what is a tangent plane. This subsection does break from the larger pattern of thought in this chapter. I include it here to try to remind how surfaces and tangent planes are described in $\mathbb{R}^3$. We need some amount of generalization beyond this section because the solution of max/min problems with constraints will take us into higher dimensional surfaces even for problems that only involve two or three spatial dimensions. We treat those questions in the next chapter.

There are three main methods to describe surfaces:

1. As a **graph**: $S = \{(x, y, z) \mid z = f(x, y) \text{ where } (x, y) \in dom(f)\}$.
2. As a **level surface**: $S = \{(x, y, z) \mid F(x, y, z) = 0\}$
3. As a **parametrized surface**: $S = \{X(u, v) \mid (u, v) \in dom(X)\}$

Let me remind you we found the tangent plane at $(x_o, y_o, z_o) \in S$ for each of these formalisms as follows (continuing to use the same notation as above):

1. For the **graph**: $z = z_o + f(x_o, y_o) + f_x(x_o, y_o)(x - x_o) + f_y(x_o, y_o)(y - y_o)$.
2. For the **level surface**: plane through (x_o, y_o, z_o) with normal $(\nabla F)(x_o, y_o, z_o)$
3. For the **parametrized surface**: find (u_o, v_o) with $X(u_o, v_o) = (x_o, y_o, z_o)$, the tangent plane goes through $X(u_o, v_o)$ and has normal $N(u_o, v_o) = X_u(u_o, v_o) \times X_v(u_o, v_o)$.

Perhaps you recall that the normal vector field to the surface S was important in the formulation of surface integrals to calculate the flux of vector fields.

Example 8.1.1. *The plane through the point $\vec{r}_o$ with normal $\vec{n} =< a, b, c >$ can be described as:*

1. *all $\vec{r} \in \mathbb{R}^3$ such that $(\vec{r} - \vec{r}_o) \cdot \vec{n} = 0$.*
2. *all $(x, y, z) \in \mathbb{R}^3$ such that $a(x - x_o) + b(y - y_o) + c(z - z_o) = 0$*
3. *if $c \neq 0$, the graph $z = f_3(x, y)$ where $f_3(x, y) = z_o + \frac{a}{c}(x - x_o) + \frac{b}{c}(y - y_o)$*
4. *if $b \neq 0$, the graph $y = f_3(x, z)$ where $f_2(x, z) = y_o + \frac{a}{b}(x - x_o) + \frac{c}{b}(z - z_o)$*
5. *if $a \neq 0$, the graph $x = f_1(y, z)$ where $f_1(y, z) = x_o + \frac{b}{a}(y - y_o) + \frac{c}{a}(z - z_o)$*
6. *given any two linearly independent vectors $\vec{a}, \vec{b}$ in the plane, the plane is the image of the mapping $X : \mathbb{R}^2 \to \mathbb{R}^3$ defined by $X(u, v) = \vec{r}_o + u\vec{a} + v\vec{b}$*

Example 8.1.2. *The sphere of radius R centered about the origin can be described as:*

1. *all $(x, y, z) \in \mathbb{R}^3$ such that $F(x, y, z) = x^2 + y^2 + z^2 = R^2$*
2. *the graphs of $z = f_\pm(x, y)$ where $f_\pm(x, y) = \pm\sqrt{R^2 - x^2 - y^2}$*
3. *for $(u, v) \in [0, 2\pi] \times [0, \pi]$, $X(u, v) = (R\cos u \sin v, R \sin u \sin v, R\cos v)$*

You may recall that the level surface concept allowed by far the easiest computation of the normal of the tangent plane for a particular point. For example, $\nabla F =< 2x, 2y, 2z >$ in the preceding example. Contrast that to calculation of $X_u \times X_v$ where the $\times$ denotes the dreaded cross-product. Of course each formalism has its place in calculus III.

Remark 8.1.3.

In this warm-up section we have hopefully observed this much about surfaces in $\mathbb{R}^3$:

1. the tangent plane is always 2-dimensional, it is really a plane in the traditional sense of the term.
2. the normal to the tangent plane is always 1-dimensional, the normal through a particular point on the surface is just a line which is orthogonal to all possible tangents through the point.
3. the dimension of the tangent plane and normal give the total dimension of the ambient space; $2 + 1 = 3$.

8.2 manifolds as level sets

We will focus almost exclusively on the level surface formulation of a manifold in the remainder of this chapter. We say $M \subseteq \mathbb{R}^n$ is a **manifold** of dimension $p \leq n$ if M has a p-dimensional tangent plane for each point on M. In other words, M is a p-dimensional manifold if it can be locally approximated by $\mathbb{R}^p$ at each point on M. Moreover, the set of all vectors normal to the tangent space will be $n - p$ dimensional.

These are general concepts which encompasses lines, planes volumes and much much more. Let me illustrate by example:

Example 8.2.1. *Let $g : \mathbb{R}^2 \to \mathbb{R}$ be defined by $g(x, y) = y - x - 1$ note that $g(x, y) = 0$ gives the line $y - x - 1 = 0$ commonly written as $y = x + 1$; note that the line has direction vector $< -1, 1 >$. Furthermore, $\nabla g =< 1, -1 >$ which is orthogonal to $< -1, 1 >$.*

Example 8.2.2. *Let $g : \mathbb{R}^3 \to \mathbb{R}$ be defined by $g(x, y, z) = y - x - 1$ note that $g(x, y, z) = 0$ gives the plane $y - x - 1 = 0$. Furthermore, $\nabla g =< 1, -1, 0 >$ which gives the normal to the plane $g = 0$.*

Example 8.2.3. *Let $g : \mathbb{R}^4 \to \mathbb{R}$ be defined by $g(x, y, z, t) = y - x - 1$ note that $g(x, y, z, t) = 0$ gives the hyperplane $y - x - 1 = 0$. Furthermore, $\nabla g =< 1, -1, 0, 0 >$ which gives the normal to the hyperplane $g = 0$. What does that mean? It means that if I take any vector in the hyperplane it is orthogonal to $< 1, -1, 0, 0 >$. Let $\vec{r}_1, \vec{r}_2$ be points in the solution set of $g(x, y, z, t) = 0$. Denote $\vec{r}_1 = (x_1, y_1, z_1, t_1)$ and $\vec{r}_1 = (x_2, y_2, z_2, t_2)$, we have $y_1 = x_1 + 1$ and $y_2 = x_2 + 1$. The vector in the hyperplane is found from the difference of these points:*

$$\vec{v} = \vec{r}_2 - \vec{r}_1 = (x_2, x_2 + 1, z_2, t_2) - (x_1, x_1 + 1, z_1, t_1) = (x_2 - x_1, x_2 - x_1, z_2 - z_1, t_2 - t_1).$$

It's easy to see that $\vec{v} \cdot \nabla g = 0$ hence ∇g is perpendicular to an arbitrary vector in the hyperplane

If you've begun to develop an intuition for the story we're telling this last example ought to bug you a bit. Why is the difference of points a tangent vector? What happened to the set of all tangent vectors pasted together or the differential or the column space of the derivative? All those concepts still apply but since we were looking at a linear space the space itself matched the tangent hyperplane. The point of the triple of examples above is just to constrast the nature of the equation $g = 0$ in various contexts. We find the dimension of the ambient space changes the dimension of the level set. Basically, we have one equation $g = 0$ and n-unknowns then the inverse image of zero gives us a $(n-1)$-dimensional manifold. If we wanted to obtain a $n-2$ dimensional manifold then we would need two equations which were independent. Before we get to that perhaps I should give a curvy example.

Example 8.2.4. *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 $s = 0$,*

$$\begin{aligned} \frac{d}{ds}\big(\ g \circ \gamma(s)\ \big) = \big(\nabla g\big)(\gamma(s)) \cdot \gamma'(s) = 0 \quad &\Rightarrow \quad \nabla g(2, 2, 1, 0) \cdot < a, b, c, d >= 0 \\ &\Rightarrow \quad < 4, 4, -4, 1 > \cdot < a, b, c, d >= 0 \\ &\Rightarrow \quad 4a + 4b - 4c + d = 0 \end{aligned}$$

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 $\nabla g(2, 2, 1, 0)$.

One more example before we dive into the theory of Lagrange multipliers. (which is little more than this section applied to word problems plus the powerful orthogonal complement theorem from linear algebra)

Example 8.2.5. *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:

$$(G_1 \circ \gamma)'(0) = \nabla G_1(\gamma(0)) \cdot < a, b, c, d >= 0 \quad \Rightarrow \quad < 2, 2, 1, 0 > \cdot < a, b, c, d >= 0$$
$$(G_2 \circ \gamma)'(0) = \nabla G_2(\gamma(0)) \cdot < a, b, c, d >= 0 \quad \Rightarrow \quad < 0, 2, 1, 1 > \cdot < a, b, c, d >= 0$$

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, $rref \begin{bmatrix} 2 & 2 & 1 & 0 \\ 0 & 2 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & -1/2 \\ 0 & 1 & 1/2 & 1/2 \end{bmatrix}$. 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

$$< a, b, c, d >=< d/2, -c/2 - d/2, c, d >= \tfrac{c}{2} < 0, -1, 2, 0 > + \tfrac{d}{2} < 1, -1, 0, 2 >$$

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.*

Remark 8.2.6.

The reason I am bothering with these seemingly bizarre examples is that the method of Lagrange multipliers comes down to the observation that both the constraint and objective function's gradient vectors should be normal to the tangent plane of the constraint surface. This means they must both reside in the normal to the tangent plane and hence they will either be colinear or for several constraints they will be linearly dependent. The geometry we consider here justifies the method. Linear algebra supplies the harder part which is that if two vectors are both orthogonal to the tangent plane then they must both be in the orthogonal complement to the tangent plane. The heart of the method of Lagrange multipliers is the orthogonal complement theory from linear algebra. Of course, you can be heartless and still sucessfully apply the method of Lagrange.

8.3 Lagrange multiplier method for one constraint

PROBLEM: FIND EXTREME VALUES OF $f(x_1, x_2, \ldots, x_n)$ on the level set $g(x_1, x_2, \ldots, x_n) = 0$.

We suppose the objective function f and the constraint g share some common domain $U \subseteq \mathbb{R}^n$ and $f, g : U \to \mathbb{R}$.

1.) If $P \in U$ gives max/min $f(P)$ then if we take any smooth path $\gamma : \mathbb{R} \to M$ (where $M = g^{-1}(\{0\})$) then $f \circ \gamma : \mathbb{R} \to \mathbb{R}$ and we see $f \circ \gamma$ should have maximum when $\gamma(t) = P$. Usually we set up $\gamma : \mathbb{R} \to M$ with $\gamma(0) = P$, and smooth $\Rightarrow \gamma'(0) \neq 0$. Consider then

$$\frac{d}{dt}\Big[(f \circ \gamma)(t)\Big] = \big[\nabla f\,(\gamma(t))\big] \cdot \gamma'(t)$$

But when $t = 0$ we have a critical pt. for $f \circ \gamma$
Hence $\underbrace{\boxed{(\nabla f)(P) \cdot \gamma'(0) = 0}}$.

this eq$^{\underline{n}}$ says $(\nabla f)(P)$ is orthogonal to an arbitrary tangent vector to constraint surface at P.

2.) Likewise, consider the path $\gamma : \mathbb{R} \to M$ composed with g; $g \circ \gamma : \mathbb{R} \to \mathbb{R}$. By def$^{\underline{n}}$ of M we have $(g \circ \gamma)(t) = 0 \quad \forall t$.

$$0 = \frac{d}{dt}\Big((g \circ \gamma)(t)\Big) = \Big[(\nabla g)(\gamma(t))\Big] \cdot \gamma'(t)$$

$$\therefore \quad \boxed{(\nabla g)(P) \cdot \gamma'(0) = 0}$$

Again this suggests $(\nabla g)(P)$ is likewise orthogonal to all tangent vectors in the tangent plane to $g = 0$.

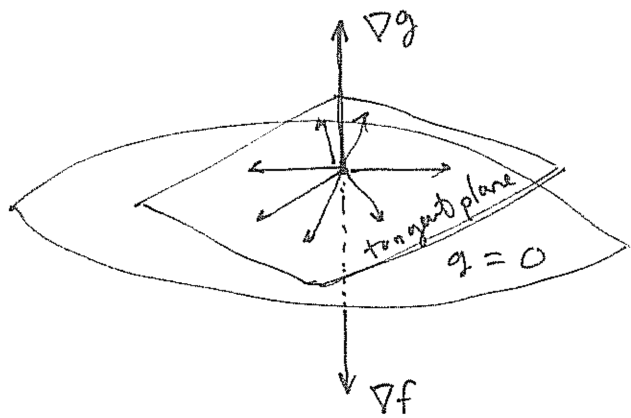
Continuing:

Think about this, $g(x_1, x_2, \ldots, x_n) = 0$ gives 1-eq$^{\underline{n}}$ in n-unknowns. It turns out $(\nabla g)(P) \neq 0$ makes $g(x_1, x_2, \ldots, x_n) = 0$ an $(n-1)$-dimensional surface. This means there is only 1-dimension left. Thus

$(\nabla f)(P) \cdot \gamma'(0) = 0$ AND $(\nabla g)(P) \cdot \gamma'(0) = 0$

$$\Rightarrow \boxed{\nabla f(P) = \lambda \nabla g(P)}$$

for some constant "Lagrange Multiplier" λ.



8.4 Lagrange multiplier method for several constraints

$\boxed{\text{PROBLEM: FIND EXTREMA FOR } f : U \subseteq \mathbb{R}^n \longrightarrow \mathbb{R} \text{ subject to constraints } G_1 = 0,\ G_2 = 0, \ldots, G_m = 0 \text{ where } G_j : U \subseteq \mathbb{R}^n \longrightarrow \mathbb{R} \text{ for } j = 1, 2, \ldots, m.}$

Observation: $G = (G_1, G_2, \ldots, G_m) = (0, 0, \ldots, 0)$ imposes m- constraints at once. A convenient def$^{\underline{n}}$ for the constraint surface is $G^{-1}\{(0,0,\ldots,0)\} = M$. This will give an $(n-m)$-dimensional surface or "manifold" in U provided a condition on $\nabla G_1, \nabla G_2, \ldots, \nabla G_m$ is met. (need $\text{rank}(G'(P)) = m$)

1.) Let $\gamma : I \subseteq \mathbb{R} \longrightarrow M$ with $\gamma(0) = P$ and $\gamma'(0) \neq 0$. Suppose f has extrema at P; $f \circ \gamma : \mathbb{R} \longrightarrow \mathbb{R}$ has critical pt (max/min maybe) at $t = 0$.

$$0 = \frac{d}{dt}\Big[(f \circ \gamma)(t)\Big] = \Big[\nabla f(\gamma(t))\Big] \cdot \gamma'(t)$$

Again $\boxed{\nabla f(P) \cdot \gamma'(0) = 0}$

2.) Note $G_j \circ \gamma : I \subseteq \mathbb{R} \longrightarrow \mathbb{R}$ is identically zero since γ is a curve whose image $\gamma(I) \subset M$. Thus

$$\frac{d}{dt}(0) = \frac{d}{dt}\Big[(G_j \circ \gamma)(t)\Big] = \Big[\nabla G_j(\gamma(t))\Big] \cdot \gamma'(t) = 0$$

This holds for $j = 1, 2, \ldots, m$ hence

$$\boxed{(\nabla G_j(P)) \cdot \gamma'(0) = 0 \quad \text{for } j = 1, 2, \ldots, m}$$

Continuing:

Provided $\text{rank}(G'(p)) = m$ it follows the set of tangent vectors at p span a $(n-m)$-dimensional subspace translated to the point p. Moreover, we note $(\nabla f)(p) \cdot \gamma'(0) = 0$ says $(\nabla f)(p)$ is in the orthogonal complement to the tangent space. On the other hand $\text{rank}(G'(p)) = m$ implies $\nabla G_1(p), \nabla G_2(p), \dots, \nabla G_m(p)$ gives a basis for the complement as each $\nabla G_j(p)$ is in the orthogonal complement of the tangent space by $\underline{\nabla G_j(p) \cdot \gamma'(0) = 0}$. It follows that ∇f must be a linear combination of the basis for the complement;

$$\boxed{\nabla f = \lambda_1 \nabla G_1 + \lambda_2 \nabla G_2 + \cdots + \lambda_m \nabla G_m}$$

Indeed, if p gives a critical point for $f \circ \gamma$ it must be subject to the boxed condition above for reasonably posed constraint $G = 0$.

Example: find points on circle $x^2+y^2=1$ and parabola $y^2 = 2(4-x)$ which are closest.

