Vector Calculus

DeVon Young

Spring 2022

These are my notes for *Vector Calculus, sixth edition*, by Jerrold E. Marsden and Anthony Tromba. If you reference the table of contents, I denote by (MT ($x.y$)) the chapter/section from *Vector Calculus* it involves. I use the same notation for problems and examples. Not all sections follow this, so for ones that do, I will surround them with "quotation marks."

Let me make this clear: Unlike the majority of course notes, this document does not consist of notes I've taken for lectures. While this is intended to roughly follow the curriculum of a course, Purdue's MA 36200 (Topics in Vector Calculus), these notes were not written to follow lectures, and they were not written as I took the course. I began writing these notes to comprehensively prepare for the final exam, rather than as a reference of prior lecture content. However, as a consequence, I can be considerably more forward-sighted in perspective when discussing certain topics and give far more introspection, such as ways I found best to do certain procedures on paper. I imagine this to actually be more universal than my other notes in the sense that the ideas here should be a bit more independent of professor and topic coverage, as well as ordering. Since these were not made in parallel with a course,

You should not use these in favor of going to class/reading your textbook.

While I don't recommend that people use any of my notes in going to class, I emphasize it especially so for this set.

I took notes for when I took Purdue's MA 26100 (Multivariate Calculus), but they're unfinished and hence unreleased. I hope that I finish this as a substitute for those. I used to worry greatly about my legacy at Purdue, which is honestly quite silly, since it's not the objective of a student to leave one behind, and perhaps distracting from more important goals. However, I know my linear algebra notes are shared to people I have never met and may perhaps never meet, and I'm very happy to know that people enjoy it – it really is the magnum opus of my time as an undergraduate, despite my distancing and departure from mathematical studies since then. Let this be my last hurrah. There are many students that take a vector calculus course here, and although there now exists many more resources now than there were before when I took it, I hope this can join the ranks.

I wrote the linear algebra notes the summer after my freshman year, and I write these notes as a final semester senior. Time has changed me, but more importantly, the style and how I write. These notes are for me, first and foremost, and so I apologize if you don't like them;

they're bound to have a different "flavor" than the linear algebra ones. I definitely talk a lot less here.

That's enough. The end is in sight. Let's begin.

Contents

1 "The Geometry of Euclidean Space (MT 1)"

Traditionally, vector calculus texts develop the theory of vectors on $\mathbb{R}^2$ and $\mathbb{R}^3$ before investigating the general case of $\mathbb{R}^n$. While this makes sense for drawing examples, I don't find the examples of vector addition and scalar multiplication truly complex enough to warrant considering this only in $\mathbb{R}^3$ and lower, and then generalizing later. In fact, if you just avoid saying "vector in $\mathbb{R}^2$ or whatever, virtually everything transfers over. So even though the first section is going to be "Vectors in Two- and Three-dimensional space", everything works in higher dimensions unless stated otherwise.

1.1 "Vectors in Two- and Three-Dimensional Space (MT 1.1)"

If you know linear algebra, we are going to define $\mathbb{R}^n$ as a vector space. That is, how to write down points, and then how to add them.

1.1.1 Cartesian coordinates and points

Working in the regular Cartesian coordinate system, the xy-plane, we define a point p on the plane by an ordered pair/tuple of two real numbers: $p = (a, b)$. Then a and b form the Cartesian coordinates of P. a is denoted the x component of p, and similarly with b and y. The x and y axes, which we take to be infinitely long perpendicular lines, are used, along side their corresponding components, to denote the position of the point p on that axis. That is, the point $(2, 3)$ is 2 on the x axis and 3 on the y axis.

This can be generalized to the xyz-space, using a 3 element ordered tuple. In fact, we describe $\mathbb{R}^n$ as an n-tuple of real numbers.

For simplicity, when we describe arbitrary tuples, we use a sequence inspired notation. Instead of $P = (p_1, p_2, \ldots, p_n)$, we can instead write $P = (p_i)$, although we may need to be careful, as we do not include n explicitly in the latter notation. This is also inspired from linear algebra.

1.1.2 Operations in $\mathbb{R}^n$

Now that we have defined the elements, we can now define the binary operators. Let $A = (a_i)$ and $B = (b_i)$, and c be a real number.

Vector Addition and Scalar Multiplication

$$A + B = (a_i + b_i)$$
$$cA = (ca_i)$$

Just like how we worked with vectors in linear algebra, we do these component wise.

Addition and scalar multiplication together form a vector space–after all, $\mathbb{R}^n$ is one.

We previously described a given tuple (a_i) as denoting the displacements onto axes to describe a point in space. Such tuples can also be used to describe vectors, in the physics sense, corresponding to actual displacements. Consider the point p. Then the vector p is the arrow that points from the origin to p.

Geometrically, vector addition of A and B corresponds to taking the two arrows, let's say B, and moving it so that instead of B starting from the origin, B starts at the point A points to. Scalar multiplication is just scaling the length of the vector. For posterity, direction is maintained when the scalar is greater than zero, and flipped if its less than zero.

1.1.3 Standard Basis Vectors

Definition 1.1: Standard Basis Vectors

We define the standard basis vectors $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$, as

$$\begin{aligned}\mathbf{i} &= (1,0,0)\\ \mathbf{j} &= (0,1,0)\\ \mathbf{k} &= (0,0,1)\end{aligned}$$

A consequence of this is that for any given vector $a = (a_1, a_2, a_3)$, we can express the vector as the sum $a = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$. I actually find this to be rather useless, and only useful when defining the cross product.

This notation can potentially save space, but becomes rather cumbersome when instead of a_i being constants, they become actual functions, which we will see later.

Example 1.1: MT 1.1 Problem 4

Compute

$$(2,3,5) - 4\mathbf{i} + 3\mathbf{j}$$

Since $\mathbf{i} = (1,0,0)$ and $\mathbf{j} = (0,1,0)$, this is just

$$\begin{aligned}(2,3,5) - 4(1,0,0) + 3(0,1,0) &= (2,3,5) + (-4,0,0) + (0,3,0)\\ &= (2-4, 3+3, 5+0)\\ &= (-2,6,5).\end{aligned}$$

1.1.4 Vectors between points, lines between points, parallelograms

These equations will show up time and time again.

Theorem 1.1: Vector joining two points

Let A and B be two vectors. Then the vector from A to B, denoted as $\overrightarrow{AB}$, is given as

$$\overrightarrow{AB} = B - A.$$

A way to think of this is that $\overrightarrow{AB}$ refers to the displacement to get to the point at B if we start at A. If A were the origin, this would be quite easy, and $\overrightarrow{AB} = B$. Instead, we can just shift our perspective and pretend the origin is A. Then B, if the origin is now A, is given by $B - A$.

Theorem 1.2: Point-Direction Form of a Line

The equation of a line from a that goes in the direction of a vector v is

$$l(t) = a + tv.$$

Decomposing these into a system of equations give the a system of parametric equations for the line. Let $a = (x_1, y_1, z_1)$ and $v = (a, b, c)$. Then expanding these gives

$$\begin{aligned} x &= x_1 + at \\ y &= y_1 + bt \\ z &= z_1 + ct. \end{aligned}$$

Example 1.2: MT 1.1 Problem 16

Use set theoretic or vector methods or both to describe the points that lay in the line passing through $(0, 2, 1)$ in the direction of $2\mathbf{i} - \mathbf{k}$.