Let $f(x,y,u,v) = (x-u)^2 + (y-v)^2$ and construct $G(x,y,u,v) = (x^2+y^2-1,\ v^2+2u-8)$. Max/Min f subject to constraints $G = 0$.

$$\nabla f = \lambda_1 \nabla G_1 + \lambda_2 \nabla G_2$$

$$\langle 2(x-u), 2(y-v), -2(x-u), -2(x-v)\rangle = \lambda_1 \langle 2x, 2y, 0, 0\rangle + \lambda_2 \langle 0, 0, 2, 2v\rangle$$

Yields,

$$\left.\begin{aligned} 2(x-u) &= 2\lambda_1 x \\ 2(y-v) &= 2\lambda_1 y \end{aligned}\right\} \longrightarrow \underline{\frac{y}{x} = \frac{y-v}{x-u}} \quad \text{I}$$

$$-2(x-u) = 2\lambda_2 \longrightarrow -\lambda_2 = x-u$$

$$-2(y-v) = 2\lambda_2 v \longrightarrow \underline{y - v = (x-u)v} \quad \text{II}$$

Combining I & II

$$\frac{y}{x} = \frac{y-v}{x-u} = \frac{\cancel{(x-u)}v}{\cancel{(x-u)}} = v \quad \therefore \quad \underline{y = xv} \quad \text{III}$$

Given together with II $\longrightarrow$ $xv - v = (x-u)v$

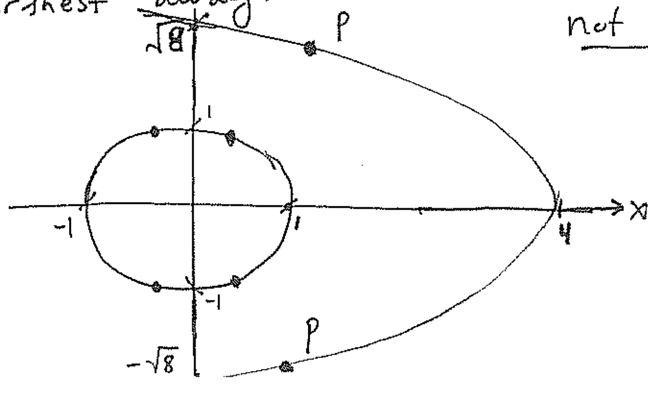
$$\Rightarrow xv - v = xv - uv$$
$$\Rightarrow v = uv$$
$$\Rightarrow \underline{u = 1}.$$

But, $v^2 = 2(4-u) = 2(3) = 6 \quad \therefore \quad \underline{v = \pm\sqrt{6}}$.

Hence, $y = (\pm\sqrt{6})x \Rightarrow x^2+y^2 = \underbrace{x^2 + 6x^2 = 1}_{7x^2=1} \quad \therefore \quad x = \pm 1/\sqrt{7}$

We find points $(1, \pm\sqrt{6})$ on the parabola and $(\pm 1/\sqrt{7}, \pm\sqrt{6}/\sqrt{7})$ on the circle are closest or furthest away.

<u>not to scale !</u>



$\boxed{\text{Example: find points on } x^2+y^2+z^2 = 1 \text{ and plane } x+y+z=3 \text{ which are closest.}}$

$g_1(x,y,z) = (1-x^2-y^2-z^2) = 0$ gives sphere. as $g_1^{-1}\{0\}$.
$g_2(u,v,w) = 3-u-v-w = 0$ gives plane as $g_2^{-1}\{0\}$.

Consider $f(\vec{x},\vec{u}) = \|\vec{x}-\vec{u}\|^2$. We'd like to find min/max for f subject the constraints

$$G(\vec{x},\vec{u}) = \langle g_1(\vec{x}), g_2(\vec{u})\rangle = \langle 0,0\rangle.$$

To say $G=0$ is to place $\vec{x}$ on the sphere and $\vec{u}$ on the plane. Let $G_1(\vec{x},\vec{u}) = g_1(\vec{x})$ and $G_2(\vec{x},\vec{u}) = g_2(\vec{u})$ then

$$\nabla G_1 = \langle \nabla g_1, 0\rangle = \langle g_{1x}, g_{1y}, g_{1z}, 0,0,0\rangle$$
$$\nabla G_2 = \langle 0, \nabla g_2\rangle = \langle 0,0,0, g_{2u}, g_{2v}, g_{2w}\rangle$$

Likewise, $f(\vec{x},\vec{u}) = \sum_{j=1}^{n}(x_j-u_j)^2 = (x-u)^2+(y-v)^2+(z-w)^2$

$$\nabla f(x,y,z,u,v,w) = \langle 2(x-u), 2(y-v), 2(z-w), -2(x-u), -2(y-v), -2(z-w)\rangle$$

Then $\nabla f = \lambda_1 \nabla G_1 + \lambda_2 \nabla G_2$ yields,

$$\left.\begin{aligned} 2(x-u) &= \lambda_1 g_{1x} = -2\lambda_1 x \\ 2(y-v) &= \lambda_1 g_{1y} = -2\lambda_1 y \\ 2(z-w) &= \lambda_1 g_{1z} = -2\lambda_1 z \end{aligned}\right\} \quad \frac{x-u}{x} = \frac{y-v}{y} = \frac{z-w}{z}$$

$$\left.\begin{aligned} -2(x-u) &= \lambda_2 g_{2u} = -\lambda_2 \\ -2(y-v) &= \lambda_2 g_{2v} = -\lambda_2 \\ -2(z-w) &= \lambda_2 g_{2w} = -\lambda_2 \end{aligned}\right\} \quad x-u = y-v = z-w$$

Thus, $\dfrac{x-u}{x} = \dfrac{x-u}{y} = \dfrac{x-u}{z} \Rightarrow \underline{x=y=z} \Rightarrow \underline{u=v=w}$

$$\underbrace{3x^2=1}_{x=\pm 1/\sqrt{3}} \qquad \underbrace{3u=3}_{u=1}$$

$\boxed{\text{We obtain the point } (1,1,1) \text{ on the plane is closest to } (1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3}) \text{ and furthest from } (-1/\sqrt{3}, -1/\sqrt{3}, -1/\sqrt{3}) \text{ on the sphere.}}$

<u>Remark</u>: p. 116 of Edwards derives this on basis of the <u>amazing</u> Example 10 of Edwards.

$\boxed{\text{Example: Find max/min of } w = x + z \text{ where } x^2+y^2+z^2 \leq 1.}$

Notice $f(x,y,z) = x + z$ has $\nabla f = \langle 1, 0, 1 \rangle$ thus f has no critical pts. It follows f must attain max/min on boundary $g(x,y,z) = x^2+y^2+z^2-1 = 0$. Use method of Lagrange,

$$\nabla f = \lambda \nabla g \quad \text{where } g = 0$$

$$\langle 1, 0, 1 \rangle = \lambda \langle 2x, 2y, 2z \rangle$$

$$\left.\begin{array}{l} 1 = 2\lambda x \\ 0 = 2\lambda y \\ 1 = 2\lambda z \end{array}\right\} \longrightarrow y = 0 \ \ \& \ \ 1 = \frac{x}{z} \Rightarrow z = x$$

$$\Rightarrow x^2 + 0^2 + x^2 = 1$$

$$\Rightarrow x^2 = \frac{1}{2}$$

$$\Rightarrow \underline{x = \pm 1/\sqrt{2}}.$$

Thus $(1/\sqrt{2}, 0, 1/\sqrt{2})$ or $(-1/\sqrt{2}, 0, -1/\sqrt{2})$ yield extremal values of f on $g = 0$.

$$f(1/\sqrt{2}, 0, 1/\sqrt{2}) = 1/\sqrt{2} + 1/\sqrt{2} = 2/\sqrt{2} = \sqrt{2}.$$

$$f(-1/\sqrt{2}, 0, -1/\sqrt{2}) = -1/\sqrt{2} - 1/\sqrt{2} = -2/\sqrt{2} = -\sqrt{2}.$$

$\boxed{\begin{array}{l}\text{The max. value is } \sqrt{2} \text{ reached at } (1/\sqrt{2}, 0, 1/\sqrt{2}) \text{ and} \\ \text{the min. value is } -\sqrt{2} \text{ reached at } (-1/\sqrt{2}, 0, -1/\sqrt{2}).\end{array}}$

Example: A plane wave in the z-direction has the form

$$\vec{E} = E_o \cos(kz - \omega t)\,\hat{z}$$

where E_o, k, ω are constants and $\hat{z} = \frac{\nabla z}{|\nabla z|} = \langle 0, 0, 1\rangle$.
Here $\vec{E} : \mathbb{R}^4 \longrightarrow \mathbb{R}^3$ is a time-dependent vector field.
One interesting function built from $\vec{E}$ is the square of its length;
define $f(t, x, y, z) = \vec{E}\cdot\vec{E} = E_o^2\cos^2(kz - \omega t)$.

PROBLEM: find critical points for f and find its max/min relative to $t = 0$ or $x^2+y^2+z^2 = R^2$ or $x^2+y^2+z^2 = (Rt)^2$

I choose the ordering $x_0 = t,\ x_1 = x,\ x_2 = y,\ x_3 = z$ for this problem. With respect to this ordering,

$$\begin{aligned}
\nabla f &= \left\langle \frac{\partial f}{\partial t}, \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}\right\rangle \\
&= \left\langle -2E_o^2\cos(kz-\omega t)\sin(kz-\omega t)(-\omega),\ 0,\ 0,\ -2E_o^2\cos(kz-\omega t)\sin(kz-\omega t)\,k\right\rangle \\
&= E_o^2\left\langle \omega\sin(2[kz-\omega t]),\ 0,\ 0,\ -k\sin(2[kz-\omega t])\right\rangle
\end{aligned}$$

Critical points must have $\nabla f = 0$ since f is everywhere differentiable. We see that both of the sine terms should vanish if we are to obtain $\nabla f = 0$. Hence,

$$\begin{aligned}
\nabla f = 0 &\iff n\pi = 2(kz - \omega t), \quad n \in \mathbb{Z}. \\
&\iff kz - \omega t = \frac{n\pi}{2}, \quad n \in \mathbb{Z}. \\
&\iff \underline{z = \frac{\omega t + n\pi/2}{k}, \quad n \in \mathbb{Z}.}
\end{aligned}$$

Think about it, for a fixed time there are planes of critical points every $\frac{\pi}{2k}$ units in the z-direction. On the other hand, if we imagine time flowing then you can envision these planes of critical points flowing upward with a speed of $\frac{dz}{dt} = \frac{\omega}{k}$.

Now let's think about the constraints

(i.) $t = 0$, (ii) $x^2+y^2+z^2 = R^2$, (iii) $x^2+y^2+z^2 = (Rt)^2$

see over

(i.) $t = 0$ gives $g_1(t,x,y,z) = t = 0$ study $\nabla f = \lambda \nabla g$,

$$\nabla f = \lambda \nabla g \longrightarrow E_0^2 \langle w\sin(2kz), 0, 0, -k\sin(2kz)\rangle = \lambda\langle 1, 0, 0, 0\rangle$$

$$\hookrightarrow \quad \begin{aligned} E_0^2 w \sin(2kz) &= \lambda \\ -E_0^2 w \sin(2kz) &= 0 \end{aligned}$$

These equations are inconsistent unless $2kz = n\pi$ for some $n \in \mathbb{Z}$. However, in that case $\nabla f = 0$ thus Lagrange multipliers don't apply. We can still analyze the function despite the failure of the method, for $t = 0$ and $2kz = n\pi$ we have

$$f\left(0, x, y, \frac{n\pi}{2k}\right) = E_0^2 \cos^2\left[k\left(\frac{n\pi}{2k} - w(0)\right)\right] = E_0^2 \cos^2\left(\frac{n\pi}{2}\right)$$

Note for $n \in 2\mathbb{Z} + 1$ (odd n) we have $f = 0$ whereas for $n \in 2\mathbb{Z}$ (even n) we have $f = E_0^2$.

Clearly these give max/min for f relative to the constraint $t = 0$.

$$f\left(0, x, y, \frac{n\pi}{2k}\right) = \begin{cases} 0 & \text{if } n \text{ odd (minimum)} \\ E_0^2 & \text{if } n \text{ even (maximum)} \end{cases}$$

The planes of critical pts. are now separated into max/min

(ii.) $x^2 + y^2 + z^2 = R^2$ encoded by $g(x,y,z) = x^2 + y^2 + z^2 - R^2 = 0$

study $\nabla f = \lambda \nabla g$, where $g = 0$

$$E_0^2 \langle w\sin[2(kz - wt)], 0, 0, -k\sin 2(kz - wt)\rangle = \lambda\langle 0, 2x, 2y, 2z\rangle$$

$$E_0^2 w \sin(2(kz - wt)) = 0 \quad \Rightarrow \quad kz - wt = \frac{n\pi}{2}, \; n \in \mathbb{Z}$$

$$\left.\begin{aligned} 0 &= 2\lambda x \\ 0 &= 2\lambda y \end{aligned}\right\} \Longrightarrow \underline{x = y = 0} \Rightarrow \underline{z^2 = R^2}.$$

$$-E_0^2 k \sin(2(kz - wt)) = 2\lambda z \quad \Rightarrow \quad 0 = 2\lambda z$$

$$\Rightarrow \lambda = 0 \text{ or } z = 0$$

Either $\lambda = 0$, $kz - wt = \frac{n\pi}{2}$.

or $x = y = z = 0$ and $-2wt = \frac{n\pi}{2}$. ← not possible.

$$\boxed{z = \frac{wt + n\pi/2}{k} \quad \text{and} \quad x^2 + y^2 = R^2 - z^2}$$

(iii) $x^2+y^2+z^2 = (Rt)^2 \leftarrow$ spherical wave front

$$g(t,x,y,z) = R^2t^2 - x^2 - y^2 - z^2$$

$$\nabla f = \lambda \nabla g \quad , \quad g = 0$$

$$E_o^2 \langle \omega \sin 2(kz-\omega t), 0, 0, -k\sin 2(kz-\omega t)\rangle = \lambda \langle 2R^2 t, -2x, -2y, -2z\rangle$$

$$\begin{cases} E_o^2 \omega \sin 2(kz-\omega t) = 2\lambda R^2 t \\ 0 = -2x\lambda \\ 0 = -2y\lambda \\ -E_o^2 k \sin 2(kz-\omega t) = -2\lambda z \end{cases}$$

$$\underbrace{\frac{R^2t}{\omega} = \frac{z}{k}} = \frac{E_o^2 \sin 2(kz-\omega t)}{2\lambda} \quad \& \quad x=y=0.$$

$$\underline{z = \left(\frac{kR^2}{\omega}\right)t} \quad \text{or} \quad \underline{t = \left(\frac{\omega}{kR^2}\right)z}.$$

where $x=y=0$. Hence, $z^2 = R^2t^2$

$$\Longrightarrow \left(\frac{kR^2}{\omega}t\right)^2 = R^2t^2$$

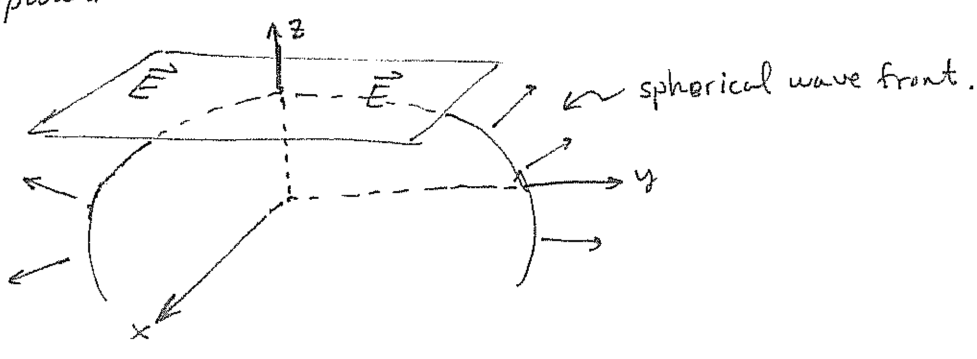
$$\Longrightarrow \left(\frac{k^2R^4}{\omega^2} - R^2\right)t^2 = 0$$

$$\Longrightarrow \left(\frac{k^2R^2}{\omega^2} - 1\right)R^2t^2 = 0$$

$$\therefore \; R^2 = \frac{\omega^2}{k^2} \quad \text{or} \quad t = 0.$$

The spherical wave-front $x^2+y^2+z^2=(Rt)^2$ collides with a max/min of the plane-wave $\vec{E}$ provided that $R = \frac{\omega}{k}$ (we assume $\omega, k > 0$ for physical reasons)

Remark: the two waves match at time zero then flow upward with same speed of $\frac{\omega}{k}$ along z-axis.