The direction is $(2, 0, -1)$. Then the line is given as

$$\begin{aligned} r(t) &= (0, 2, 1) + t(2, 0, -1) \\ &= (2t, 2, 1 - t). \end{aligned}$$

This can also be written as

$$\{(2t, 2, 1 - t) \mid t \in \mathbb{R}\}.$$

An important special case of this theorem is achieved by taking the point-direction form of a line, and obtaining the vector from one vector to another vector by the vector joining two points is the point-point form of a line:

Theorem 1.3: Point-Point Form of a Line

For a point A and a point B, the line is given by

$$l(t) = A + (B - A)t.$$

This also has a system of parametric equations variation as well.

Example 1.3: MT 1.1 Problem 18

Use set theoretic or vector methods or both to describe the points that lay in the line passing through $(-5, 0, 4)$ and $(6, -3, 2)$.

Take $A = (-5, 0, 4)$ and $B = (6, -3, 2)$. Then $B - A = (11, -3, -2)$, and then

$$r(t) = (-5, 0, 4) + t(11, -3, -2)$$

$$= (-5 + 11t, -3t, 4 - 2t).$$

When discussing equations of lines, or really any mathematical object in higher dimensional space, equations are not unique. We do not have this issue in 2 dimensions, really, since most of our objects (except circles, incredibly), can be completely described by $y = f(x)$. We shall see when we discuss surfaces that this is not a complete picture (it's already incomplete, because this fails for circles), but problematically, one cannot define a line as $y = f(x)$ in higher dimensional space. Importantly, there's a z term. We can hope for something like $y = f(x)$ and $z = g(x)$, but beyond the fact that this doesn't work (ah, now spheres!), here's an example where non-uniqueness kicks in.

Consider the line from $(-1, 2, 3)$ pointing in the direction of the vector $(4, 5, 6)$. This gives the line

$$l(t) = (-1 + 4t, 2 + 5t, 3 + 6t).$$

However, the magnitude of the direction of the vector is irrelevant to the direction of the line, so the direction of $(8, 10, 12)$ is the same. This gives the line

$$l(t) = (-1 + 8t, 2 + 10t, 3 + 12t)$$

which is actually the same line, but "moves twice as fast" with respect to t (this idea is made more rigorous in line integrals). But they're still the same line.

Theorem 1.4: Points Inside Spanned Parallelogram

Consider the parallelogram formed by a, b, the origin, and $a + b$. Then the set of all points, including the edges, is described as

$$as + bt \quad 0 \le s, t \le 1.$$

Example 1.4: MT 1.1 Problem 14

Use set theoretic or vector methods or both to describe the points that lay in the plane spanned by $v_1 = (3, -1, 1)$ and $v_2 = (0, 3, 4)$.

$$\begin{aligned} r(s,t) \quad 0 \le s,t \le 1 &\implies sv_1 + tv_2 \quad 0 \le s,t \le 1 \\ &\implies s(3,-1,1) + t(0,3,4) \quad 0 \le s,t \le 1 \\ &\implies (3s, -s + 3t, s + 4t) \quad 0 \le s,t \le 1. \end{aligned}$$

This can also be written as

$$\{(3s, -s + 3t, s + 4t) \mid s, t \in [0, 1]\}.$$

1.2 "The Inner Product, Length and Distance" (MT 1.2)

What makes $\mathbb{R}^3$ *somewhat* special is how closely it models the real world (unfortunately, it does assume zero-curvature if we model the surface of the Earth). Though we have means to describe position and the translation of an object as a consequence, we still do not have a way to define how far things apart are. We do have the distance formula and its higher dimensional analogues, but we can now put them on slightly more rigorous footing.

1.2.1 The Inner Product and Norms

Definition 1.2: The Inner Product

The inner product of two vectors A and B, also known as the dot product, denoted $A \cdot B$, is defined as

$$A \cdot B = \sum_{\forall i} a_i b_i$$

That is, the sum of all corresponding component products.

Example 1.5: MT 1.2 Problem 2a

Calculate $a \cdot b$ where $a = 2\mathbf{i} + 10\mathbf{j} - 12\mathbf{k}$ and $b = -3\mathbf{i} + 4\mathbf{k}$.

Converting to vector notation, $a = (2, 10, -12)$ and $b = (-3, 0, 4)$. This gives $a \cdot b = -6 + 0 - 48 = -54$.

We have the following properties:

- $A \cdot A \geq 0$, and this is zero when $A = 0$. (The inner product is a norm)
- $a(A \cdot B) = aA \cdot aB$ and $aA \cdot B = a(A \cdot B)$. (The inner product distributes over scalar multiplication)
- $A \cdot (B + C) = A \cdot B + A \cdot C$. (The inner product distributes over addition)
- $A \cdot B = B \cdot A$. (The inner product is commutative).

The inner product allows us to completely define length, distance, and angles. In that sense, it is the inner product, together with the standard vector space definition, that forms the basis of this text.

We can now define the length – or norm – of a vector.

Definition 1.3: The Norm of a Vector

We define the norm of a vector, $\|A\|$ (or $|A|$ if I'm lazy), as

$$\|A\| = \sqrt{A \cdot A}$$

As a tip, it's generally easier, to find the norm of a vector, to first find the norm of the vector *squared*. That is, compute $\|A\|^2$, then take the square root at the end. For more complicated expressions, this saves you from having to write down the square root sign over and over.

Example 1.6: MT 1.2 Problem 6

Compute $\|u\|$, $\|v\|$ and $u \cdot v$ if $u = (15, -2, 4)$ and $v = (\pi, 3, -1)$.

$$\begin{aligned} \|u\|^2 &= 225 + 4 + 16 \\ &= 245 \\ \|u\| &= \sqrt{245} \\ \|v\|^2 &= \pi^2 + 9 + 1 \\ &= \pi^2 + 10 \\ \|v\| &= \sqrt{\pi^2 + 10} \\ u \cdot v &= 15\pi - 6 - 4 \\ &= 15\pi - 10. \end{aligned}$$

Example 1.7: MT 1.2 Problem 9

Compute $\|u\|$, $\|v\|$ and $u \cdot v$ if $u = -\mathbf{i} + 3\mathbf{j} + \mathbf{k}$ and $v = -2\mathbf{i} - 3\mathbf{j} - 7\mathbf{k}$.

Converting to vector notation, $u = (-1, 3, 1)$ and $v = (-2, -3, -7)$. Then

$$\begin{aligned} \|u\|^2 &= 1 + 9 + 1 \\ &= 11 \\ \|u\| &= \sqrt{11} \\ \|v\|^2 &= 4 + 9 + 49 \\ &= 62 \\ \|v\| &= \sqrt{62} \\ u \cdot v &= 2 - 9 - 7 \\ &= -14. \end{aligned}$$

Example 1.8: MT 1.5 Problem 1

Calculate the dot product of $(1, -1, 0, 2)$ and $(1, 2, 3, 4)$

This is $1 - 2 + 0 + 8 = 7$.

Example 1.9: MT 1.2 Problem 14b

Let $v_1 = (0, 3, 0)$, $v_2 = (2, 2, 0)$ and $v_3 = (1, 1, 3)$. These three vectors with their tails at the origin determine a parallelepiped P.
Determine the length of the main diagonal (from the origin to the opposite vertex).

Using the head-tail addition for vectors, this the opposite vertex is $v_1 + v_2 + v_3 = (3, 6, 3)$, which is also the vector to the opposite vertex from the origin. This gives that $\|v_1+v_2+v_3\|^2 = 9 + 36 + 9 = 54$, so $\|v_1 + v_2 + v_3\| = 3\sqrt{6}$.

1.2.2 Unit Vectors

In some vector calculus texts, unit vectors are particularly important, namely when we get to surface integrals. This is not the case in this text.

We often speak of vectors as having a magnitude and a direction. This often leads to non-uniqueness, as we frequently use the directional portion of a vector to parameterize objects, which means that an infinite family of vectors can satisfy the parameterization if direction is the only relevant component. Hence, we can scale a vector by its length to obtain a unit vector–a vector of length 1 to obtain a "true" direction vector, in some senses.

Definition 1.4: Unit Vector

We define a unit vector as any vector with norm 1.
To take an arbitrary vector A and scale it by multiplying by the appropriate factor to obtain a unit vector is known as *normalization*, and is done by

$$\frac{A}{\|A\|}.$$

1.2.3 Distance

We have defined displacement vectors, that is, the vector that corresponds to "getting to" B "from" A, $\overrightarrow{AB}$, as $B - A$. This makes defining distance very straightforward as the norm of this displacement vector.

Definition 1.5: Distance between two vectors

We define the distance between two vectors A and B as the norm of their displacement vector:

$$d(A, B) = \|B - A\|.$$

1.2.4 Angles

In non traditional vector settings, the definition of the inner product is used for angles specifically. The cosine similarity metric does this. Since many mathematical objects can be easily represented as vectors, even as (honestly, I would say obviously) as one-hot encoding or an ordered multi-set, we can define the angles between them, which more or less says how far apart they are. We don't really make use of that property, but since that's generally the context where I encounter angles of vectors, I think it might be interesting to note.

Theorem 1.5: Dot-Product and Angles and Proofs

Let A and B be vectors and let θ be the smallest angle between the two. Then we have that

$$A \cdot B = \|A\|\|B\| \cos(\theta).$$

The proof traditionally follows from the law of cosines. We have that

$$\|B - A\| = \|A\|^2 + \|B\|^2 - 2\|A\|\|B\| \cos(\theta)$$

Rewriting this partly using norms gives

$$(B - A) \cdot (B - A) = A \cdot A + B \cdot B - 2\|A\|\|B\| \cos(\theta).$$

By algebra, we can expand $(B - A)$ using the summation definition (an advantage of defining it as such):

$$\begin{aligned}(B - A) \cdot (B - A) &= \sum_{\forall i} (b_i - a_i)(b_i - a_i) \\ &= \sum_{\forall i} b_i^2 - 2a_i b_i + a_i^2 \\ &= \sum_{\forall i} a_i^2 + \sum_{\forall i} b_i^2 - 2 \sum_{\forall i} a_i b_i \\ &= A \cdot A + B \cdot B - 2A \cdot B\end{aligned}$$

Which then gives that $2A \cdot B = 2\|A\|\|b\| \cos(\theta)$, and dividing by 2 gives the desired equality.

Example 1.10: MT 1.2 Problem 3

Find the angle between $7\mathbf{j} + 19\mathbf{k}$ and $-2\mathbf{i} - \mathbf{j}$ to the nearest degree.

Call the vectors A and B. Converting to vector form gives the vectors $A = (0, 7, 19)$ and $B = (-2, -1, 0)$. Then $\|A\| = \sqrt{49 + 361} = \sqrt{410}$ and $\|B\| = \sqrt{4 + 1} = \sqrt{5}$. $A \cdot B = -7$. Putting this altogether gives

$$\begin{aligned}-7 &= \sqrt{410} \cdot \sqrt{5} \cos(\theta) \\ \cos(\theta) &= \frac{-7}{5\sqrt{82}} \\ \theta &= \arccos\left(\frac{-7}{5\sqrt{82}}\right) \\ &\approx 99^\circ.\end{aligned}$$

We have a special name for perpendicular vectors,

Definition 1.6: Orthogonal Vectors

Let A and B have inner angle $\theta = 90^\circ = \pi/2$. We say A and B are *orthogonal*.

Example 1.11: MT 1.2 Problem 18

Find all values of x such that $(x, 1, x)$ and $(x, -6, 1)$ are orthogonal.

If $(x, 1, x)$ and $(x, -6, 1)$ are orthogonal, then $(x, 1, x) \cdot (x, -6, 1) = 0$. This dot product is given as $x^2 - 6 + x$, which factors as $(x + 3)(x - 2)$ and hence has zeros at $x = 2$ and $x = 3$.

Most dot product related proofs work the same as the cosine one mechanically. That a dot product is a sum of products is actually extremely easy to work with, since all algebra that we are familiar with can be used.

As an example,

Example 1.12: MT 1.5 Problem 2a

Show that

$$2\|X\| + 2\|Y\|^2 = \|X + Y\|^2 + \|X - Y\|^2.$$

This is known as the parallelogram law.

We can really pick either side, but since the left side is simpler, we will transform the right into the left. Using the dot product formula,

$$\begin{aligned}
\|X + Y\|^2 + \|X - Y\|^2 &= \sum_{\forall i}(x_i + y_i)^2 + \sum_{\forall i}(x_i - y_i)^2 \\
&= \sum_{\forall i} x_i^2 + 2x_iy_i + y_i^2 + \sum_{\forall i} x_i^2 - 2x_iy_i + y_i^2 \\
&= \sum_{\forall i} 2x_i^2 + 2y_i^2 \\
&= 2\sum_{\forall i} x_i^2 + 2\sum_{\forall i} y_i^2 \\
&= 2\|X\| + 2\|Y\|
\end{aligned}$$

as desired.

Example 1.13: MT 1.5 Problem 7

If $\|V\| = \|W\|$, show that $\|V + W\|$ and $\|V - W\|$ are orthogonal.

It suffices to show $\|V + W\| \cdot \|V - W\| = 0$.

$$\begin{aligned}
\|V + W\| \cdot \|V - W\| &= (v_i + w_i) \cdot (v_i - w_i) \\
&= \sum_{\forall i} v_i^2 - w_i^2 \\
&= \sum_{\forall i} v_i^2 - \sum_{\forall i} w_i^2 \\
&= \|V\| - \|W\| \\
&= 0.
\end{aligned}$$

1.2.5 The Cauchy-Schwarz Inequality

Theorem 1.6: The Cauchy-Schwarz Inequality

$$\|A \cdot B\| \leq \|A\|\|B\|.$$

The proof comes from looking at the dot-product angle formula and taking the absolute value, $\|A \cdot B\| \leq \|A\|\|B\| \cos(\theta)$, and noting that $-1 \leq \cos(\theta) \leq 1$, so $\|A\|\|B\| \leq \|A\|\|B\|$.