Chapter 9

theory of differentiation

In the last chapter I began by announcing I would apply the central theorems of this course to solve interesting applied problems. If you remembered that I said that you may be a bit perplexed after completing the preceding chapter. Where did we use these theorems? It would seem we mostly just differentiated and pulled a magic λ from the thin air. Where did we use the inverse or implicit mapping theorems? It's subtle. These theorems go to the existence of a mapping, or the solution of a system of equations. Often we do not even care about finding the inverse or solving the system. The mere existence justifies other calculations we do make explicit. In this chapter I hope to state the inverse and implicit function theorems carefully. I leave the complete proofs for Edward's text, we will just discuss portions of the proof. In particular, I think it's worthwhile to discuss Newton's method and the various generalizations which reside at the heart of Edward's proof. In contrast, I will take it easy on the analysis. The arguments given in Edward's generalize easily to the infinite dimensional case. I do think there are easy arguments but part of his game-plan is set-up the variational calculus chapter which is necessarily infinite-dimensional. Finally, I conclude this chapter by examining a few examples of constrained partial differentiation.

9.1 Newton's method for solving the insolvable

I'll begin with a quick review of Newton's method for functions.

Problem: given a function $f : \mathbb{R} \to \mathbb{R}$ which is continuously differentiable on $[a, b]$ and $f(a) < 0 < f(b)$ with $f'(x) > 0$ for each $x \in [a, b]$ **how** can we find the solution to $f(x) = 0$ w.r.t. the interval $[a, b]$?

Solution: Newton's Method. In a nutshell, the idea is to guess some point in $x_o \in [a, b]$ and then replace the function with the tangent line to $(x_o, f(x_o))$. Then we can easily calculate the zero of the tangent line through elementary algebra.

$$y = L_f^{x_o}(x) = f(x_o) + f'(x_o)(x - x_o) = 0 \qquad \Rightarrow \qquad x = x_o - \frac{f(x_o)}{f'(x_o)}$$

Now, this is just the first approximation, we can apply the idea again to our new guess $x_1 = x$; that is define $x_1 = x_o - \frac{f(x_o)}{f'(x_o)}$ and think of x_1 as our new "x_o". The zero of the tangent line to $(x_1, f(x_1))$ is called x_2 and we can calculate,

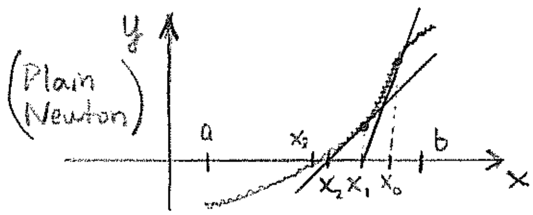
$$y = L_f^{x_1}(x) = f(x_1) + f'(x_1)(x - x_1) = 0 \qquad \Rightarrow \qquad \boxed{x_2 = x_1 - \frac{f(x_1)}{f'(x_1)}}$$

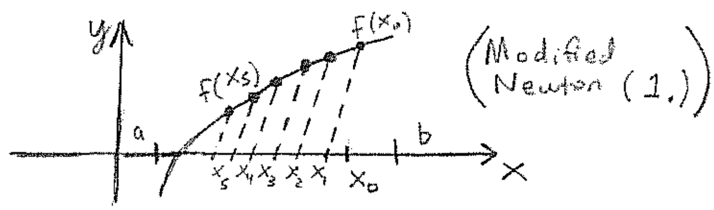
Notice that if $f(x_1) = 0$ then we found the zero and the method just gives $x_2 = x_1$. The idea then is to continue in this fashion and define the n-th guess iteratively by

Newton's Method:

$$\boxed{x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}}$$

If for some particular n we actually find the exact value of the zero then the iteration just stays on that value. Otherwise, it can be shown that $\lim_{n\to\infty} x_n = x_*$ where $f(x_*) = 0$.





This is the simplest form of Newton's method but it is also perhaps the hardest to code. We'd have to calculate a new value for the derivative for each step. Edwards gives two modifications of the method and proves convergence for each.

Modified Newton Methods:

1. $\boxed{x_{n+1} = x_n - \frac{f(x_n)}{M}}$ where we know $0 < m < f'(x) < M$.

2. $\boxed{x_{n+1} = x_n - \frac{f(x_n)}{f'(a)}}$ where we know $f'(a) \neq 0$.

In case (1.) Edwards uses the concept of a **contraction mapping** to prove that the sequence converges and he even gives an estimate to bound the error of the guess (see Theorem 1.2 on pg. 164). Then he cautions against (2.) because it is possible to have Fig. 3.2 on pg. 162 occur, in other words if we guess badly to begin we might never find the root x_*. The remedy is fairly simple, you just look on smaller intervals. For (2.) he states the result concretely only in a local case (see Theorem 1.3 on pg. 165). I actually have only stated a particular case of his Theorem since I have made $b = 0$. The proof of the inverse function theorem builds from method (2.) but I'll give an example of (1.) because it's interesting and it should help make this whole discussion a little more tangible.

Example 9.1.1. Let $f(x) = \sin(x) + \frac{x}{2} - 1$. Find zero on $[-\pi/2, \pi/2]$.
Notice $f'(x) = \cos(x) + \frac{1}{2}$ hence $f'(x) > 0$ for $-\frac{\pi}{2} \leq x \leq \frac{\pi}{2}$.
Moreover, $\frac{1}{2} \leq f'(x) \leq \frac{3}{2}$ for $x \in [-\pi/2, \pi/2]$.
Guess $x_o = 0$ and use $M = 3/2$.

$$x_1 = x_o - \tfrac{2}{3}f(x_o) = \tfrac{2}{3}.$$
$$x_2 = x_1 - \tfrac{2}{3}f(x_1) \cong 0.6989$$
$$x_3 = x_2 - \tfrac{2}{3}f(x_2) \cong 0.7037$$
$$x_4 = x_3 - \tfrac{2}{3}f(x_3) \cong 0.7044$$
$$x_5 = x_4 - \tfrac{2}{3}f(x_4) \cong \underbrace{0.7045}_{\text{last digit uncertain}}$$

Remark: I had to zoom-in about 5 times on a TI-83 to verify this result

$$\boxed{x \cong 0.705}$$

In case (2.) we can actually solve the equation $f(x) = y$ for a given value y close to b provided $f(a) = b$ and $f'(a) \neq 0$. The idea here is just to replace the function at $(x_o, f(x_o))$ with the line $L(x) = f(x_o)+f'(a)(x-x_o)$ and then we solve $L(x) = y$ to obtain $x = x_o - \frac{f(x_o)-y}{f'(a)}$. Note here we use the slope from the point (a, b) throughout the iteration, in particular we say $x_1 = x$ and start iterating as usual: $x_{n+1} = x_n - \frac{f(x_n)-y}{f'(a)}$ (see Theorem 1.3 on pg. 165 in Edwards for proof this converges)

> **Problem:** given a function $f : \mathbb{R} \to \mathbb{R}$ which is continuously differentiable near a and $f'(a) \neq 0$, can we find a function g such that $f(g(y))) = y$ for y near the image $f(a)$?

Solution: Modified Newton's Method. we seek to solve $f(g(y)) = y$ for y in some neighborhood of a, simply define $g_o(y) = a$ and apply the method

$$\boxed{g_{n+1}(y) = g_n(y) - \frac{f(g_n(y)) - y}{f'(a)}}$$

Notice this can be done for each y near $f(a)$, in other words, we have a **sequence of functions** $\{g_n\}_{n=0}^{\infty}$. Moreover, if we take $n \to \infty$ this sequence **uniformly** converges to an exact solution g. This gives us an iterative way to construct local inverse functions for some given continuously differentiable function at a point a such that $f'(a) \neq 0$.

The idea of convergence of functions begs the question of what precisely is the "length" or "norm" of a function. Again, I postpone such discussion until the very end of the course. For now just accept that the idea of convergence of sequences of functions is well defined and intuitively it just means that the sequence matches the limiting function as $n \to \infty$. You encountered this idea in the discussion of Taylor series in calculus II, one can ask whether the sequence of Taylor polynomials for f does converge to f relative to some interval of convergence.

The calculations that follow here amaze me.

Example 9.1.2. (this is also in Edwards' Text)

$$f(x) = x^2 - 1, \quad a = 1, \quad b = 0 \Rightarrow f'(1) = 2(1) = 2.$$

$$g_0(y) = 1$$

$$g_1(y) = g_0(y) - \tfrac{1}{2}[f(g_0(y)) - y] = 1 - \frac{0-y}{2} = 1 + \tfrac{1}{2}y$$

$$g_2(y) = (1 + \tfrac{1}{2}y) - \tfrac{1}{2}[(1 + y/2)^2 - y] = 1 + \tfrac{1}{2}y - \tfrac{1}{8}y^2$$

Thus the inverse of $f(x)$ near $a = 1$ is,

$$g(y) = 1 + \tfrac{1}{2}y - \tfrac{1}{8}y^2 + \cdots$$

Recall $(1+y)^k = 1 + ky + \frac{1}{2}k(k-1)y^2 + \cdots$ observe our $g(y)$ fits the binomial series with $k = 1/2$ thus

$$\boxed{g(y) = \sqrt{1+y} = f^{-1}(y)} \leftarrow$$ technically I should say $f|_{(-1,1)}^{-1}$

Example 9.1.3.

Usually this algorithm gives much weirder results. For example, $f(x) = \sin(x)$ take $(0,0)$ as base point note $f'(0) = \cos(0) = 1$.

$$g_0(y) = 0$$

$$g_1(y) = g_0(y) - [f(g_0(y)) - y] = 0 - \sin(0) + y = y.$$

$$g_2(y) = y - [\sin(y) - y] = 2y - \sin y$$

$$g_3(y) = 2y - \sin y - [\sin(2y - \sin y) - y]$$

Hence $\boxed{f^{-1}(y) \approx 3y - \sin(y) - \sin(2y - \sin y)}$

As a quick check try $f(g_3(y))$, use $\sin\theta \approx \theta$ for $\theta \approx 0$,

$$\begin{aligned} f(g_3(y)) &= \sin(3y - \sin y - \sin(2y - \sin y)) \\ &\cong \sin(3y - y - 2y + y) \\ &\cong \sin(y) \approx y. \end{aligned}$$

It would be interesting to implement this algorithm in Mathematica.

9.1.1 local solutions to level curves

Next, we try a similar technique to solve equations of the form $G(x,y) = 0$. You should recall that the solution set of $G(x,y) = 0$ is called a **level curve**. Usually we cannot make a global solution for y; in other words, there does not exist $f(x)$ such that $G(x, f(x)) = 0$ for all x in the solution set of G. For example, $G(x,y) = x^2 + y^2 - 1$ allows us to cast the unit circle as the solution set of the equation $G(x,y) = 0$. But, the unit circle is not the graph of a single function since it fails the vertical line test. Instead we need a pair of functions to cover the circle. Generally the situation can get quite complicated. Let's pause to notice there are two points where we cannot find a solution to $G(x,y) = 0$ on an open disk about the point: these points are $(-1,0)$ and $(1,0)$. We have trouble at the vertical tangents, note $G_y(x,y) = 2y$ has $G_y(-1,0) = G_y(1,0) = 0$ [1].

Idea: use the Newton's method approach to find solution, however, the approach here is slightly indirect. We'll use the mean value theorem to replace a function with its tangent line. Consider a fixed x_* near a then we have an function of y alone: $h(y) = G(x_*, y)$. Apply the mean value theorem to h for a y-value y_* such that point (x_*, y_*) has $G(x_*, y_*) = 0$,

$$G_y(x_*, b) = \frac{G(x_*, y_*) - G(x_*, b)}{y_* - b} = -\frac{G(x_*, b)}{y_* - b}$$

We can solve for y_* to obtain:

$$y_* = b - \frac{G(x_*, b)}{G_y(x_*, b)}$$

Define $f_o(x) = b$ and define $f_1(x)$ by

$$f_1(x) = f_o(x) - \frac{G(x, f_o(x))}{G_y(x, f_o(x))} \quad \text{and} \quad f_2(x) = f_1(x) - \frac{G(x, f_1(x))}{G_y(x, f_1(x))} \quad \text{and so forth...}$$

Foruntately, Edwards proves we can use an easier formula where the denomimator is replaced with $G_y(a,b)$ which is pretty close to the formula we have above provided the point considered is close to (a,b).

Theorem 9.1.4. *(Theorem 1.4 in Edwards's Text)*

> Let $G : \mathbb{R}^2 \to \mathbb{R}$ be continuously differentiable and (a,b) a point such that $G(a,b) = 0$ and $G_y(a,b) \neq 0$ then we can find a function f on some closed interval J centered at a which covers the solution set of $G(x,y) = 0$ near all points close to (a,b). Moreover, this **local solution** is the limit of the sequence of functions inductively defined below:
>
> $$f_o(x) = b \quad \text{and} \quad f_{n+1}(x) = f_n(x) - \frac{G(x, f_n(x))}{G_y(a,b)}$$
>
> for all $n \in \mathbb{N}$. **We can calculate solutions iteratively!**

[1] yes, if we used closed disks then we could find a solution on a disk where $(-1,0)$ or $(1,0)$ was on the boundary, the point of the discussion is to motivate the implicit function theorem's langauge

Look at Example 2 on page 170 for a nice straight-forward application of Theorem 1.4. Perhaps you're not too excited by this example. Certainly, algebra solves the problem with ease anyway, we just have to take care with the algebraic steps. I intend for the next example to confound algebraic techniques and yet we can find an approximate solution:

Example 9.1.5. *Let $G(x,y) = exp(x^2+y^2)+x-e$. Notice that $G(0,1)=0$ and $G_y(0,1)=2$. Apply the algorithm:*

$$\begin{aligned} f_o(x) &= 1 \\ f_1(x) &= 1-\tfrac{1}{2}G(x,1) = 1-\tfrac{1}{2}(exp(x^2+1)+x-e) \\ f_2(x) &= f_1(x)-\tfrac{1}{2}[exp(x^2+[f_1(x)]^2+x-e] \end{aligned}$$

I'd go on but it just gets ugly. What is neat is that

$$y = f_1(x) = 1-\tfrac{1}{2}(exp(x^2+1)+x-e)$$

gives an approximation of a local solution of $exp(x^2+y^2)+x-e=0$ for points near $(0,1)$.

Example 9.1.6. *Let $G(x,y) = x^2+y^2+y-1$ note $G_y = 2y+1$. Note that $G(1,0)=0$ and $G_y(1,0)=1$. Calculate the local solution by the algorthim:*

$$\begin{aligned} f_o(x) &= 0 \\ f_1(x) &= 0-G(x,0) = 1-x^2 \\ f_2(x) &= 1-x^2-G(x,1-x^2) = x^2-x^4 \\ f_3(x) &= x^2-x^4-G(x,x^2-x^4) = 1-x^2-x^4+2x^6-x^8 \end{aligned}$$

Now, these formulas are somewhat bizarre because we are writing an approximation centered at $x=1$ as polynomials centered at zero. It is probable that a nicer pattern emerges if we were to write all of these as polynomials in $(x-1)$. Notice that $f_n(1)=0$ for $n=0,1,2,3$.

Example 9.1.7. *Let $G(x,y) = x^2+y^2+y-2$ note $G_y = 2y+1$. Note that $G(0,1)=0$ and $G_y(0,1)=3$. Calculate the local solution by the algorthim:*

$$f_o(x) = 1$$

$$\begin{aligned} f_1(x) &= 1-\tfrac{1}{3}G(x,1) \\ &= 1-\tfrac{1}{3}x^2 \end{aligned}$$

$$\begin{aligned} f_2(x) &= 1-\tfrac{1}{3}x^2-G(x,1-\tfrac{1}{3}x^2) \\ &= 1-\tfrac{1}{3}x^2-[x^2+(1-\tfrac{1}{3}x^2)^2+(1-\tfrac{1}{3}x^2)-2] \\ &= 1-\tfrac{1}{3}x^2-\tfrac{1}{9}x^4 \end{aligned}$$

Note how the approximation unfolds order by order when the center matches the format in which we write the expansion.

If the center $a \neq 0$ then what can happen is that the terms of a particular order get spread across all orders in the Newton's Method approximation. I've found the expansions generated from the Newton's method are not easy to write in a nice form in general... of course, this shouldn't be that surprising, the method just gave us a way to solve problems that defy closed-form algebraic solution.

9.1.2 from level surfaces to graphs

In the preceding section we found that $G(x, y) = 0$ could be understood as a graph of a function of a single variable locally, in other words we found a 1-manifold. When we have an equation of n-variables it will likewise find $(n-1)$ free variables. This means that $G(x, y, z) = x^2+y^2+z^2-1 = 0$ gives us a level-surface (the sphere), or $G(t, x, y, z) = -t^2+x^2+y^2+z^2 = 0$ gives a level-volume (the light cone[2]). If we can solve the equation $G(x_1, x_2, \ldots, x_n)$ for x_j then we say we have re-written the level surface as a graph. This is important because graphs are a special case of a parametrized manifold, the parametric formalism allows us to set-up integrals over higher-dimensional surfaces and so forth. These things will become clearer when we study integration of differential forms later in this course. I state Theorem 1.5 in Edwards here for completeness. The essential point is this, if $\nabla G(p) \neq 0$ then there exists j such that $\frac{\partial G}{\partial x_j}(p) \neq 0$ and we can solve for x_j by using basically the same the iterative process we just worked out in the $n = 2$ case in the preceding subsection.