An important special case of the Cauchy-Schwarz inequality is the triangle inequality.

Theorem 1.7: The Triangle Inequality

For any two vectors A and B,

$$\|A + B\| \leq \|A\| + \|B\|.$$

Example 1.14: MT 1.5 Problem 13

Use induction on k to show that

$$\|x_1 + \ldots + x_k\| \leq \|x_1\| + \ldots \|x_k\|$$

Base case $k = 1$. This is just $\|x_1\| \leq \|x_1\|$, which is true.

Base case $k = 2$. This is $\|x_1 + x_2\| \leq \|x_1\| + \|x_2\|$, which is true by the triangle inequality.

Induction step $k = n$. We have that

$$\|x_1 + \ldots + x_{n-1}\| \leq \|x_1\| + \ldots \|x_{n-1}\|$$

It suffices to show

$$\|x_1 + \ldots + x_n\| \leq \|x_1\| + \ldots \|x_n\|$$

The trick is that since vectors form a vector field, vectors are closed under addition, so the addition of two vectors is also a vector, and we can convert the sum of n vectors into $n - 1$. Let $y = x_{n-1} + x_n$, so it suffices to show

$$\|x_1 + \ldots + x_{n-2} + y\| \leq \|x_1\| + \ldots \|x_{n-2}\| + \|x_{n-1}\| + \|x_n\|$$

By the induction hypothesis, we have that

$$\|x_1 + \ldots + x_{n-2} + y\| \leq \|x_1\| + \ldots \|x_{n-2}\| + \|y\|$$

What remains is to show $\|y\| \leq \|x_{n-1}\| + \|x_n\|$. Since $y = x_{n-1} + x_n$, this is true by the triangle inequality, or case $k = 2$.

1.2.6 Orthogonal Projection

A projection generally refers to "dropping" something in higher dimensional space into lower dimensional space. For instance, if we take the point $(1, 2, 3)$ and just ignore the z-component, we view it as $(1, 2)$. This is "projecting" it into the remaining space, the xy-plane. In other words, the xy-plane projection of a point in xyz-space is it's "shadow", if we pretend the sun exists at $(0, 0, \infty)$.

We also talk about projecting vectors onto other vectors, or the projection of a vector V *onto* another vector A. I think of this as "the amount of V that agrees with A", and hence from our discussion from cosine similarity, you can imagine we will use the inner product for this.

Definition 1.7: The projection of a vector onto another vector

We define the projection of a vector V onto a vector A, denoted $\text{proj}_A V$ as

$$\text{proj}_A V = \frac{A \cdot V}{A \cdot A} A$$

This is more conventionally presented as

$$\frac{A \cdot V}{\|A\|} A.$$

This takes how close V and A are, and scales A correspondingly.

Example 1.15: MT 1.2 Problem 20

Find the projection of $u = -\mathbf{i} + \mathbf{j} + \mathbf{k}$ onto $v = 2\mathbf{i} + \mathbf{j} - 3\mathbf{k}$.

Firstly, $u = (-1, 1, 1)$ and $v = (2, 1, -3)$. Then

$$\begin{aligned}
\text{proj}_v u &= \frac{u \cdot v}{v \cdot v} v \\
&= \frac{-2 + 1 - 3}{4 + 1 + 9}(2, 1, -3) \\
&= \frac{-4}{14}(2, 1, -3) \\
&= \left(\frac{-4}{7}, \frac{-2}{7}, \frac{6}{7} \right)
\end{aligned}$$

1.3 "Matrices, Determinants and the Cross Product (MT 1.3)"

I'm going to assume you know linear algebra, and will skip over many details.

We define the determinant by the co-factor expansion. As an example, the 3×3 case is given as

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = a_1 \begin{vmatrix} \square & \square & \square \\ \square & b_2 & c_2 \\ \square & b_3 & c_3 \end{vmatrix} - b_1 \begin{vmatrix} \square & \square & \square \\ a_2 & \square & c_2 \\ a_3 & \square & c_3 \end{vmatrix} + c_1 \begin{vmatrix} \square & \square & \square \\ a_2 & b_2 & \square \\ a_3 & b_3 & \square \end{vmatrix}$$

$$= a_1 \begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix} \qquad - b_1 \begin{vmatrix} a_2 & c_2 \\ a_3 & c_3 \end{vmatrix} \qquad + c_1 \begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix}$$

A very useful properties of the determinant is that you can factor things out *row-wise* or *column-wise*.

Example 1.16: MT 1.2 Problem 2a

Evaluate the determinant

$$\begin{vmatrix} 2 & -1 & 0 \\ 4 & 3 & 2 \\ 3 & 0 & 1 \end{vmatrix}$$

$$\begin{aligned} \begin{vmatrix} 2 & -1 & 0 \\ 4 & 3 & 2 \\ 3 & 0 & 1 \end{vmatrix} &= 2(3) - (-1)(4-6) + 0 \\ &= 6 - 2 \\ &= 4. \end{aligned}$$

1.3.1 Cross Products

The cross product is an incredibly *useful* operator. I contrast this to the dot product, which is moreso *fundamental*. You need the dot product to even get the ground running, and once you do, the cross product makes many many calculations feasible.

Definition 1.8: (Fake) The Cross Product of Two Vectors

Let $A = (a_1, a_2, a_3)$ and $B = (b_1, b_2, b_3)$ be vectors in $\mathbb{R}^3$. Then we define the cross product of A and B, denoted $A \times B$, as

$$A \times B = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}.$$

This is pretty much the only time we use the standard basis vectors.

Since determinants require use to alternate signs and the cross product is really only used for two vectors in $\mathbb{R}^3$, we can do some optimization. Instead of flipping the result of the second cofactor expansion, we can instead just flip the diagonals we take the difference of. Instead of doing the equivalent of $-a_2(b_1c_3 - b_3c_1)$, we can instead do $a_2(b_3c_1 - b_1c_3)$.

An important restriction is that the cross product is only defined in 3-dimensions. Well, this isn't completely true. Though we work in $\mathbb{R}^3$, we can represents vectors in $\mathbb{R}^2$ as vectors with z-component 0, and we can then get the cross product to work in 2-dimensions. This is also how we'll define the scalar curl later.

However, for other dimensions, the cross product is not well defined except $\mathbb{R}^7$ oddly enough, and the determinant definition breaks in all dimensions. This is why much of vector calculus takes place in $\mathbb{R}^3$.

I still haven't explained why the cross product is useful. It's useful for this one reason, given by the "true" definition of the cross product

Definition 1.9: (True) The Cross Product of Two Vectors

Let A and B be as before, let $\mathbf{n}$ be a unit vector perpendicular to to the plane that contains both A and B obeying the right-hand rule, and let θ be their internal angle. Then

$$A \times B = \|A\|\|B\| \sin(\theta)\mathbf{n}.$$

The important consequence of the $\mathbf{n}$ is to not be understated.

The cross product gives a vector orthogonal to both vectors.

Taking the norm of both sides gives a cross product angle formula analogous to the inner product.

Example 1.17: MT 1.3 Problem 10

Describe all unit vectors orthogonal to both $-5\mathbf{i} + 9\mathbf{j} - 4\mathbf{k}$ and $7\mathbf{i} + 8\mathbf{j} + 9\mathbf{k}$.

Converting to vector notation gives the vectors (-5, 9, -4) and $(7, 8, 9)$. A vector orthogonal to both is given by their cross product,

$$\begin{aligned} \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ -5 & 9 & -4 \\ 7 & 8 & 9 \end{vmatrix} &= (81 + 32, -28 + 45, -40 - 63) \\ &= (113, 17, -103). \end{aligned}$$

I don't really want to finish this; normalizing gives one such unit normal vector. Flipping the sign of that vector gives the other.

Some properties of the cross product:

- $A \times B = -B \times A$ (the cross product is anticommutative)
- $A \times (B + C) = A \times B + A \times C$ (the cross product distributes over addition)
- $aA \times B = a(A \times B)$ (the cross product distributes over scalar multiplication)

Theorem 1.8: Areas of parallelepipeds

Let A, B and C be vectors in $\mathbb{R}^3$. Then the area of the paralleogram spanned by A and B is given by

$$\|A \times B\|$$

and the area of the parallelepiped spanned by A and B and C is given by the *signed tripled product*,

$$A \cdot (B \times C).$$

Example 1.18: MT 1.3 Problem 4

Compute $a \cdot (b \times c)$ where $a = \mathbf{i} - 2\mathbf{j} + \mathbf{k}$, $b = 2\mathbf{i} + \mathbf{j} + \mathbf{k}$ and $c = 3\mathbf{i} - \mathbf{j} + 2\mathbf{k}$.

Converting these to vector notation, $a = (1, -2, 1)$, $b = (2, 1, 1)$ and $c = (3, -1, 2)$. Then,

$$\begin{aligned} a \cdot (b \times c) &= (1, -2, 1) \cdot \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 2 & 1 & 1 \\ 3 & -1 & 2 \end{vmatrix} \\ &= (1, -2, 1) \cdot (3, -1, -5) \\ &= 3 + 2 - 5 \\ &= 0. \end{aligned}$$

Example 1.19: MT 1.3 Problem 5

Find the area of the parallelogram given with sides $a = \mathbf{i} - 2\mathbf{j} + \mathbf{k}$ and $b = 2\mathbf{i} + \mathbf{j} + \mathbf{k}$.

Converting to vector notation, $a = (1, -2, 1)$ and $b = (2, 1, 1)$. Then,

$$\begin{aligned} |a \times b| &= \left| \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & -2 & 1 \\ 2 & 1 & 1 \end{vmatrix} \right| \\ &= |(-3, 1, 5)| \\ &= \sqrt{9 + 1 + 25} \\ &= \sqrt{35}. \end{aligned}$$

Example 1.20: MT 1.3 Problem 8

What is the volume of the parallelepiped with sides $\mathbf{i}$, $3\mathbf{j} - \mathbf{k}$ and $4\mathbf{i} + 2\mathbf{j} - \mathbf{k}$?

Converting to vector notation gives the sides $(1, 0, 0)$, $(0, 3, -1)$ and $(4, 2, -1)$. Computing the triple product,

$$\left| (1, 0, 0) \cdot \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 0 & 3 & -1 \\ 4 & 2 & -1 \end{vmatrix} \right| = |(1, 0, 0) \cdot (-1, -4, -12)|$$

$$= |-1|$$
$$= 1.$$

1.3.2 Equation of a plane

We can now finally start defining objects that aren't lines.

Theorem 1.9: Equation of a plane

Let $\mathcal{P}$ be a plane containing the point (x_0, y_0, z_0) and a normal vector $\mathbf{n} = (A, B, C)$. Then a point (x, y, z) is on the plane if and only if it satisfies

$$A(x - x_0) + B(y - y_0) + C(z - z_0) = 0.$$

To see why this is true, take the point (x_0, y_0, z_0) and point the vector $\mathbf{n}$ out from that point. Now, consider all vectors that are orthogonal to it that are of unit length (the length doesn't actually matter for this argument, but it keeps it simpler). This should form a circle around the point. Now, since scaling a vector doesn't affect whether or not it's orthogonal to another, we can stretch these arbitrarily. This gives a plane.

Two planes are parallel if their normal vectors are parallel. Similarly, two planes are orthogonal if their normal vectors are orthogonal.

Example 1.21: MT 1.3 Problem 15a

Find an equation for the plane that is perpendicular to $v = (1, 1, 1)$ and passes through the point (1, 0, 0)

$$\begin{aligned} 1(x-1) + 1(y-0) + 1(z-0) = 0 &\implies x - 1 + y + z = 0 \\ &\implies x + y + z - 1 = 0. \end{aligned}$$

Example 1.22: MT 1.3 Problem 35

Find an equation for the plane that contains the line $v = (-1, 1, 2) + t(3, 2, 4)$ and is perpendicular to the plane $2x + y - 3z + 4 = 0$.

Call the target plane $\mathcal{P}$ and the plane $2x + y - 3z + 4 = 0$ $\mathcal{P}'$

Since the line has directional vector $(3, 2, 4)$, $(3, 2, 4)$ is parallel to $\mathcal{P}$.

$\mathcal{P}'$ is orthogonal to the vector $(2, 1, -3)$. Since $\mathcal{P}$ is orthogonal to $\mathcal{P}'$, $(2, 1, -3)$ is parallel to $\mathcal{P}$.

Since $(3, 2, 4)$ and $(2, 1, -3)$ are both parallel to $\mathcal{P}$ (and are not scalar multiples), their nonzero cross product will be orthogonal to the plane.

$$\begin{aligned}\begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 3 & 2 & 4 \\ 2 & 1 & -3 \end{vmatrix} &= (-6-4, 8+9, 3-4) \\ &= (-10, 17, -1).\end{aligned}$$

Since the line has starting position $(-1, 1, 2)$, this point is on the plane. Then our plane is given as

$$-10(x+1) + 17(y-1) - 1(z-2) = 0.$$

1.3.3 Distance from a point to a plane

Theorem 1.10: Distance from a point to a plane

Let $\mathcal{P}$ be a plane with equation $Ax + By + Cz + D = 0$ and let (x_1, y_1, z_1) be a point. Then the distance from the point to $\mathcal{P}$ is given as

$$\frac{|Ax_1 + By_1 + Cz_1 + D|}{\sqrt{A^2 + B^2 + C^2}}.$$

I remember this by taking the numerator to be "evaluating the point" within the plane equation.

Example 1.23: MT 1.3 Problem 34

Find the distance from the point $(2, 1, -1)$ to the plane $x - 2y + 2z + 5 = 0$.

$$\begin{aligned}\frac{|2-2-2+5|}{\sqrt{1+4+4}} &= \frac{|3|}{\sqrt{5}} \\ &= \frac{3\sqrt{5}}{5}.\end{aligned}$$

2 "Differentiation" (MT 2)

I assume you know single variable calculus.

2.1 "The Geometry of Real Valued Functions" (MT 2.1)

This is unfortunately a really boring subsection, and is basically only definitions. These have actual tangible meaning and uses, but are really only introduced for their own sake until much much later.