Theorem 9.1.8. *(Theorem 1.5 in Edwards's Text)*

Let $G : \mathbb{R}^n \to \mathbb{R}$ be continuously differentiable and $p = (a_1, a_2, \ldots, a_n)$ a point such that $G(p) = 0$ and $G_j(p) \neq 0$ then we can find a function f on some closed interval J centered at a_j which covers the solution set of $G(x_1, x_2, \ldots, x_n) = 0$ near all points close to p. Moreover, this **local solution** is the limit of the sequence of multivariate functions inductively defined below:

$$f_o(\vec{x}) = a_j \quad \text{and} \quad f_{n+1}(\vec{x}) = f_n(\vec{x}) - \frac{G(x_1, \ldots, f(\vec{x}), \ldots, x_n)}{G_{x_j}(p)}$$

for all $n \in \mathbb{N}$. If $f = \lim_{n\to\infty} f_n$ then $G(x_1, \ldots, f(\vec{x}), \ldots, x_n) = 0$ for points near p.

Something interesting happens when we apply this theorem to examples which allow explicit closed-form algebraic solution.

Example 9.1.9. *Consider* $G(x, y, z) = x+y+2z-4 = 0$. *Note that* $G_z = 2 \neq 0$ *and* $G(1, 1, 1) = 0$. *Apply the algorithm:*

$$\begin{aligned}
&f_o(x, y) = 1 \\
&f_1(x, y) = 1 - \tfrac{1}{2}G(x, y, 1) = 1 - \tfrac{1}{2}(x + y + 2 - 4) = -\tfrac{1}{2}(x + y - 4) \\
&f_2(x, y) = -\tfrac{1}{2}(x + y - 4) - \tfrac{1}{2}G(x, y, -\tfrac{1}{2}(x + y - 4)) = f_1(x, y)
\end{aligned}$$

You can clearly see that $f_n = f_1$ *for all* $n \geq 1$ *thus* $\lim_{n\to\infty} f_n = f_1$. *In other words, we found the exact solution is* $z = -\frac{1}{2}(x + y - 4)$.

[2]physically this represents the border of the spacetime which we can interact with in the future or the past, granting that special relativity actually describes nature without exception...

You might wonder if this just happened because the preceding example was linear, in fact, it has little to do with it. Here's another easy example,

Example 9.1.10. *Consider* $G(x,y,z) = x^2+y^2-z = 0$. *Note that* $G_z = -1 \neq 0$ *and* $G(0,0,0) = 0$. *Apply the algorithm:*

$$\begin{aligned}
f_o(x,y) &= 0 \\
f_1(x,y) &= 0 + G(x,y,0) = x^2 + y^2 \\
f_2(x,y) &= x^2 + y^2 + G(x,y,x^2+y^2) = x^2 + y^2 + [x^2 + y^2 - (x^2 + y^2)] = f_1(x,y)
\end{aligned}$$

You can clearly see that $f_n = f_1$ *for all* $n \geq 1$ *thus* $\lim_{n\to\infty} f_n = f_1$. *In other words, we found the exact solution is* $z = x^2 + y^2$.

Part of the reason both of the preceding examples were easy is that the solutions were not just local solutions, in fact they were global. When the solution is the level surface equation breaks up into cases it will be more complicated.

Example 9.1.11. *Suppose* $G(x,y,z) = \sin(x+y-z) = 0$ *then solutions must satisfy* $x+y-z = n\pi$ *for* $n \in \mathbb{Z}$. *In other words, the algorithm ought to find* $z = x+y-n\pi$ *where the choice of* n *depends on the locality we seek a solution. This level-set is actually a whole family of disconnected paralell planes. Let's see how the algorithm deals with this, feed it* $(0,0,2\pi)$ *as the starting point (this ought to select the* $n = -2$ *surface. Apply the algorithm to* $G(x,y,z) = \sin(x+y-z)$ *where clearly* $G(0,0,2\pi) = 0$ *and* $G_z = -\cos(-2\pi) = -1$ *hence:*

$$\begin{aligned}
f_o(x,y) &= 2\pi \\
f_1(x,y) &= 2\pi + G(x,y,2\pi) = 2\pi + \sin(x+y+2\pi) = 2\pi + \sin(x+y) \\
f_2(x,y) &= 2\pi + \sin(x+y) + \sin(x+y+\sin(x+y)) \\
f_3(x,y) &= 2\pi + \sin(x+y) + \sin(x+y+\sin(x+y)) \\
&\quad + \sin(x+y+\sin(x+y) + \sin(x+y+\sin(x+y)))
\end{aligned}$$

I deem these formulas weird. Perhaps I can gain some insight by expanding f_1,

$$f_1(x,y) = 2\pi + x + y - \tfrac{1}{3!}(x+y)^3 + \cdots$$

I'm a little scared to look at f_2. *There must be some sort of telescoping that happens in order for us to obtain the real solution of* $z = x + y + 2\pi$.

It's not at all obvious to me how the formula above telescopes in the limit that $n \to \infty$. However, unless I'm missing something or making a silly mistake, it seems clear that G is continuously differentiable at $(0,0,2\pi)$ and $G_z(0,0,2\pi) \neq 0$. Therefore, Theorem 1.5 applies and the sequence of function f_n should uniformly converge to the solution we know exists through direct argument in this example. Anyway, my point in this section is not to make a blanket endorsement that you solve all equations by the algorthim. I am merely trying to illustrate how it works.

9.2 inverse and implicit mapping theorems

In the preceding section we began by motivating the inverse function theorem for functions of one variable. In short, if the derivative is nonzero at a point then the function is 1-1 when restricted to a neighborhood of the point. Newton's method, plus a bunch of careful analysis about contraction mappings which we skipped this semester, then gave an algorithm to calculate the local inverse for a function of one variable. After that we essentially applied the local inverse idea to the problem of solving a level curve $G(x, y) = 0$ locally for an explicit solution of y. The result that such a solution is possible near points where $G_y \neq 0$ is known as the **implicit function theorem**. We then concluded by observing that almost the same mathematics allowed us to find an explicit solution of $G(x_1, \dots, x_{n+1}) = 0$ for one of the variables provided the partial derivative in that direction was nonzero. This result is also called the **implicit function theorem**. We used these theorems implicitly when I pulled parametrizations from my imagination, typically it is the implicit function theorem that justifies such a step. Moreover, to insist $\nabla g(p) \neq 0$ means that there exists at least one partial derivative nonzero so the implicit function theorem applies. All of that said, this section is basically the same story again. Difference is we have to deal with a little extra notation and linear algebra since a mapping is actually an ensemble of functions dealt with at once.

9.2.1 inverse mapping theorem

Suppose $f : \mathbb{R}^n \to \mathbb{R}^n$ has an inverse $f^{-1} = g$ then we have $f \circ g = Id$ so the chain rule yields $df \circ dg = d(Id) = Id$ since the identity is a linear map and hence it is its own best linear approximation. Note that we find that $f'g' = I_n$ thus $(f')^{-1} = g'$ or in other notation $[f']^{-1} = [f^{-1}]'$. With this in mind we wish to find a formula to calculate the inverse function. The definition seems like a good place to start:

$$\begin{aligned}
f(g(y)) = y \quad &\Rightarrow \quad g(y) = f^{-1}(y) \\
&\Rightarrow \quad g(y) \approx g(f(a)) + g'(a)[y - f(a)] \\
&\Rightarrow \quad g(y) \approx a + [f'(a)]^{-1}[y - f(a)] \\
&\Rightarrow \quad g_1(y) = g_o(y) + [f'(a)]^{-1}[y - f(g_o(y))] \quad \text{where} \quad g_o(y) = a \\
&\Rightarrow \quad g_{n+1}(y) = g_n(y) + [f'(a)]^{-1}[y - f(g_n(y))] \quad \text{where} \quad g_o(y) = a
\end{aligned}$$

Theorem 9.2.1. *(Theorem 3.3 in Edwards's Text see pg 185)*

> 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
>
> $$g_o(y) = a \quad \text{and} \quad g_{n+1}(y) = g_n(y) - [f'(a)]^{-1}[f(g_n(y)) - y]$$
>
> for all $y \in V$.

Notice this theorem gives us a way to test coordinate mappings for invertibility, we can simply calculate the derivative matrix then calculate its determinant to check to see it is nonzero to insure invertibility and hence the local invertibility of the coordinate map. There still remains the danger that the mapping doubles back to the same value further out so if we insist on a strict one-one correspondance then more analysis is needed to make sure a given transformation is indeed a coordinate system. (see Ex 1 on pg. 183 for a function which is everywhere locally invertible and yet not an injective mapping)

Example 9.2.2.

$$f(r,\theta) = (r\cos\theta,\ r\sin\theta) \longrightarrow f'(r,\theta) = \begin{bmatrix} \cos\theta & \sin\theta \\ -r\sin\theta & r\cos\theta \end{bmatrix}$$

$$\det\big(f'(r,\theta)\big) = r\cos^2\theta + r\sin^2\theta = r.$$

It follows f is locally invertible everywhere except the origin. We know from previous experience $\cos(\theta + 2\pi) = \cos\theta$ spoils the existence of a global inverse on $\mathbb{R}^2 - \{(0,0)\}$.

Example 9.2.3.

$f(\vec{v}) = \vec{v} \times \vec{a}$ where $\vec{a} \neq 0$ is a fixed vector in $\mathbb{R}^3$ thus $f : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$. Let $\vec{a} = \langle a_x, a_y, a_z \rangle$

$$f(x,y,z) = \langle x,y,z \rangle \times \langle a_x, a_y, a_z \rangle$$

$$= \langle \underbrace{y a_z - z a_y}_{f_1},\ \underbrace{z a_x - x a_z}_{f_2},\ \underbrace{x a_y - y a_x}_{f_3} \rangle$$

$$f'(x,y,z) = \begin{bmatrix} 0 & a_z & -a_y \\ -a_z & 0 & a_x \\ a_y & -a_x & 0 \end{bmatrix}$$

$$\det\left[f'(x,y,z)\right] = a_z a_x a_y - a_y a_x a_z = 0.$$

Therefore, f is nowhere invertible. For example, $\vec{a} = \langle 0,0,1 \rangle$

$$f(\vec{v}) = \vec{v} \times \langle 0,0,1 \rangle = \langle v_y, -v_x, 0 \rangle$$

for $\vec{a} = \langle 0,0,1 \rangle$ this always happens.

9.2.2 implicit mapping theorem

Let me begin by stating the problem we wish to consider:

Given continuously differentiable functions $G_1, G_2, \dots, G_n$

$$\begin{aligned} G_1(x_1, \dots, x_m, y_1, \dots, y_n) &= 0 \\ G_2(x_1, \dots, x_m, y_1, \dots, y_n) &= 0 \\ &\vdots \\ G_n(x_1, \dots, x_m, y_1, \dots, y_n) &= 0 \end{aligned}$$

Locally solve $y_1, \dots, y_n$ as functions of $x_1, \dots x_m$. That is, find a mapping $h : \mathbb{R}^m \to \mathbb{R}^n$ such that $G(x, y) = 0$ iff $y = h(x)$ near some point $(a, b) \in \mathbb{R}^{m+n}$ such that $G(a, b) = 0$. In this section we use the notation $x = (x_1, x_2, \dots x_m)$ and $y = (y_1, y_2, \dots, y_n)$.

It is convenient to define partial derivatives with respect to a whole vector of variables,

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

Consider $h : \mathbb{R}^m \to \mathbb{R}^n$ such that $G(x, y) = 0$ iff $y = h(x)$ near some point $(a, b) \in \mathbb{R}^{m+n}$ such that $G(a, b) = 0$. In other words, suppose $G(x, h(x)) = 0$. The chain rule reads:

$$0 = \frac{\partial G}{\partial x} + \frac{\partial G}{\partial y} h'(x)$$

Or, provided the matrix $\frac{\partial G}{\partial y}$ is invertible we can calculate,

$$h'(x) = -\left[\frac{\partial G}{\partial y}\right]^{-1} \frac{\partial G}{\partial x}$$

Theorem 9.2.4. *(Theorem 3.4 in Edwards's Text see pg 190)*

Let $G : \mathbb{R}^{n+m} \to \mathbb{R}^n$ be continuously differentiable in a open ball about the point (a, b) where $G(a, b) = 0$. 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 $\mathbb{R}^{n+m}$ and a continuously differentiable mapping $h : U \to \mathbb{R}^n$ such that $G(x, y) = 0$ 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

$$h_o(x) = b, \quad h_{n+1} = h_n(x) - [\tfrac{\partial G}{\partial y}(a, b)]^{-1} G(x, h_n(x))$$

for all $x \in U$.

I have given barely enough details to understand the notation here. If you read pages 188-194 of Edwards you can have a much deeper understanding. I will not attempt to recreate his masterpiece here. One important notation I should mention is the so-called Jacobian of G with respect to y. It is the determinant of the partial derivative matrix $\frac{\partial G}{\partial y}$ which is denoted $det\frac{\partial G}{\partial y} = \frac{\partial(G_1,G_2,\dots,G_n)}{\partial(y_1,y_2,\dots,y_n)}$. This gives us an easy criteria to check on the invertibility of $\frac{\partial G}{\partial y}$. Note that if this Jacobian is nonzero then we may judge the level set $G(x,y) = 0$ is an n-dimensional space since it is in one-one correspondence of some open ball in $\mathbb{R}^n$.

Remark 9.2.5.

You may recall the strange comments in red from my section 6.2. I discussed the rank of various derivative matrices. In this section we put the free variables (x) at the start of the list and the dependent variables (y) at the end, however, this is just a notational choice. In practice if we can select any set of n-variables for $G(z_1, z_2, \dots, z_{m+n}) = 0$ such that $det[G_{i_1}|G_{i_2}|\cdots|G_{i_n}] \neq 0$ then we can solve for $z_{i_1}, \dots, z_{i_n}$ in terms of the remaining variables. Thus, in retrospect, showing full rank of the derivative matrix could justifies the local invertibility of certain mappings.

Example 9.2.6. (using Ex. 6.2.10 from my Chapter 6)

$F : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ where $F(x,y,z) = (x^2+z^2, yz)$

$$\frac{\partial(F_1,F_2)}{\partial(x,y)} = det\begin{bmatrix} \partial F_1/\partial x & \partial F_1/\partial y \\ \partial F_2/\partial x & \partial F_2/\partial y \end{bmatrix} = det\begin{bmatrix} 2x & 0 \\ 0 & z \end{bmatrix} = 2xz$$

This means we can solve $F(x,y,z) = (0,0)$ in terms of z at points where $2xz \neq 0$.
We can find $f_1(z)$ and $f_2(z)$ such that
$F(f_1(z), f_2(z), z) = 0$. (I see $f_1(z) = -z^2$, $f_2(z) = 0$).

$\boxed{F = 0 \text{ gives curve } \vec{r}(z) = \langle -z^2, 0, z\rangle}$

Example 9.2.7.

Same function, different choice of free parameter,

$$\frac{\partial(F_1,F_2)}{\partial(y,z)} = det\begin{bmatrix} F_{1y} & F_{1z} \\ F_{2y} & F_{2z} \end{bmatrix} = det\begin{bmatrix} 0 & 2z \\ z & y \end{bmatrix} = -2z^2$$

We can solve $F = 0$ where $z \neq 0$ in terms of the remaing variable x; find $y = f_1(x)$ and $z = f_2(x)$ such that $F(x, f_1(x), f_2(x)) = (0,0)$. A little thinking reveals $f_1(x) = 0$ and $f_2(x) = -x^2$ hence we've found $F = 0$ has sol^{n} $\underbrace{\vec{r}(x) = \langle x, 0, -x^2\rangle}_{\text{parametrization of sol}^{n}}$.

9.3 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".

Example 9.3.1. Suppose $\overbrace{xyz + 2x^2z + 3xz^2 - 1}^{G} = 0$. Calculate $\left(\frac{\partial z}{\partial x}\right)_y$ or $\left(\frac{\partial z}{\partial y}\right)_x$. Here the notation $\left(\frac{\partial z}{\partial x}\right)_y$ announces y is independent from x. We have $G(x,y,z) = 0$ thus $dG = \frac{\partial G}{\partial x}dx + \frac{\partial G}{\partial y}dy + \frac{\partial G}{\partial z}dz$ and $\frac{\partial G}{\partial z} = xy + 2x^2 + 6xz \neq 0$ thus the implicit function Th^{m} applies and we can find $z = h(x,y)$ for some function h at least locally where $G_z \neq 0$

$$0 = G_x\, dx + G_y\, dy + G_z\, dz \quad , \text{ if } G_z \neq 0 \text{ then}$$

$$dz = \frac{-G_x}{G_z}dx - \frac{G_y}{G_z}dy \implies \underline{\left(\frac{\partial z}{\partial x}\right)_y = \frac{-G_x}{G_z} = \frac{-yz - 4xz - 3z^2}{xy + 2x^2 + 6xz}}.$$

Example 9.3.2.