2.1.1 Functions and Mappings

Definition 2.1: Function Notation

When we say

$$f : U \subset A \mapsto B,$$

this means f maps points from U, a subset of A, to B.

2.1.2 Graphs of Functions

Definition 2.2: Graph of a Function

For a function $f : U \subset \mathbb{R}^n \mapsto \mathbb{R}$, we define the the graph of f, as an object in $\mathbb{R}^{n+1}$ space as the set of points

$$\{(*X, f(X))\}$$

where $(*)$ here is an unpacking operator.

2.1.3 Level Objects

Definition 2.3: Level Set

For a function $f : U \subset \mathbb{R}^n \mapsto \mathbb{R}$, a level set of f is any set of the form

$$\{X \mid f(X) = c\}$$

for some constant c.

2.1.4 Sections

Definition 2.4: Section of a graph

For a function $f : U \subset \mathbb{R}^n \mapsto \mathbb{R}$, a section is defined as the intersection of a flat plane and the graph.

Example 2.1: MT Section 2.1 Problem 8

Sketch level sets of values $c = 0, 1, 4, 9$ for both $f(x, y) = x^2+y^2$ and $g(x, y) = \sqrt{x^2 + y^2}$. How are the graphs of f and g different? How are their sections different?

I'll leave the sketching to you.

Let's consider a level set first. Suppose we hold $f(x, y) = c$ and $g(x, y) = c$ for some constant c. For f, this gives $x^2 + y^2 = c$ and for g this gives $\sqrt{x^2 + y^2} = c$ or $x^2 + y^2 = c^2$.

f is then a circle of radius $\sqrt{c}$ and g is a circle of radius c–the circles are wider for g. In the reverse sense, this means that f rises faster than g.

For a section, let's take $x = 0$. This gives $z = x^2$ for f and $z = \sqrt{x^2}$ or $z = x$ for g. This affirms our prior discussion; f is a cylindrical paraboloid whereas g is a flat cone. This then answers the graph question.

Example 2.2: MT Section 2.1 Problem 9

Let S be the surface in $\mathbb{R}^3$ defined by the equation $x^2y^6 - 2z = 3$.

(a) Find a real-valued function $f(x, y, z)$ of three variables and a constant c such that S is the level set of f of value c.

(b) Find a real-valued function $g(x, y)$ of two variables such that S is the graph of g.

(a) Since $x^2y^6 - 2x = 3$, we already have an expression in terms of x and y on the left hand side and a constant on the right. So take $f(x, y, z) = x^2y^6 - 2x$ and $c = 3$.

(b) If S is the graph of some $g(x, y)$, this is the set of points $\{(x, y, g(x, y)) \mid x, y \in \mathbb{R}\}$. Hence we need to find z in terms of x and y. By algebra, we have $-2z = 3 - x^2y^6$ and then $z = (x^2y^6 - 3)/2 = g(x, y)$.

2.2 Limits and Continuity (MT 2.2)

There's a lot of technical things in this section, most of which is useful for developing a theory of differentiation of multivariate functions, but not much else. There is a single exception.

2.2.1 Developing Continuity and Differentiation

Definition 2.5: Open Disk

We denote the open disk of radius r around a point X, $D_r(X_0)$ as all points X that satisfy

$$d(X, X_0) < r.$$

We often say $D_r(x_0)$ is a neighborhood around x_0.

Definition 2.6: Open Set

A subset U of $\mathbb{R}^n$ is an open set if for any $x_0 \in U$, there is some r such that $D_r(x_0) \subset U$.

What this means is that an open set is one where for any point, it's "reasonably inside". For any point you pick, you can always find some point that is "closer" to being outside than the selected point.

Definition 2.7: Boundary Point

Let $A \subset \mathbb{R}^n$. A point x is called a boundary point if for any r, $D_r(x)$ has a point both inside and outside A.

In other words, boundary points are points on the edge of the surface.

2.2.2 Limits

Definition 2.8: Limit of a function

Let $f : A \subset \mathbb{R}^m \mapsto \mathbb{R}^n$ where U is an open set. Let x_0 be in A, and let $N = D_r(b)$ where $b \in \mathbb{R}^n$.
We say f is eventually in N as x approaches x_0 if there exists some $U = D_r(x_0)$ where if $x \neq x_0$, $x \in U$ and $x \in A$ implies that $f(x) \in N$.

In other words, $\lim_{x \to x_0} f(x) = b$.

Let $\lim_{x \to x_0} f(x) = b$ and $\lim_{x \to x_0} g(x) = c$ Then we have the following properties

- $\lim_{x \to x_0} kf(x) = kb$. (Limits are linear)
- $\lim_{x \to x_0} f(x) + g(x) = b + c$. (Limits are linear)
- If $c, b \in \mathbb{R}$, then $\lim_{x \to x_0} f(x)g(x) = bc$. (Results of limits multiply)
- If $b \neq 0$, then $\lim_{x \to x_0} 1/f(x) = 1/b$. (Limits respect division)

The limit of a multivariable function can also be understood as the tuple of the limits of the individual component functions.

Limits in multivariable functions are super super finnicky, and evaluating them without resorting to epsilon-delta methods is difficult. When a function is continuous, we can simply use regular limit formulae, but if we don't know if a function is continuous, things get harder. This is often a point of major confusion, and was for me:

> To show that $\lim_{x \to x_0} f(x) = b$, this must be true *for every possible path* to x_0. In contrast, to show that $\lim_{x \to x_0} f(x)$ does not exist, one must find two paths that disagree.

Example 2.3: MT Section 2.1 Problem 6

Let

$$f(x,y) = \begin{cases} \frac{xy^3}{x^2+y^6} & (x,y) \neq (0,0) \\ 0 & (x,y) = (0,0) \end{cases}$$

(a) Compute the limit as $(x,y) \to (0,0)$ of f along the path $x = 0$.
(b) Compute the limit as $(x,y) \to (0,0)$ of f along the path $x = y^3$.

(a) The path $x = 0$ allows us to consider $f(0,y)$, which gives

$$f(x,y) = \begin{cases} \frac{0y^3}{0^2+y^6} = \frac{0}{y^6} = 0 & y \neq 0 \\ 0 & y = 0 \end{cases}$$

Which is just $f(x,y) = 0$, so $\lim_{(x,y)\to(0,0)} f(x,y)$ on this path is 0.

(b) The path $x = y^3$ allows us to consider $f(y^3,y)$, which gives

$$f(x,y) = \begin{cases} \frac{y^3y^3}{(y^3)^2+y^6} = \frac{y^6}{2y^6} = \frac{1}{2} & y \neq 0 \\ 0 & y = 0 \end{cases}$$

Then $\lim_{(x,y)\to(0,0)} f(x,y) = 1/2$ on this path.

So the limit of f doesn't exist at $(0,0)$.

Example 2.4: MT Section 2.2 Problem 9a

Compute the following limit if it exists:

$$\lim_{(x,y)\to(0,0)} \frac{e^{xy}-1}{y}.$$

Ideally, if we were instead considering something like $(e^{xy}-1)/xy$, we could make a substitution like $u = xy$ and use L'Hopital's rule. So we want to multiply the denominator by x. So let's just multiply both the numerator and the denominator, then.

$$\lim_{(x,y)\to(0,0)} \frac{e^{xy}-1}{y} = \lim_{(x,y)\to(0,0)} \frac{e^{xy}-1}{y} \cdot \frac{x}{x} = \lim_{(x,y)\to(0,0)} \frac{x(e^{xy}-1)}{xy}.$$

The key step is to view the expression $(x(e^{xy}-1))/xy$ as the product of x and $(e^{xy}-1)/xy$. This let's us split the limit as such:

$$\lim_{(x,y)\to(0,0)} \frac{e^{xy}-1}{y} = \left(\lim_{(x,y)\to(0,0)} x\right) \cdot \left(\lim_{(x,y)\to(0,0)} \frac{e^{xy}-1}{xy}\right)$$

The left one is 0, which finishes the problem, but for posterity, under the substitution $u = xy$, the right gives 1 by L'Hopital's rule.

2.2.3 Continuity

Definition 2.9: Continuity

Let $f : A \mapsto B$. Then f is continuous at x_0 if

$$\lim_{x \to x_0} f(x) = f(x_0)$$

All the same properties from limits more or less also hold for continuity–if functions f and g are continuous at x_0, so are linear combinations of f and g, as well as the reciprocals if they are non-zero at x_0.

Example 2.5: MT Section 2.2 Problem 1

Let $f : \mathbb{R}^2 \mapsto \mathbb{R}$ and suppose that $\lim_{(x,y)\to(1,3)} f(x, y) = 5$. What can you say about the value $f(1, 3)$?

Nothing. We have no reason to believe that $\lim_{(x,y)\to(1,3)} f(x, y) = f(1, 3)$.

Example 2.6: MT Section 2.2 Problem 2

Let $f : \mathbb{R}^2 \mapsto \mathbb{R}$ be continuous and suppose that $\lim_{(x,y)\to(1,3)} f(x, y) = 5$. What can you say about the value $f(1, 3)$?

$f(1, 3) = 5$. Since f is continuous (on its entire domain), we have that $\lim_{(x,y)\to(1,3)} f(x, y) = f(1, 3)$.

Example 2.7: MT Section 2.2 Problem 6c

Let

$$f(x, y) = \begin{cases} \frac{xy^3}{x^2+y^6} & (x, y) \neq (0, 0) \\ 0 & (x, y) = (0, 0) \end{cases}$$

Show that f is not continuous at $(0, 0)$.

We have previously found that on the path $x = 0$, that $\lim_{(x,y)\to(0,0)} f(x, y) = (0, 0)$ and that on path $x = y^3$ $\lim_{(x,y)\to(0,0)} f(x, y) = 1/2$, so the limit does not exist, so it cannot be continuous.

Example 2.8: MT Section 2.2 Problem 25

(a) Can
$$\frac{\sin(x+y)}{x+y}$$
be made continuous by suitably defining it at $(0,0)$?

(b) Can
$$\frac{xy}{x^2+y^2}$$
be made continuous by suitably defining it at $(0,0)$?

(c) Prove that $f : \mathbb{R}^2 \mapsto R$, $(x,y) \mapsto ye^x + \sin(x) + (xy)^4$ is continuous.

(a) Yes. Make the substitution $z = x + y$, and this is the limit of $\sin(z)/z$ as z approaches 0, which we know is 1.

(b) No. On the path $x = 0$, this is 0. However, on the path $y = x$, this is $1/2$.

(c) It suffices to show each individual term is continuous, since the a linear combination of continuous functions is continuous. y and e^x are both continuous, so their product is too. $\sin(x)$ is continuous. xy is continuous, so its iterated product is also continuous.

Definition 2.10: Continuity by ε and δ

$$\lim_{x\to x_0} f(x) = x_0$$

if and only if for any $\varepsilon > 0$, there exists some δ such that for any $x \in A$ that satisfies $0 \leq d(x, x_0) < \delta$ we have that $d(f(x), b) < \varepsilon$.

Example 2.9: MT Section 2.2 Problem 33

A function $f : A \subset \mathbb{R}^n \mapsto R^m$ is called Hölder-continuous if there exists positive constants K and α such that for all x and y in A

$$\|f(x) - f(y)\| \leq K\|x - y\|^\alpha$$

In the case where $\alpha = 1$, f is also described as Lipschitz-continuous.
Show that a Hölder-continuous function is continuous.

Worry about expressing δ later. Let $\varepsilon > 0$. Suppose $0 < d(x,y) < \delta$. Let d be the regular norm, so this is really just the equivalent of saying let $\|x - y\| < \delta$. Then we must show that $d(f(x), f(y)) = \|f(x) - f(y)\| < \varepsilon$.

We have that since f is Hölder continuous, $\|f(x)-f(y)\| \leq K\|x-y\|^\alpha$ for some K and α. We want $\|f(x) - f(y)\| < \varepsilon$, and since $\|f(x) - f(y)\| \leq K\|x - y\|^\alpha$, if we can get $\varepsilon \geq K\|x - y\|^\alpha$, we will then have that $\|f(x) - f(y)\| \leq K\|x - y\|^\alpha \leq \varepsilon$.

However, ε is arbitrary and we do not have freedom to pick it. We must find δ in terms of ε

such that this inequality does hold for all ε.

The first main idea is to note that $K\|x-y\|^\alpha$ is expressed in terms of $d(x,y)=\|x-y\|$, and by hypothesis, $0<\|x-y\|<\delta$. The Hölder-continuous condition then gives that $\|f(x)-f(y)\|<K\delta^\alpha$. Then final main idea is that since we have $\|f(x)-f(y)\|<\varepsilon$, and we have that $\|f(x)-f(y)\|<K\delta^\alpha$, we can pick some δ such that $K\delta^\alpha$ is less than ε. If we solve for δ in this inequality, we have that $\delta=\sqrt[\alpha]{\varepsilon/k}$. We can now write the proof up.

We claim that f is continuous. Let $\varepsilon>0$. We claim that $\delta=\sqrt[\alpha]{\varepsilon/k}$ suffices.

Suppose $0<\|x-y\|<\delta$. Then, as f is Hölder-continuous, for all $x,y\in A$, $\|f(x)-f(y)\|\le K\|x-y\|^\alpha$ for some K and α, so

$$\begin{aligned}\|f(x)-f(y)\| &\le K\|x-y\|^\alpha\\ &\le K\delta^\alpha\\ &\le k\left(\sqrt[\alpha]{\frac{\varepsilon}{k}}\right)^\alpha\\ &\le \varepsilon\end{aligned}$$

as desired. Since ε was arbitrary, this holds for all ε. Hence f is continuous.

2.2.4 Compositions

For compositions of functions, if $g(x_0)=y_0$, $f\circ g$ is continuous at x_0 if g is continuous at x_0 and f is continuous at y_0.

This is because if we care about $f\circ g$ at some x_0, and $g(x_0)=y_0$, then when we evaluate $f\circ g$ at x_0, we must evaluate f at y_0.

2.3 "Differentiation" (MT 2.3)

2.3.1 Partial Derivatives

Since we deal with functions that take in multiple variables, there's not really a notion of a single derivative, since the function's derivative is different depending on the variable we choose. To this end, we call a derivative of a function to a particular variable the partial derivative of that function with respect to that variable.