Likewise, $dz = \left(\frac{\partial z}{\partial x}\right)_y dx + \left(\frac{\partial z}{\partial y}\right)_x dy$ hence,

$$\left(\frac{\partial z}{\partial y}\right)_x = \frac{-G_y}{G_z} = \frac{-xz}{xy + 2x^2 + 6xz} = \underline{\frac{-z}{y + 2x + 6z}}.$$

The idea is this: the implicit function Th^{m} gives us the existence of a sol^{n} to $G(x_1, \dots, x_j, \dots, x_n) = 0$ provided $G_j \neq 0$ then partial derivatives w.r.t. the variables $x_1, x_2, \dots, x_{j-1}, x_{j+1}, \dots, x_n$ can be obtained by solving $dG = G_1 dx_1 + \cdots + G_j dx_j + \cdots + G_n dx^n$ for dx_j.

Example 9.3.3. Same G, think of x, z as independent & y dependent,

$$dy = \frac{-G_x}{G_y}dx - \frac{G_z}{G_y}dz$$

$$\underline{\left(\frac{\partial y}{\partial z}\right)_x = \frac{-G_z}{G_y} = \frac{-xy - 2x^2 - 6xz}{xz}}.$$

Example 9.3.4. Let $\begin{cases} 2x + y - 3z - 2u = 0 \\ x + 2y + z + u = 0 \end{cases}$ find $\left(\frac{\partial x}{\partial y}\right)_z$.

Note we have 2 eq^{n}'s and 4 unknowns. The implicit mapping Th^{m} may provide an implicit sol^{n} locally. Notice y, z are free if we are to calculate $\left(\frac{\partial x}{\partial y}\right)_z$ hence we should suspect $(x, u) = h(y, z)$. $(h : \mathbb{R}^2 \longrightarrow \mathbb{R}^2)$

$$\text{I} \quad 2dx + dy - 3dz - 2du = 0$$
$$\text{II} \quad dx + 2dy + dz + du = 0$$

Add I and 2II to obtain (trying to eliminate du)

$$4dx + 5dy - dz = 0$$
$$\therefore \; dx = \frac{-5}{4}dy - \frac{1}{4}dz \implies \boxed{\left(\frac{\partial x}{\partial y}\right)_z = \frac{-5}{4}}$$

Example 9.3.5.

Generally suppose
$$G_1(x_1, \dots, x_m, y_1, \dots, y_n) = 0$$
$$G_2(\vec{x}, \vec{y}) = 0$$
$$\vdots$$
$$G_n(\vec{x}, \vec{y}) = 0$$

We have n-eq^{n}'s and $(m+n)$-unknowns. We can solve for $\vec{y}$ in terms of $\vec{x}$ provided $\frac{\partial G}{\partial \vec{y}}$ is invertible. This is equivalent to saying we can solve for $d\vec{y}$. the eq^{n} $G_{\vec{x}}\,d\vec{x} + G_{\vec{y}}\,d\vec{y}$ we need solve for $G_{\vec{y}} = \frac{\partial G}{\partial \vec{y}}$ to have $\det\left(\frac{\partial G}{\partial \vec{y}}\right) = \frac{\partial(G_1, \dots, G_n)}{\partial(y_1, \dots, y_n)} \neq 0$.

Again we wish to solve for the dy's,

$$dG_1 = \frac{\partial G_1}{\partial x_1}dx_1 + \cdots + \frac{\partial G_1}{\partial x_m}dx_m + \frac{\partial G_1}{\partial y_1}dy_1 + \cdots + \frac{\partial G_1}{\partial y_n}dy_n = 0$$
$$\vdots$$
$$dG_n = \frac{\partial G_n}{\partial x_1}dx_1 + \cdots + \frac{\partial G_n}{\partial x_m}dx_m + \frac{\partial G_n}{\partial y_1}dy_1 + \cdots + \frac{\partial G_n}{\partial y_n}dy_n = 0$$

$$\underbrace{\begin{bmatrix} \partial G_1/\partial y_1 & \partial G_1/\partial y_2 & \cdots & \partial G_1/\partial y_n \\ \partial G_2/\partial y_1 & - & - & \partial G_2/\partial y_n \\ \vdots & & & \vdots \\ \partial G_n/\partial y_1 & - & - & \partial G_n/\partial y_n \end{bmatrix}}\begin{bmatrix} dy_1 \\ dy_2 \\ \vdots \\ dy_n \end{bmatrix} = \begin{bmatrix} -\frac{\partial G_1}{\partial x_1}dx_1 - \cdots - \frac{\partial G_1}{\partial x_m}dx_m \\ \vdots \\ -\frac{\partial G_n}{\partial x_1}dx_1 - \cdots - \frac{\partial G_n}{\partial x_m}dx_m \end{bmatrix}$$

$$\left[\frac{\partial G}{\partial y_1} \middle| \frac{\partial G}{\partial y_2} \middle| \cdots \middle| \frac{\partial G}{\partial y_n}\right]\left[d\vec{y}\right] = \left[-\frac{\partial G}{\partial \vec{x}}d\vec{x}\right]$$

Example 9.3.6. I'll solve for dy_1

$$dy_1 = \frac{\det\left[-\frac{\partial G}{\partial \vec{x}} d\vec{x} \,\middle|\, \frac{\partial G}{\partial y_2} \,\middle|\, \cdots \,\middle|\, \frac{\partial G}{\partial y_n} \right]}{\det\left[\frac{\partial G}{\partial \vec{y}} \right]}$$

$$dy_2 = \frac{\det\left[\frac{\partial G}{\partial y_1} \,\middle|\, -\frac{\partial G}{\partial \vec{x}} d\vec{x} \,\middle|\, \frac{\partial G}{\partial y_3} \,\middle|\, \cdots \,\middle|\, \frac{\partial G}{\partial y_n} \right]}{\det\left[\frac{\partial G}{\partial \vec{y}} \right]}$$

etc... (this notation is unfolded on the previous page, see Ex. 9.3.5.)

Example 9.3.7.

Chapter 14

variational calculus

14.1 history

The problem of variational calculus is almost as old as modern calculus. Variational calculus seeks to answer questions such as:

Remark 14.1.1.

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. Hopefully we can derive the coriolis force in this manner. I also plan to include a problem or two about Maxwell's equations from the variational viewpoint. There must be at least a dozen different ways to phrase Maxwell's equations, one reason I revisit them is to give you a concrete example as to the fact that physics has many formulations.

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.

14.2 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:

$$\begin{aligned}
\frac{\partial J(\alpha)}{\partial \alpha} &= \frac{\partial}{\partial \alpha}\left[\int_{x_1}^{x_2} f(y(x, \alpha), y(x, \alpha)', x)\, dx \right] \\
&= \int_{x_1}^{x_2} \left(\frac{\partial f}{\partial y}\frac{\partial y}{\partial \alpha} + \frac{\partial f}{\partial y'}\frac{\partial y'}{\partial \alpha} + \frac{\partial f}{\partial x}\frac{\partial x}{\partial \alpha} \right) dx \\
&= \int_{x_1}^{x_2} \left(\frac{\partial f}{\partial y}\eta + \frac{\partial f}{\partial y'}\frac{d\eta}{dx} \right) dx \qquad (14.1)
\end{aligned}$$

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 14.1 in view of the product rule above,

$$\begin{aligned}
\frac{\partial J(\alpha)}{\partial \alpha} &= \int_{x_1}^{x_2} \left(\frac{\partial f}{\partial y}\eta + \frac{d}{dx}\left[\frac{\partial f}{\partial y'}\eta\right] - \frac{d}{dx}\left[\frac{\partial f}{\partial y'}\right]\eta \right) dx \\
&= \frac{\partial f}{\partial y'}\eta\bigg|_{x_1}^{x_2} + \int_{x_1}^{x_2} \left(\frac{\partial f}{\partial y}\eta - \frac{d}{dx}\left[\frac{\partial f}{\partial y'}\right]\eta \right) dx \\
&= \int_{x_1}^{x_2} \left(\frac{\partial f}{\partial y} - \frac{d}{dx}\left[\frac{\partial f}{\partial y'}\right] \right)\eta \, dx
\end{aligned} \tag{14.2}$$

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 14.2 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:**

$$\boxed{\frac{\partial f}{\partial y} - \frac{d}{dx}\left[\frac{\partial f}{\partial y'}\right] = 0 \quad \text{Euler-Lagrange Equations for} \quad J[y] = \int_{x_1}^{x_2} f(y, y', x)\, dx}$$

14.3 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,

$$\begin{aligned}
\delta J &= \delta\left[\int_{x_1}^{x_2} f(y,y',x)\,dx\right] \\
&= \left[\int_{x_1}^{x_2} \delta f(y,y',x)\,dx\right] \\
&= \int_{x_1}^{x_2} \left(\frac{\partial f}{\partial y}\delta y + \frac{\partial f}{\partial y'}\delta\left(\tfrac{dy}{dx}\right) + \frac{\partial f}{\partial x}\delta x\right) dx \\
&= \int_{x_1}^{x_2} \left(\frac{\partial f}{\partial y}\delta y + \frac{\partial f}{\partial y'}\frac{d}{dx}(\delta y)\right) dx \\
&= \frac{\partial f}{\partial y'}\delta y\Big|_{x_1}^{x_2} + \int_{x_1}^{x_2}\left(\frac{\partial f}{\partial y} - \frac{d}{dx}\left[\frac{\partial f}{\partial y'}\right]\right)\delta y\,dx
\end{aligned} \tag{14.3}$$

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 14.3.1.

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.

14.4 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.

14.4.1 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,

$$J[x] = \int_{y_1}^{y_2} \sqrt{1 + (x')^2}\, dy$$

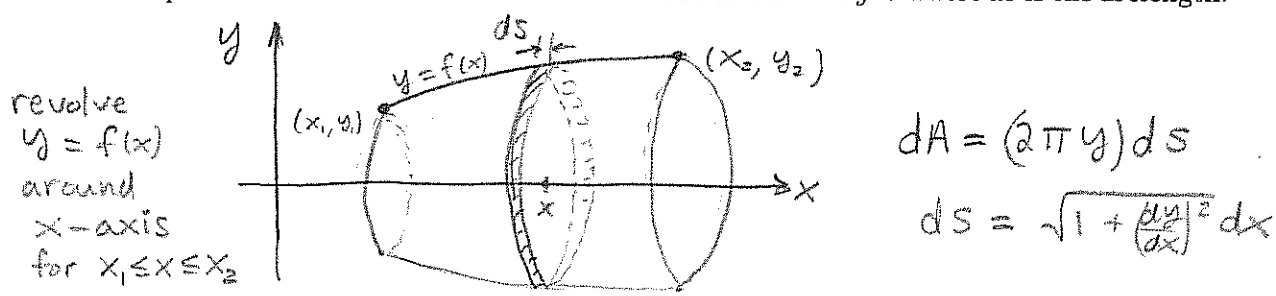
where $x' = dx/dy$ and the Euler Lagrange equations would yield the solution

$$x = x_1 + \frac{x_2 - x_1}{y_2 - y_1}(y - y_1).$$

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.

14.4.2 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.



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).

14.4.3 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.

$$t = \int \frac{ds}{v} = \int \frac{\sqrt{dx^2+dy^2}}{\sqrt{2g(x-x_1)}}$$

then we note by 6.26 $\begin{cases} x = a(1-\cos\theta) \\ y = a(\theta - \sin\theta) \end{cases}$ $\begin{array}{l} dx = a\sin\theta\, d\theta \\ dy = a(1-\cos\theta)d\theta \end{array}$

$$= \int \left(\frac{1+y'^2}{2g(x-x_1)} \right)^{1/2} dx$$

$$y' = \frac{dy}{dx} = \frac{a(1-\cos\theta)d\theta}{a\sin\theta\, d\theta} = \frac{1}{\sin\theta} - \frac{1}{\tan\theta} = y'$$

$$= \int \left(\frac{1 + \left(\frac{1}{\sin\theta} - \frac{1}{\tan\theta}\right)^2}{2g\left(a(1-\cos\theta) - a(1-\cos\theta_1)\right)} \right)^{1/2} a\sin\theta\, d\theta$$

$$= \int \left(\frac{\sin^2\theta + \left(\frac{\sin\theta}{\sin\theta} - \frac{\sin\theta}{\tan\theta}\right)^2}{2g\left(a(1-\cos\theta) - a(1-\cos\theta_1)\right)} \right)^{1/2} a\, d\theta$$

$$= \int \left(\frac{\left(\sin^2\theta + (1-\cos\theta)^2\right) a^2}{2g\left(a(1-\cos\theta) - a(1-\cos\theta_1)\right)} \right)^{1/2} d\theta$$

$$= \int \sqrt{\frac{a}{g}} \left(\frac{\sin^2\theta + 1 - 2\cos\theta + \cos^2\theta}{2(\cos\theta_1 - \cos\theta)} \right)^{1/2} d\theta$$

$$= \int \sqrt{\frac{a}{g}} \left(\frac{2 - 2\cos\theta}{2(\cos\theta_1 - \cos\theta)} \right)^{1/2} d\theta$$

$$= \int \sqrt{\frac{a}{g}} \left(\frac{1 - \cos\theta}{\cos\theta_1 - \cos\theta} \right)^{1/2} d\theta$$

$$= \int \sqrt{\frac{a}{g}} \left(\frac{2\sin^2(\theta/2)}{\cos\theta_1 - 1 + 2\sin^2(\theta/2)} \right)^{1/2} d\theta$$

$$= \int \sqrt{\frac{a}{g}} \left(\frac{2\sin^2(\theta/2)}{\tilde{\alpha} + 2\sin^2(\theta/2)} \right)^{1/2} d\theta$$

$\tilde{\alpha} = \cos\theta_1 - 1$

let $u = \cos(\theta/2)$

$du = -\frac{1}{2}\sin(\theta/2)\, d\theta$

$$= \int \sqrt{\frac{a}{g}} \left(\frac{\sin(\theta/2)\, d\theta}{(\alpha^2 - \cos^2(\theta/2))^{1/2}} \right)$$

$\alpha^2 = \frac{\cos\theta_1 + 1}{2}$

$$= \int \sqrt{\frac{a}{g}} \left(\frac{-2\, du}{(\alpha^2 - u^2)^{1/2}} \right) = \sqrt{\frac{a}{g}}(-2) \int \frac{du}{(\alpha^2 - u^2)^{1/2}} = \left\{ \sin^{-1}\left(\frac{u}{\alpha}\right) \right\} \sqrt{\frac{a}{g}}(-2)$$

$$= \sqrt{\frac{a}{g}}(-2)\left[\sin^{-1}\left(\frac{\cos\theta_2/2}{\alpha}\right) - \sin^{-1}\left(\frac{\cos\theta_1/2}{\alpha}\right) \right]$$

evaluating from $\theta_1 \longrightarrow \theta_2$

$\theta_2 = \pi$

$$= \sqrt{\frac{a}{g}}(-2)\left[\sin^{-1}(0) - \sin^{-1}\left(\frac{\cos\theta_1/2}{\alpha}\right) \right]$$

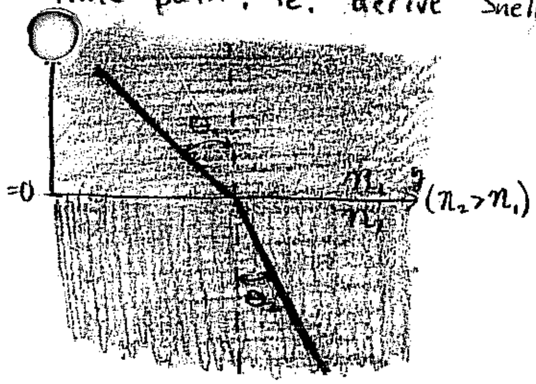
$$= \left(\sqrt{\frac{a}{g}}\right)(2)\left[\sin^{-1}\left(\frac{\cos\theta_1/2}{\sqrt{\frac{\cos\theta_1+1}{2}}} \right) \right] = \sqrt{\frac{a}{g}}(2)\left\{ \sin^{-1}\left(\frac{\cos\theta_1/2}{\cos\theta_1/2} \right) \right\}$$

as $\cos(\theta_1/2) = \sqrt{\frac{\cos\theta_1 + 1}{2}}$

$$= \sqrt{\frac{a}{g}}(2)\left(\sin^{-1}(1)\right) = \sqrt{\frac{a}{g}}(2)\frac{\pi}{2} = \boxed{\pi\sqrt{\frac{a}{g}} = t}$$

SNELL'S LAW DERIVED VIA VARIATIONAL CALCULUS

Consider light passing from medium with index of refraction n_1 into another medium of index of refraction n_2. Use Fermat's principle to find least time path, ie. derive Snell's Law $n_1 \sin\theta_1 = n_2 \sin\theta_2$.



$n = c/v$ thus $v_1 = c/n_1$

$v_2 = c/n_2$

c = speed of light in vaccum

$$t = \int \frac{ds}{v(x)} = \int \frac{dx^2+dy^2}{v(x)} = \int \frac{\sqrt{dx^2+dy^2}}{v(x)} = \int \frac{\sqrt{1+y'^2}}{v(x)}\,dx \qquad y' = \frac{dy}{dx}$$

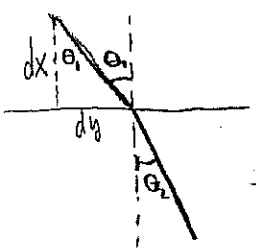
the function $f = \dfrac{\sqrt{1+y'^2}}{v}$ then $\dfrac{\partial f}{\partial y} = 0$ and $\dfrac{\partial f}{\partial y'} = \dfrac{y'}{v\sqrt{1+y'^2}}$

Euler-Lagrange yields minimum functional with $\dfrac{d}{dx}\left(\dfrac{\partial f}{\partial y'}\right) - \dfrac{\partial f}{\partial y} = 0$

$$\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) = 0 \quad \Rightarrow \quad \frac{\partial f}{\partial y'} = \text{const. with respect to } x = a = \frac{y'}{v\sqrt{1+y'^2}}$$

$$a = \frac{y'}{v\sqrt{1+y'^2}} = \frac{-\tan\theta}{v\sqrt{1+\tan^2\theta}} = \frac{-\tan\theta}{v\sqrt{\sec^2\theta}} = \frac{-\sin\theta}{\cos\theta}\left(\frac{1}{\frac{v}{\cos\theta}}\right) = -\sin\theta$$

$$-a = \frac{\sin\theta}{v} = \frac{n\sin\theta}{c}$$



$\Rightarrow \tan\theta = \dfrac{dx}{-dy}$

$\Rightarrow \tan\theta = \dfrac{-dx}{dy}$

$\Rightarrow$ Const $= n\sin\theta$

and n could well have been a function of x. That is $n = n(x)$ which could be used to desribe the gradual bending of starlight by the differing types of air at different altitudes

In general then $\boxed{n(x)\sin\theta = \text{const.}}$

specifically if $n(x) = \begin{cases} n_1 & \text{for } x \geq 0 \\ n_2 & \text{for } x < 0 \end{cases}$ then $\boxed{n_1 \sin\theta_1 = n_2 \sin\theta_2}$

14.5 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,

$$\begin{aligned}
\frac{\partial J(\alpha)}{\partial \alpha} &= \frac{\partial}{\partial \alpha}\left[\int_{t_1}^{t_2} f(y_i(t, \alpha), \dot{y}_i(t, \alpha), t)\, dt \right] \\
&= \int_{t_1}^{t_2} \sum_{j=1}^{n} \left(\frac{\partial f}{\partial y_j}\frac{\partial y_j}{\partial \alpha} + \frac{\partial f}{\partial \dot{y}_j}\frac{\partial \dot{y}_j}{\partial \alpha} \right) dt \\
&= \int_{t_1}^{t_2} \sum_{j=1}^{n} \left(\frac{\partial f}{\partial y_j}\eta_j + \frac{\partial f}{\partial \dot{y}_j}\frac{d\eta_j}{dt} \right) dt \qquad (14.4) \\
&= \sum_{j=1}^{n} \frac{\partial f}{\partial \dot{y}_j}\eta \bigg|_{t_1}^{t_2} + \int_{t_1}^{t_2} \sum_{j=1}^{n} \left(\frac{\partial f}{\partial y_j} - \frac{d}{dt}\frac{\partial f}{\partial \dot{y}_j} \right) \eta_j\, dt
\end{aligned}$$

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.

14.5.1 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.

14.5.2 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 $0 \le t \le 1$. These are the parametric equations for the line segment from the origin to q.

14.6 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 14.6.1.

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.

14.7 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.

14.7.1 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:

$$S = \int (R^2\dot{\theta}^2 + \dot{z}^2)\, dt$$

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 $0 \leq t \leq 1$. 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 $z_1 \leq z \leq z_2$

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 $0 \leq t \leq 1$.

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.

14.7.2 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:

$$S = \int (\underbrace{R^2 \sin^2(\theta)\dot{\phi}^2 + R^2\dot{\theta}^2}_{f(\theta,\phi,\dot{\theta},\dot{\phi})}) \, dt$$

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 I will give you points. I have some solutions but nothing looks too pretty.

Remark 14.7.1.

I'd like to add a few more examples here, but time is up. There are a few more examples in homework. In particular, the homework has the geodesic problem set-up in a more tractable manner. It's easier to solve the geodesic problem if we use one of the coordinates on the sphere as the parameter for calculation of arclength. I should have anticipated this in view of the examples I've already given. The parametric equations for a geodesic will be more general, for example in the case of the plane we found horizontal and vertical lines at once whereas one or the other is lost if x or y is taken as the parameter, and hence harder to solve.

Kinetic Energy In Other Coordinates

Basically, can just divide metric by dt^2. Or we can derive these through chain/product rules applied to coordinate transformations.

POLAR COORDINATES:

$$x = r\cos\theta$$
$$y = r\sin\theta$$

Supposing x, y & r, θ are all functions of time t with $\dot{x}, \dot{y}, \dot{r}, \dot{\theta}$ denoting their respective t-derivatives we find they must satisfy the following relations,

$$\dot{x} = \dot{r}\cos\theta - r\sin\theta\,\dot{\theta} \longrightarrow \dot{x}^2 = \dot{r}^2\cos^2\theta - 2r\dot{r}\sin\theta\cos\theta\,\dot{\theta} + r^2\sin^2\theta\,\dot{\theta}^2$$
$$\dot{y} = \dot{r}\sin\theta + r\cos\theta\,\dot{\theta} \longrightarrow \dot{y}^2 = \dot{r}^2\sin^2\theta + 2r\dot{r}\sin\theta\cos\theta\,\dot{\theta} + r^2\cos^2\theta\,\dot{\theta}^2$$

Note the cross-terms cancel once we add, we find the kinetic energy in polar coordinates from $T = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2)$,

$$\boxed{T = \frac{1}{2}m\left(\dot{r}^2 + r^2\dot{\theta}^2\right)} \quad \text{or} \quad \frac{1}{2}m\left(\dot{r}^2 + r^2\dot{\phi}^2\right)$$

(notation sometimes altered)

Spherical Coordinates: $r = \sqrt{x^2+y^2+z^2}$, ϕ = polar angle, θ = azimuthal angle

$$\underbrace{\boxed{T = \frac{1}{2}m\left(\dot{r}^2 + r^2\dot{\theta}^2 + r^2\sin^2\theta\,\dot{\varphi}^2\right)}}$$

follows from differentiating $x = r\cos\phi\sin\theta$, $y = r\sin\phi\sin\theta$, $z = r\cos\theta$ and substituting into $T = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2)$.

Likewise, CYLINDRICAL COORDINATES

$$\boxed{T = \frac{1}{2}m\left(\dot{r}^2 + r^2\dot{\varphi}^2 + \dot{z}^2\right)}$$

Supposing that
$$\left.\begin{aligned} x &= r\cos\varphi \\ y &= r\sin\varphi \end{aligned}\right\} \text{ makes } \varphi \text{ a "polar" angle.}$$
$$z = z$$

14.8 Lagrangian mechanics

14.8.1 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,

$$m\ddot{x} = F_x \qquad m\ddot{y} = F_y \qquad m\ddot{z} = F_z$$

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.

14.8.2 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 14.5.1). 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[1] 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.

Definition 14.8.1. 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.

[1] 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!

14.8.3 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 14.8.2. 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:

$$\frac{d}{dt}\bigg[m\dot{x}\bigg] = 0 \qquad \frac{d}{dt}\bigg[m\dot{y}\bigg] = 0 \qquad \frac{d}{dt}\bigg[m\dot{z}\bigg] = \frac{\partial}{\partial z}(-mgz) = -mg.$$

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 14.8.3. 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

$$L = \frac{1}{2}ml^2\dot{\theta}^2 + mgl\cos(\theta)$$

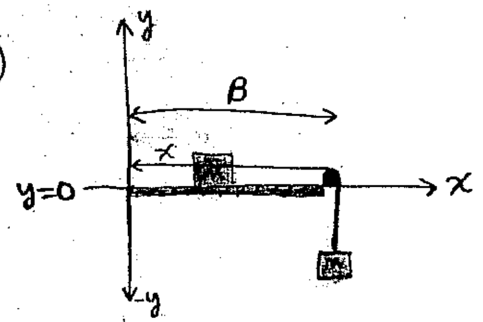
Then, the Euler-Lagrange equation in θ is simply:

$$\frac{d}{dt}\left[\frac{\partial L}{\partial \dot{\theta}}\right] = \frac{\partial L}{\partial \theta} \qquad \Rightarrow \qquad \frac{d}{dt}(ml^2\dot{\theta}) = -mgl\sin(\theta) \qquad \Rightarrow \qquad \ddot{\theta} + \frac{g}{l}\sin(\theta) = 0.$$

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 $\omega = \sqrt{g/l}$

Ex TWO BLOCKS OF MASS M FALL OR SLIDE AS PICTURED W/O FRICTION

(a)



$(\beta - x) - y = l$ where β, l are just constants

$-x - y = l - \beta \Rightarrow -\dot{x} - \dot{y} = 0 \Rightarrow \dot{x} = -\dot{y}$

$$T = \frac{1}{2}M\dot{x}^2 + \frac{1}{2}M\dot{y}^2 = \frac{1}{2}M\dot{y}^2 + \frac{1}{2}M\dot{y}^2 = M\dot{y}^2$$

$$U = Mgy$$

$$L = T - U = M\dot{y}^2 - Mgy = 0$$

$$\frac{\partial L}{\partial \dot{y}} = 2M\dot{y} \qquad \frac{\partial L}{\partial y} = -Mg$$

$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{y}}\right) - \frac{\partial L}{\partial y} = 2M\ddot{y} + Mg = 0 \Rightarrow \ddot{y} = \frac{-g}{2}$, this says the acceleration in y is constant thus integrating twice we obtain the standard form constant acceleration formulae we know and love from freshman physics,

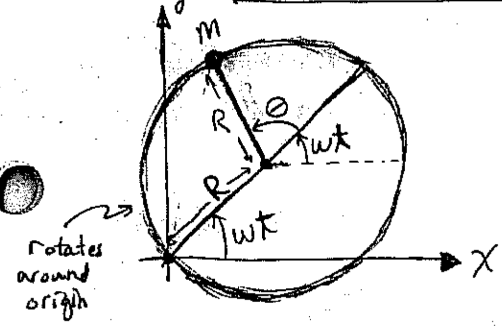
$\left(\begin{array}{l}\dot{y}_0 = \text{intial } V_y \\ y_0 = \text{intial position}\end{array}\right)$

$$\int \ddot{y}\,dt = \int -\frac{g}{2}\,dt \Rightarrow \dot{y} = -\frac{g}{2}(t) + \dot{y}_0$$

$$\int\left(\dot{y}_0 - \frac{g}{2}t\right)dt = y_0 - \frac{gt^2}{4} + \dot{y}_0 t, \quad \boxed{y(t) = y_0 + \dot{y}_0 t - \frac{g}{4}t^2}$$

(in general we would have $y(t) = c_1 + c_2 t - \frac{g}{4}t^2$ but I applied the intial conditions)

Example: Pendulum mounted on center of wheel rotating around point on rim.



$$x = R\cos(\omega t) + R\cos(\omega t + \theta)$$
$$y = R\sin(\omega t) + R\sin(\omega t + \theta)$$
$$\dot{x} = -R\omega\sin(\omega t) - R(\omega + \dot{\theta})\sin(\omega t + \theta)$$
$$\dot{y} = R\omega\cos(\omega t) + R(\omega + \dot{\theta})\cos(\omega t + \theta)$$

$$\dot{x}^2 = \left(-R\omega\sin(\omega t) - R(\omega + \dot{\theta})\sin(\omega t + \theta)\right)\left(-R\omega\sin(\omega t) - R(\omega + \dot{\theta})\sin(\omega t + \theta)\right), \quad \text{Let } \beta = \omega t.$$
$$= R^2\omega^2\sin^2(\beta) + 2R(\omega + \dot{\theta})\sin(\beta + \theta)R\omega\sin(\beta) + R^2(\omega + \dot{\theta})^2\sin^2(\omega t + \theta)$$

$$\dot{y}^2 = \left(R^2\omega^2\cos^2(\beta) + 2R^2\omega(\omega + \dot{\theta})\cos(\beta)\cos(\beta + \theta) + R^2(\omega + \dot{\theta})^2\cos^2(\omega t + \theta)\right)$$

$$T = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) = \frac{1}{2}m\left\{R^2\omega^2(\sin^2\beta + \cos^2\beta) + 2R^2\omega(\omega + \dot{\theta})\left[\sin(\beta)\sin(\beta + \theta) + \cos(\beta)\cos(\beta + \theta)\right]\right.$$
$$\left. + R^2(\omega + \dot{\theta})^2\left[\sin^2(\beta + \theta) + \cos^2(\beta + \theta)\right]\right\}$$

$$\sin(\beta)\sin(\beta + \theta) + \cos(\beta)\cos(\beta + \theta) = \sin(\beta)\left[\sin\beta\cos\theta + \cos\beta\sin\theta\right] + \cos\beta\left[\cos\beta\cos\theta - \sin\theta\sin\beta\right]$$
$$= \cos\theta(\sin^2\beta + \cos^2\beta) + \sin\beta\cos\beta\sin\theta - \cos\beta\sin\theta\sin\beta$$
$$= \cos\theta$$

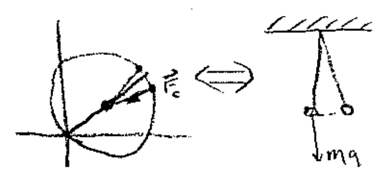
$$T = \frac{1}{2}m\left(R^2\omega^2 + 2R^2\omega(\omega + \dot{\theta})\cos\theta + R^2(\omega + \dot{\theta})^2\right)$$

$L = T - U = T.$ as no U was given.

$$\frac{\partial L}{\partial \dot{\theta}} = mR^2\omega\cos\theta + mR^2(\omega + \dot{\theta})$$
$$\frac{\partial L}{\partial \theta} = -mR^2\omega(\omega + \dot{\theta})\sin\theta$$

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\theta}}\right) - \frac{\partial L}{\partial \theta} = \frac{d}{dt}\left(mR^2\omega\cos\theta + mR^2(\omega + \dot{\theta})\right) + mR^2\omega(\omega + \dot{\theta})\sin\theta$$
$$= -mR^2\omega\dot{\theta}\sin\theta + mR^2\ddot{\theta} + mR^2\omega\dot{\theta}\sin\theta + mR^2\omega^2\sin\theta$$
$$= mR^2\ddot{\theta} + mR^2\omega^2\sin\theta = 0$$
$$\Rightarrow \boxed{\ddot{\theta} + \omega^2\sin\theta = 0} \quad (*)$$

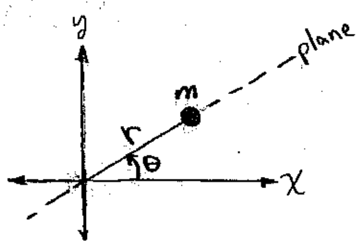
now if θ is small $\sin\theta \approx \theta \Rightarrow (*) \longrightarrow \ddot{\theta} + \omega^2\theta = 0$
which is the equation of a pendulum for small angular dispacements
WITH THE NORMAL PENDULUM $\omega = -\frac{g}{l}$. Our solution reduces to the classic pendulum equation provided θ is small, that is to say the mass is "about" the end of the diameter



Example: moving plane, mass slides on moving plane.

A PARTICLE OF MASS m RESTS ON A SMOOTH PLANE. THE PLANE IS RAISED TO AN INCLINATION ANGLE θ AT A CONSTANT RATE α.

$$\dot{\theta} = \alpha \Rightarrow \theta = \alpha t + \text{const}, \quad \theta(0) = 0 \Rightarrow \theta(t) = \theta = \alpha t$$



SUPPOSE THE PLANE TILTS ABOUT THE z AXIS. THE TRANSFORMATIONS ARE

$$\begin{Bmatrix} x = r\cos\theta \\ y = r\sin\theta \end{Bmatrix} \quad \text{and as } \theta = \alpha t \Rightarrow \begin{Bmatrix} x = r\cos\alpha t \\ y = r\sin\alpha t \end{Bmatrix}$$

$$\begin{Bmatrix} \dot{x} = \dot{r}\cos\alpha t - r\alpha\sin\alpha t \\ \dot{y} = \dot{r}\sin\alpha t + r\alpha\cos\alpha t \end{Bmatrix}$$

$$T = \tfrac{1}{2}m(\dot{x}^2 + \dot{y}^2) = \tfrac{1}{2}m(\dot{r}\cos\alpha t - r\alpha\sin\alpha t)^2 + \tfrac{1}{2}m(\dot{r}\sin\alpha t + r\alpha\cos\alpha t)^2$$

$$= \tfrac{1}{2}m\left(\dot{r}^2\cos^2\alpha t \cancel{- 2r\alpha\sin\alpha t\,\dot{r}\cos\alpha t} + r^2\alpha^2\sin^2\alpha t\right)$$

$$+ \tfrac{1}{2}m\left(\dot{r}^2\sin^2\alpha t \cancel{+ 2\dot{r}\sin\alpha t\, r\alpha\cos\alpha t} + r^2\alpha^2\cos^2\alpha t\right)$$

$$= \tfrac{1}{2}m\left(\dot{r}^2(\cos^2\alpha t + \sin^2\alpha t) + r^2\alpha^2(\sin^2\alpha t + \cos^2\alpha t)\right) = \boxed{\tfrac{1}{2}m(\dot{r}^2 + r^2\alpha^2) = T}$$

For simplicity I choose gravity to act in the y direction in the usual way so,

$$U = mgy = \boxed{mgr\sin\alpha t = U}$$

$$L = T - U = \tfrac{1}{2}m(\dot{r}^2 + r^2\alpha^2) - mgr\sin\alpha t \qquad \text{(LANGRANGIAN IN } r\text{)}$$

$$\frac{\partial L}{\partial \dot{r}} = m\dot{r} \qquad \frac{\partial L}{\partial r} = mr\alpha^2 + mg\sin\alpha t$$

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{r}}\right) - \frac{\partial L}{\partial r} = \frac{d}{dt}(m\dot{r}) - mr\alpha^2 - mg\sin\alpha t = 0$$

ELIMINATING m we have the second order differential equation with constant coefficients and a non-homogeneous term to solve.

Moving Plate Example Continued,

$$\boxed{\ddot{r} - r\alpha^2 = g\sin\alpha t}$$, first find r_H.

$\ddot{r}_H - r_H\alpha^2 = 0 \Rightarrow z^2 - \alpha^2 = 0 \Rightarrow z = \pm\alpha \Rightarrow$ sol^n by the method of constant coefficients, based on writing the characteristic equation (in z here) is just $r_H = c_1 e^{\alpha t} + c_2 e^{-\alpha t} = c_3\cosh(\alpha t) + c_4\sinh(\alpha t)$

Then suppose $r_p = a\cos\alpha t + b\sin\alpha t$ and substitute into DE.

$$\ddot{r}_p - r_p\alpha^2 = -a\alpha^2\cos\alpha t - b\alpha^2\sin\alpha t - \alpha a\cos\alpha t - \alpha b\sin\alpha t = g\sin\alpha t$$

$$\cos\alpha t(-2a\alpha^2) + \sin\alpha t(2b\alpha^2) = \sin\alpha t(g)$$

$$\Rightarrow -2a\alpha^2 = 0 \quad , \quad 2b\alpha^2 = g$$

$$\Rightarrow a = 0 \quad , \quad b = \frac{g}{2\alpha^2}$$

Then the general Solution is

$$r = r_H + r_P = C_3\cosh(\alpha t) + C_4\sinh(\alpha t) + \frac{g}{2\alpha^2}\sin\alpha t$$

Let r_0 be the intial position for the mass then,

$$r_0 = C_3 + C_4\sinh(0) + \frac{g}{2\alpha^2}\sin(0) = C_3$$

Then as the particle was intially at rest we know $\dot{r} = 0$.

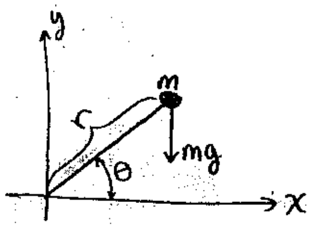
$$\dot{r} = \alpha r_0\sinh(\alpha t) + \alpha C_4\cosh(\alpha t) + \frac{g}{2\alpha}\cos(\alpha t)$$

$$\dot{r}(0) = \alpha C_4 + \frac{g}{2\alpha} = 0 \Rightarrow C_4 = \frac{-g}{2\alpha^2}$$

Finally

$$\boxed{r(t) = r_0\cosh(\alpha t) + \frac{g}{2\alpha^2}(\sin\alpha t - \sinh\alpha t)}$$

Gravity & A general radial force analyzed:



$$\vec{F} = -mg\hat{j} - Ar^{\alpha-1}\hat{r} = \vec{F}_g + \vec{F}_r$$

$$-\nabla U = -\nabla(U_g + U_r) = \vec{F}_g + \vec{F}_r$$

we may superpose the two potential energy functions (radial and gravitational) to find the total potential energy-

$U_r = \dfrac{Ar^\alpha}{\alpha}$ from problem 7.4 as the radial force is identical here.

$U_g = mgy = mgr\sin\theta$, using $y = r\sin\theta$ as before.

I choose radial coordinates r, θ.

$$T = \tfrac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2) \quad \text{from (7.4)}$$

$$U = \frac{Ar^\alpha}{\alpha} + mgr\sin\theta$$

$$L = T - U = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2) - \frac{Ar^\alpha}{\alpha} - mgr\sin\theta$$

$$\frac{\partial L}{\partial \dot{r}} = m\dot{r} \qquad \frac{\partial L}{\partial r} = mr\dot{\theta}^2 - Ar^{\alpha-1} - mg\sin\theta$$

$$\frac{\partial L}{\partial \dot{\theta}} = mr^2\dot{\theta} \qquad \frac{\partial L}{\partial \theta} = -mgr\cos\theta$$

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{r}}\right) - \frac{\partial L}{\partial r} = \frac{d}{dt}(m\dot{r}) - mr\dot{\theta}^2 + Ar^{\alpha-1} + mg\sin\theta = \boxed{m\ddot{r} - mr\dot{\theta}^2 + Ar^{\alpha-1} + mg\sin\theta = 0}$$

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\theta}}\right) - \frac{\partial L}{\partial \theta} = \frac{d}{dt}(mr^2\dot{\theta}) + mgr\cos\theta = m\frac{dr^2}{dt}\dot{\theta} + mr^2\ddot{\theta} + mgr\cos\theta = 0$$

$$\Rightarrow 2r\dot{r}\dot{\theta} + r^2\ddot{\theta} + gr\cos\theta = 0$$

$$\Rightarrow \boxed{2\dot{r}\dot{\theta} + r\ddot{\theta} + g\cos\theta = 0}$$

In general we have defined the momenta of a coordinate q as $\frac{\partial T}{\partial \dot{q}}$ here we consider θ. $P_\theta = \frac{\partial T}{\partial \dot{\theta}} = mr^2\dot{\theta}$. This time we may not claim that P_θ is conserved. $\frac{d}{dt}(mr^2\dot{\theta}) = -mgr\cos\theta \Rightarrow$ that P_θ is not the same for all time, there is some change with time thus it is not conserved, this is a consequence of gravity, if this were not the case then we would be able to spin things in vertical planes, one could "swing" little children in complete orbits around the swing set.

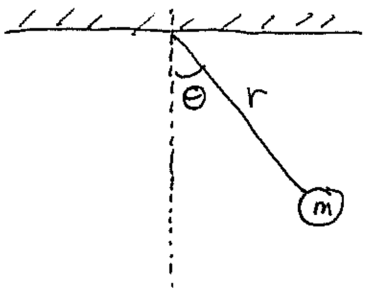
Lagrange Multipliers

The Euler–Lagrange Eq^{n}s are modified to include constraints $f_k(q_1, q_2, .., q_n, t) = 0$ for $k = 1, 2, .., m$ as follows,

$$\boxed{\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right) - \frac{\partial L}{\partial q_j} = \sum_{k=1}^{m} \lambda_k(t) \frac{\partial f_k}{\partial q_j} = Q_j} \quad (*)$$

for $j = 1, 2, \dots, n$ where q_j are generalized coordinates for the system considered.

Example: simple pendulum



$$T = \frac{1}{2} m \left(\dot{r}^2 + r^2 \dot{\theta}^2\right)$$

$$U = -mgr\cos\theta$$

$$\underline{L = \frac{1}{2} m (\dot{r}^2 + r^2\dot{\theta}^2) + mgr\cos\theta}$$

The constraint here is $r = l$ which we can write as $f_1(r, \theta) = r - l$. This constaint is time independent. Using (*) as our guide,

① $$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{r}}\right) - \frac{\partial L}{\partial r} = \lambda_r \frac{\partial f_1}{\partial r} = \lambda_r = Q_r$$

$$\Rightarrow \underline{\frac{d}{dt}(m\dot{r}) - mr\dot{\theta}^2 - mg\cos\theta = \lambda_r} \quad \text{(I)}$$

② $$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\theta}}\right) - \frac{\partial L}{\partial \theta} = \lambda_r \frac{\partial f_1}{\partial \theta} = 0$$

$$\Rightarrow \underline{\frac{d}{dt}(mr^2\dot{\theta}) + mgr\sin\theta = 0} \quad \text{(II)}$$

③ $\underline{r = l}$.

Obviously ③ simplifies things considerably. Note $\dot{r} = 0$ thus

(I) $\longrightarrow$ $-mr\dot{\theta}^2 - mg\cos\theta = \lambda_r$ $\longleftarrow$ centripetal acceleration & more.

(II) $\longrightarrow$ $\frac{d}{dt}(ml^2\dot{\theta}) = -mgl\sin\theta$ $\longrightarrow$ $\underline{\frac{dL}{dt} = -mgl\sin\theta}$

think angular mom. & torque!

Ex] $L_{free} = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2)$

Constraint: constant velocity $f_1 = x - V_1 t$, $f_2 = y - V_2 t$.

$\hookrightarrow \quad L = \frac{1}{2}m(\dot{x}^2+\dot{y}^2) - \lambda_1 f_1 - \lambda_2 f_2$

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right) - \frac{\partial L}{\partial x} = 0 \longrightarrow m\ddot{x} = \lambda_1$$

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{y}}\right) - \frac{\partial L}{\partial y} = 0 \longrightarrow m\ddot{y} = \lambda_2$$

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\lambda}_1}\right) - \frac{\partial L}{\partial \lambda_1} = f_1 = 0 \rightarrow x = v_1 t \longrightarrow \dot{x} = v_1 \rightarrow \ddot{x} = 0$$

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\lambda}_2}\right) - \frac{\partial L}{\partial \lambda_2} = f_2 = 0 \rightarrow y = V_2 t \longrightarrow \dot{y} = V_2 \longrightarrow \ddot{y} = 0$$

Hence the generalized forces λ_1 & λ_2 are both zero.

Ex] $L = \frac{1}{2}m(\dot{x}^2+\dot{y}^2) - \lambda f$ where $f = y - g(x)$.

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right) - \frac{\partial}{\partial x}\lambda f = 0 \rightarrow m\ddot{x} = \lambda\frac{\partial f}{\partial x} = -\lambda\frac{dg}{dx}.$$

$$m\ddot{y} = \lambda\frac{\partial f}{\partial y} = \lambda$$

If we want $y = g(x) = x^2$ then $\frac{dg}{dx} = 2x$

and we need $m\ddot{x} = -2x\lambda$ & $m\ddot{y} = \lambda$ for $y = x^2$.

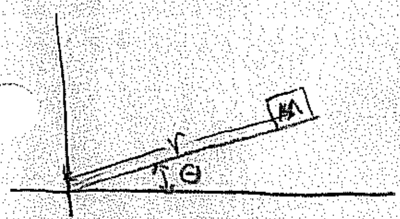
Note, $\dot{y} = 2x\dot{x}$ & $\ddot{y} = 2\dot{x}^2 + 2x\ddot{x}$

$$\Rightarrow m\ddot{y} = m(2\dot{x}^2 + 2x\ddot{x}) = \lambda$$

$$\hookrightarrow m\ddot{x} = -2x\lambda = -2x(m(2\dot{x}^2 + 2\ddot{x}))$$

$$\Rightarrow m\ddot{x} + 4x\ddot{x} = -4mx\dot{x}^2 \ldots$$

Use LAGRANGE Multipliers to find constraint forces, for Moving Plane Example



$x = r\cos\theta$
$y = r\sin\theta$

$$T_{POLAR} = \tfrac{1}{2}m(\dot r^2 + r^2\dot\theta^2) \qquad U = mgr\sin\theta$$

$$L = T - U = \tfrac{1}{2}m(\dot r^2 + r^2\dot\theta^2) - mgr\sin\theta$$

Constraint: $\dot\theta = \alpha \Rightarrow \theta = \alpha t + \text{const.}$ and $\theta(t=0) = 0 \Rightarrow \text{const} = 0 \Rightarrow \theta = \alpha t$
thus $f_1(r,\theta) = \theta - \alpha t = 0.$

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot r}\right) - \frac{\partial L}{\partial r} = \lambda_1 \frac{\partial f_1}{\partial r} = 0 = \frac{d}{dt}(m\dot r) - mr\dot\theta^2 - mg\sin\theta$$

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot\theta}\right) - \frac{\partial L}{\partial \theta} = \lambda_1 \frac{\partial f_1}{\partial \theta} = \lambda_1(1) = \lambda_1 = \frac{d}{dt}\left(mr^2\dot\theta\right) + mgr\cos\theta$$

Three Equations, Three unknowns, r, θ, λ_1

$$\left\{\begin{array}{l} \theta - \alpha t = 0 \\ m\ddot r - mr\dot\theta^2 - mg\sin\theta = 0 \\ \frac{d}{dt}\left(mr^2\dot\theta\right) + mgr\cos\theta = \lambda_1 \end{array}\right\} \quad \begin{array}{l}\text{then as} \\ \theta = \alpha t \\ \dot\theta = \alpha\end{array} \quad \left\{\begin{array}{l} m\ddot r - mr\alpha^2 - mg\sin\alpha t = 0 \\ (*)\ \frac{d}{dt}\left(mr^2\alpha\right) + mgr\cos\alpha t = \lambda_1 \end{array}\right\}$$

Now I solve $\ddot r - \alpha^2 r - g\sin\alpha t = 0$ to find $r(t)$ as before
I will omit the details.

$$r(t) = r_0\cosh(\alpha t) + \frac{g}{2\alpha^2}\left(\sin\alpha t - \sinh\alpha t\right)$$

Then we substitute $r(t)$ into $(*)$ to get

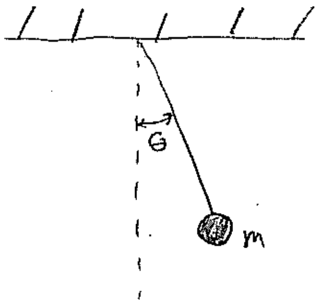
$$\frac{d}{dt}\left(mr^2\alpha\right) + mgr\cos\alpha t = \lambda_1$$

$$m(2r)\dot r\alpha + mgr\cos\alpha t = \lambda_1$$

$$\boxed{mr(2\alpha\dot r + g\cos\alpha t) = \lambda_1} = \sum_k \lambda_k \frac{\partial f_k}{\partial \theta} \equiv Q_{\theta}$$: the generalized force of constraint in θ.

and $\sum \lambda_k \frac{\partial f_k}{\partial r} \equiv Q_r = 0$: generalized constraint force for r.

Hamiltonian for Simple Pendulum



$$\dot{l} = -\alpha$$
$$l = -\alpha t + l_0$$
$$x = l\sin\theta = (-\alpha t + l_0)\sin\theta$$
$$y = -l\cos\theta = -(-\alpha t + l_0)\cos\theta$$
$$\dot{x} = \dot{l}\sin\theta + l\dot{\theta}\cos\theta = -\alpha\sin\theta + (-\alpha t + l_0)\dot{\theta}\cos\theta$$
$$\dot{y} = -\dot{l}\cos\theta - l\dot{\theta}\sin\theta = +\alpha\cos\theta - (-\alpha t + l_0)\dot{\theta}\sin\theta$$

$$L = \frac{1}{2}m\left(\dot{l}^2 + l^2\dot{\theta}^2\right) + mgl\cos\theta = \boxed{\frac{1}{2}m\left(\alpha^2 + l^2\dot{\theta}^2\right) + mgl\cos\theta = L}$$

$$p_\theta = \frac{\partial L}{\partial \dot{\theta}} = ml^2\dot{\theta} \implies \dot{\theta} = p_\theta / ml^2$$

$$H = \dot{\theta}p_\theta - L = \frac{p_\theta^2}{ml^2} - \frac{1}{2}m\left(\alpha^2 + l^2\dot{\theta}^2\right) - mgl\cos\theta$$
$$= \frac{p_\theta^2}{ml^2} - \frac{1}{2}m\left(\alpha^2 + \frac{p_\theta^2}{m^2l^2}\right) - mgl\cos\theta$$

$$\boxed{H(l,\theta,p_\theta) = \frac{p_\theta^2}{2ml^2} - \frac{1}{2}m\alpha^2 - mgl\cos\theta}$$

$$E = T + U = \frac{1}{2}m\left(\alpha^2 + l^2\dot{\theta}^2\right) - mgl\cos\theta$$
$$= \frac{1}{2}m\alpha^2 + \frac{1}{2}\frac{p_\theta^2}{ml^2} - mgl\cos\theta$$

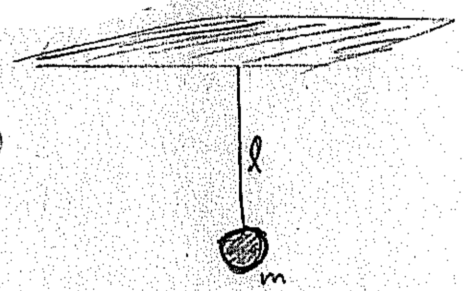
sign of $\frac{1}{2}m\alpha^2$ differs from $H \to E$.

note that the total energy is not the same as the Hamiltonian with this formulation of the problem, because I accounted for the explicit time dependence of the transformation Eq's from the rectangular coord. to the generalized coordinates.

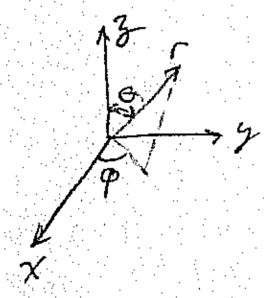
$$\frac{dE}{dt} = \frac{d}{dt}\left(\frac{1}{2}m\alpha^2 + \frac{1}{2}\frac{p_\theta^2}{ml^2} - mg(\alpha t + l_0)\cos\theta\right) \neq 0$$

thus energy is not conserved.

Hamiltonian for Spherical Pendulum



$$x = r\sin\theta\cos\phi$$
$$y = r\sin\theta\sin\phi$$
$$z = r\cos\theta$$



$$T = \frac{1}{2}m\left(\dot{r}^2 + r^2\dot{\theta}^2 + r^2\sin^2\theta\,\dot{\varphi}^2\right)$$
$$U = mgr\cos\theta$$

$$L = T - U = \frac{1}{2}m\left(\dot{r}^2 + r^2\dot{\theta}^2 + r^2\sin^2\theta\,\dot{\varphi}^2\right) - mgr\cos\theta$$

$$\left\{\begin{array}{l} p_r = \dfrac{\partial L}{\partial \dot{r}} = m\dot{r} \\ p_\theta = \dfrac{\partial L}{\partial \dot{\theta}} = mr^2\dot{\theta} \\ p_\varphi = \dfrac{\partial L}{\partial \dot{\varphi}} = mr^2\sin^2\theta\,\dot{\varphi} \end{array}\right\} \Rightarrow \left\{\begin{array}{l} \dot{r} = p_r/m \\ \dot{\theta} = p_\theta/mr^2 \\ \dot{\varphi} = p_\varphi/mr^2\sin^2\theta \end{array}\right\}$$

$$H = \dot{r}p_r + \dot{\theta}p_\theta + \dot{\phi}p_\phi - L$$
$$= \frac{p_r^2}{m} + \frac{p_\theta^2}{mr^2} + \frac{p_\varphi^2}{mr^2\sin^2\theta} - \frac{1}{2}m\left\{\dot{r}^2 + r^2\dot{\theta}^2 + r^2\sin^2\theta\,\dot{\varphi}^2\right\} + mgr\cos\theta$$
$$= \frac{p_r^2}{m} + \frac{p_\theta^2}{mr^2} + \frac{p_\varphi^2}{mr^2\sin^2\theta} - \frac{1}{2}m\left\{\frac{p_r^2}{m^2} + r^2\left(\frac{p_\theta^2}{m^2r^4}\right) + r^2\sin^2\theta\left(\frac{p_\varphi^2}{m^2r^4\sin^4\theta}\right)\right\} + mgr\cos\theta$$

$$\boxed{H = \cancel{\frac{p_r^2}{2m}}^{\,0} + \frac{p_\theta^2}{2mr^2} + \frac{p_\varphi^2}{2mr^2\sin^2\theta} + mgr\cos\theta}$$

Combining p_φ and θ terms, $U_{eff}(p_\phi, \theta) = \dfrac{p_\varphi^2}{2mr^2\sin^2\theta} + mgr\cos\theta$

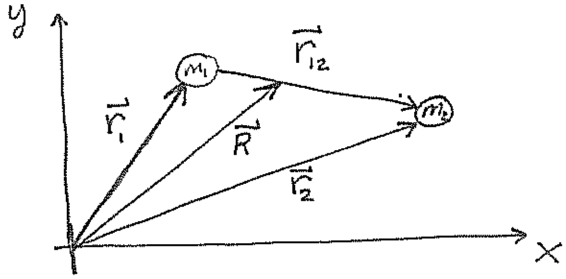
Discuss and Sketch V as a function of θ

(1) $p_\varphi = 0$

(2) several $p_\varphi \neq 0$

(3) Conical pendulum (θ = constant) with reference to $V-\theta$ digram

Central Force Problem



m_1 at $\vec{r}_1$
m_2 at $\vec{r}_2$
$\vec{r}_{12} = \vec{r}_2 - \vec{r}_1 = \vec{r}$

$$\vec{R} = \frac{1}{m_1+m_2}\left(m_1\vec{r}_1 + m_2\vec{r}_2\right)$$

We suppose the only force is an interaction whose strength depends only on the distance $|\vec{r}_{12}| = |\vec{r}_2 - \vec{r}_1|$.
Define $\vec{r} = \vec{r}_2 - \vec{r}_1$, the relative coordinate. We have two fundamental transformations:

$$\vec{r} = \vec{r}_2 - \vec{r}_1 \qquad\qquad \vec{r}_1 = \vec{R} - \frac{m_2}{m_1+m_2}\vec{r}$$

$$\vec{R} = \frac{m_1}{m_1+m_2}\vec{r}_1 + \frac{m_2}{m_1+m_2}\vec{r}_2 \qquad\qquad \vec{r}_2 = \vec{R} + \frac{m_1}{m_1+m_2}\vec{r}$$

We also define velocities of the com (center of mass) $\vec{R}$ and relative coordinate $\vec{r}$,

$$\vec{V} = \frac{d\vec{R}}{dt} \qquad\qquad \vec{V}_1 = \frac{d\vec{r}_1}{dt} = \frac{d\vec{R}}{dt} - \frac{m_2}{m_1+m_2}\frac{d\vec{r}}{dt} = \vec{V} - \frac{m_2}{m_1+m_2}\vec{v}$$

$$\vec{v} = \frac{d\vec{r}}{dt} \qquad\qquad \vec{V}_2 = \frac{d\vec{r}_2}{dt} = \vec{V} + \frac{m_1}{m_1+m_2}\vec{v}$$

We compute the Lagrangian in the com/relative coordinate system as follows,

$$L = \frac{1}{2}m_1\vec{V}_1\cdot\vec{V}_1 + \frac{1}{2}m_2\vec{V}_2\cdot\vec{V}_2 - U(|\vec{r}_2 - \vec{r}_1|)$$

$$= \frac{1}{2}m_1\left(\vec{V} - \frac{m_2}{m_1+m_2}\vec{v}\right)\cdot\left(\vec{V} - \frac{m_2}{m_1+m_2}\vec{v}\right) + \frac{1}{2}m_2\left(\vec{V} + \frac{m_1}{m_1+m_2}\vec{v}\right)\cdot\left(\vec{V} + \frac{m_1}{m_1+m_2}\vec{v}\right) - U(r)$$

$$= \frac{1}{2}(m_1+m_2)\vec{V}\cdot\vec{V} + \frac{1}{2}\frac{m_1m_2^2 + m_2m_1^2}{(m_1+m_2)^2}\vec{v}\cdot\vec{v} - U(r)$$

$$= \frac{1}{2}(m_1+m_2)V^2 + \frac{1}{2}\left(\frac{m_1m_2}{m_1+m_2}\right)v^2 - U(r).$$

Continuing, the Lagrangian in the $\vec{r}, \vec{R}$ coordinates,

$$L = \frac{1}{2}(m_1+m_2)V^2 + \frac{1}{2}\left(\frac{m_1 m_2}{m_1+m_2}\right)v^2 - U(r)$$

$$\Rightarrow \boxed{L = \frac{1}{2}MV^2 + \frac{1}{2}\mu v^2 - U(r)}$$

Where we define:

$M = m_1 + m_2$ the total mass

$\mu = \dfrac{m_1 m_2}{m_1+m_2}$ the reduced mass

A typical application is $m_1 = M_{SUN}$, $m_2 = M_{EARTH}$. In this case $\mu = \dfrac{M_E M_S}{M_E + M_S} \approx \dfrac{M_E M_S}{M_S} = M_E$. It is approximately true to say $\vec{R} = \vec{r}_{sun}$, the com is close to center of sun. However, the math here allows for an <u>exact</u> description of the motion. We need not assume the sun is the com.

<u>Equations of Motion</u>:

$\vec{R}$: Let $\vec{R} = (X, Y, Z)$ then using $L = \frac{1}{2}M(\dot{X}^2+\dot{Y}^2+\dot{Z}^2) + \frac{1}{2}\mu(\underbrace{\dot{x}^2+\dot{y}^2+\dot{z}^2}) - U(r)$

letting $\vec{r} = (x,y,z)$.

$$\left.\begin{aligned} \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{X}}\right) &= \frac{\partial L}{\partial X} \\ \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{Y}}\right) &= \frac{\partial L}{\partial Y} \\ \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{Z}}\right) &= \frac{\partial L}{\partial Z} \end{aligned}\right\} \longrightarrow \underline{M\ddot{X} = M\ddot{Y} = M\ddot{Z} = 0}.$$

$\hookrightarrow$ the center of mass moves with <u>constant velocity</u> motion.

$$\boxed{\vec{R}(t) = (X_o, Y_o, Z_o) + t(V_{xo}, V_{yo}, V_{zo})}$$

<u>Remark</u>: we've proved previously that if $\vec{F}$ is directed along $\vec{r}$ then the resulting motion falls onto a particular plane. Without loss of generality we suppose $z = 0$ and the motion is described by just x & y. Thus,

$$\underline{L = \frac{1}{2}M(\dot{X}^2+\dot{Y}^2+\dot{Z}^2) + \frac{1}{2}\mu(\dot{x}^2+\dot{y}^2) - U(r)}$$

Continuing: (we choose $z=0$ for clarity of notation)
We introduce polar coordinates for the relative coordinate,

$$x = r\cos\theta$$
$$y = r\sin\theta$$

Then,

$$L = \frac{1}{2}M(\dot{X}^2+\dot{Y}^2+\dot{Z}^2) + \frac{1}{2}\mu(\dot{r}^2+r^2\dot{\theta}^2) - U(r)$$

We find

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{r}}\right) = \frac{\partial L}{\partial r} \Rightarrow \underline{\mu\ddot{r} = -\frac{\partial U}{\partial r} + \mu r\dot{\theta}^2}.$$

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\theta}}\right) = \frac{\partial L}{\partial \theta} \Rightarrow \frac{d}{dt}\big(\underbrace{\mu r^2\dot{\theta}}\big) = 0$$

(mechanists call θ an "ignorable" coordinate)

conserved quantity!
(angular momentum $l = \mu r^2\dot{\theta} = \text{const.}$)

The radial eq$^{\underline{n}}$ simplifies now that we know $\dot{\theta} = \frac{l}{\mu r^2}$,

$$\mu\ddot{r} = \frac{-\partial U}{\partial r} + \mu r\left(\frac{l}{\mu r^2}\right)^2$$

$$\Rightarrow \underline{\mu\frac{d^2r}{dt^2} = -\frac{d}{dr}\left[U + \frac{l^2}{2\mu r^2}\right]}.$$

Multiply by $\frac{dr}{dt}$ to obtain,

$$\mu\frac{dr}{dt}\frac{d^2r}{dt^2} + \frac{dr}{dt}\frac{d}{dr}\left[U + \frac{l^2}{2\mu r^2}\right] = 0$$

$$\frac{d}{dt}\Big[\underbrace{\frac{1}{2}\mu\dot{r}^2 + U(r) + \frac{l^2}{2\mu r^2}}\Big] = 0$$

conserved quantity!
(energy $E = \frac{1}{2}\mu\dot{r}^2 + \underbrace{U(r) + \frac{l^2}{2\mu r^2}}$

The effective potential is used to lump fictitious forces into the same category as the real potential U. We can analyze as if V_{eff} is an ordinary potential function. (we ignore θ)

$$\boxed{V_{eff} = U(r) + \frac{l^2}{2\mu r^2}}$$

Equations of Motion for Central Force Problem

$$\left.\begin{aligned} l &= \mu r^2 \dot{\theta} \\ E &= \frac{1}{2}\mu \dot{r}^2 + U(r) + \frac{l^2}{2\mu r^2} \end{aligned}\right\}$$ conserved quantities for central force motion where $F = -\frac{dU}{dr}$

We have two, fairly simple differential equations which are possible to solve once we specify a choice of $U(r)$.

$$\boxed{\begin{aligned} \frac{d\theta}{dt} &= \frac{l}{\mu r^2} \\ \frac{dr}{dt} &= \pm\sqrt{E - U - \frac{l^2}{2\mu r^2}}\sqrt{\frac{2}{\mu}} \end{aligned}}$$

<u>Orbit Equation:</u>

It turns out solving for r as fnct of θ is better replaced with problem of finding $u = \frac{1}{r}$ as fnct. of θ.

Consider, $\mu \ddot{r} - \frac{l}{\mu r^3} + \frac{dU}{dr} = 0$ where $\frac{d\theta}{dt} = \frac{l}{\mu r^2}$

Chain-rule reveals,

$$\frac{dr}{dt} = \frac{dr}{d\theta}\frac{d\theta}{dt} = \frac{dr}{d\theta}\frac{l}{\mu r^2}$$

$$\Rightarrow \frac{d^2 r}{dt^2} = \frac{d}{dt}\left[\frac{l}{\mu r^2}\frac{dr}{d\theta}\right] = \frac{d}{d\theta}\left[\frac{l}{\mu r^2}\frac{dr}{d\theta}\right]\frac{d\theta}{dt} = \underbrace{\frac{l}{\mu r^2}\frac{d}{d\theta}\left[\frac{l}{\mu r^2}\frac{dr}{d\theta}\right]}_{\ddot{r}}$$

Hence,

$$\mu \frac{l}{\mu r^2}\frac{d}{d\theta}\left[\frac{l}{\mu r^2}\frac{dr}{d\theta}\right] - \frac{l}{\mu r^3} + \frac{dU}{dr} = 0$$

$$\frac{l^2}{\mu r^2}\underbrace{\frac{d}{d\theta}\left[\frac{1}{r^2}\frac{dr}{d\theta}\right]} - \frac{l}{\mu r^3} = -\frac{dU}{dr}$$

$$\frac{-l^2}{\mu r^2}\frac{d^2}{d\theta^2}\left(\frac{1}{r}\right) - \frac{l}{\mu r^3} = F(r)$$

$$\Rightarrow \frac{d^2}{d\theta^2}\left(\frac{1}{r}\right) + \frac{1}{r} = \frac{-\mu r^2}{l^2} F(r)$$

Orbit Equation Continued

We found

$$\frac{d^2}{d\theta^2}\left(\frac{1}{r}\right) + \frac{1}{r} = \frac{-\mu r^2}{l^2} F(r)$$

$$\hookrightarrow \boxed{\frac{d^2 u}{d\theta^2} + u = \frac{-\mu}{u^2 l^2} F(1/u) \quad \text{for } u = \frac{1}{r}}$$

orbit equation

Remark: I have several pages of examples based on this equation if you look in my Physics ~~___~~ 412 notes..

Example: Newton's Universal Law of Gravitation.

Suppose $F = \frac{-Gm_1m_2}{r^2}$ write $F = \frac{-k}{r^2}$ where $k = Gm_1m_2$

Note, $F\left(\frac{1}{u}\right) = \frac{-k}{(1/u)^2} = -ku^2$ hence the orbit equation

takes form:

$$\underbrace{\frac{d^2u}{d\theta^2} + u = \frac{k\mu}{l^2}}_{\text{well-known ordinary differential equation.}} \leftarrow \text{constant of motion.}$$

the sol$^{\underline{n}}$ is simply,

$$u = u_o\cos(\theta + \theta_o) + u_p$$

Note $u_p = A \Rightarrow u_p' = 0 \dashrightarrow u_p = k\mu/l^2$. therefore,

$$u = u_o\cos(\theta + \theta_o) + \frac{k\mu}{l^2} = \frac{1}{r}$$

$$\Rightarrow \boxed{r = \frac{1}{u_o\cos(\theta+\theta_o) + k\mu/l^2}}$$

this is the equation of a conic section in polar coordinates. to make it pretty it's nice to use E as a constant in the sol$^{\underline{n}}$. It's a page or three to settle notation. Ask if interested.