Definition 2.11: Partial derivatives

We define the partial derivatives of f with respect to x_j as the limit quotient

$$\frac{\partial f}{\partial x_j}f(x_1,\dots,x_n)=\lim_{h\to 0}\frac{f(x_1,\dots,x_j+h,\dots,x_n)-f(x_1,\dots,x_n)}{h}.$$

Example 2.10: MT Section 2.3 Problem 2

Find the partial derivatives $\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial y}$ for the given functions at the indicated points:

(a) $z = \sqrt{a^2 - x^2 - y^2}$ at $(0, 0)$ and $(a/2, a/2)$

(b) $z = \log\sqrt{1 + xy}$ at $(1, 2)$ and $(0, 0)$

(c) $z = e^{ax}\cos(bx + y)$ at $(2\pi/b, 0)$.

(a)

$$\begin{aligned}\frac{\partial z}{\partial x} &= \frac{\partial}{\partial x} z \\ &= \frac{\partial}{\partial x}\sqrt{a^2 - x^2 - y^2} \\ &= \frac{1}{2\sqrt{a^2 - x^2 - y^2}} \cdot -2x \\ &= -\frac{x}{\sqrt{a^2 - x^2 - y^2}}.\end{aligned}$$

Then,

$$\begin{aligned}\frac{\partial z}{\partial x}(0, 0) &= -\frac{0}{\sqrt{a^2 - 0^2 - y^2}} \\ &= 0. \\ \frac{\partial z}{\partial x}\left(\frac{a}{2}, \frac{a}{2}\right) &= -\frac{a/2}{\sqrt{a^2 - (a/2)^2 - (a/2)^2}} \\ &= -\frac{a/2}{\sqrt{a^2 - a^2/2}} \\ &= -\frac{a/2}{\sqrt{a^2/2}} \\ &= -\frac{a/2}{|a|/\sqrt{2}} \\ &= -\frac{a\sqrt{2}}{2|a|}\end{aligned}$$

Since z is symmetric across $x = y$, the results are the same for $\partial z/\partial y$.

(b)

$$\begin{aligned}\frac{\partial z}{\partial x} &= \frac{\partial}{\partial x} z \\ &= \frac{\partial}{\partial x}\log\sqrt{1 + xy} \\ &= \frac{1}{\sqrt{1 + xy}} \cdot \frac{1}{2\sqrt{1 + xy}} \cdot y\end{aligned}$$

$$= \frac{y}{2(1+xy)}.$$

Then

$$\begin{aligned}\frac{\partial z}{\partial x}(1,2) &= \frac{2}{2(1+2)}\\ &= \frac{1}{3}.\\ \frac{\partial z}{\partial x}(0,0) &= \frac{0}{2(1+0)}\\ &= 0.\end{aligned}$$

Since z is symmetric across $x = y$, $\partial z/\partial x(0,0) = \partial z/\partial y(0,0) = 0$. However, $\partial z/\partial y(1,2) = \partial z/\partial x(2,1)$, we have to compute that.

$$\begin{aligned}\frac{\partial z}{\partial y}(1,2) &= \frac{\partial z}{\partial x}(2,1)\\ &= \frac{1}{2(1+2)}\\ &= \frac{1}{6}.\end{aligned}$$

(c)

$$\begin{aligned}\frac{\partial z}{\partial x} &= \frac{\partial}{\partial x}z\\ &= \frac{\partial}{\partial x}e^{ax}\cos(bx+y)\\ &= ae^{ax}\cos(bx+y) - be^{ax}\sin(bx+y).\\ \frac{\partial z}{\partial y} &= \frac{\partial}{\partial y}z\\ &= \frac{\partial}{\partial y}e^{ax}\cos(bx+y)\\ &= -e^{ax}\sin(bx+y).\end{aligned}$$

Then,

$$\begin{aligned}\frac{\partial z}{\partial x}\left(\frac{2\pi}{b},0\right) &= ae^{a\cdot\frac{2\pi}{b}}\cos(b\cdot(2\pi/b)+0) - be^{a\cdot\frac{2\pi}{b}}\sin(b\cdot(2\pi/b)+0)\\ &= ae^{2\pi a/b}\\ \frac{\partial z}{\partial y}\left(\frac{2\pi}{b},0\right) &= -e^{a\cdot 2\pi/b}\sin(b\cdot 2\pi/b+0)\\ &= 0.\end{aligned}$$

2.3.2 Linear Approximations and Tangent Lines

In single variable calculus, we made use of tangent lines for linear approximations. In multiple dimensions, we have tangent hyperplanes for linear approximations.

Definition 2.12: Affine approximation/Tangent plane for a function

$$z = f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0)$$

Example 2.11: MT Section 2.3 Problem 5

Find the equation of the plane tangent to the surface

$$z = x^2 + y^3$$

at $(3, 1, 10)$.

$z_x = 2x$ and $z_y = 3y^2$. Then $z_x(3, 1) = 6$ and $z_y = 3$. Then the tangent plane is given as

$$z = 10 + 6(x - 3) + 3(y - 1).$$

Example 2.12: MT Section 2.3 Problem 12

Let

$$f(x, y) = e^{2x+3y}.$$

(a) Find the tangent plane to f at $(0, 0)$
(b) Use this to approximate $f(0.1, 0)$ and $f(0, 0.1)$.

(a) First we compute $f(0, 0) = 1$. Then $f_x = 2e^{2x+3y}$ and $f_y = 3e^{2x+3y}$, so we have that the tangent plane is
$$f(x, y) = 1 + 2(x - 0) + 3(y - 0).$$

(b) $f(0.1, 0) \approx 1 + 0.2 \approx 1.2$, and $f(0, 0.1) \approx 1 + 0.3 \approx 1.3$.

Example 2.13: MT Section 2.3 Problem 13

Where does the plane tangent to $z = e^{x-y}$ at $(1, 1, 1)$ meet the z-axis?

We have that $f_x = e^{x-y}$ and $f_y = -e^{x-y}$. So $f_x(1, 1) = 1$ and $f_y(1, 1) = -1$. This gives the tangent plane
$$z = 1 + 1(x - 1) - 1(y - 1) \implies z = x - y + 1.$$

The z-axis is where $x, y = 0$. This gives $z = 1$.

Example 2.14: MT Section 2.3 Problem 16

Use the linear approximation to approximate a suitable function f and thereby estimate the following:
(a) $(0.99e^{0.02})^8$
(b) $(0.99)^3 + (2.01)^3 - 6(0.99)(2.01)$
(c) $\sqrt{(4.01)^2 + (3.98)^2 + (2.02)^2}$

I won't actually compute the linear approximations.

(a) This looks like $f(x, y) = (xe^y)^8 = x^8e^{8y}$ at $x = 1$ and $y = 0$. $f(1, 0) = 1$. $f_x = 8x^7e^{8y}$, so $f_x(1, 0) = 8$. $f_y = 8x^8e^{8y}$, so $f_y(1, 0) = 8$. This gives the linear approximation $f(x, y) = 1 + 8(x - 1) + 8(y - 0)$. Then $f(0.99, 0.02)$ is approximated as $1 - 0.08 + 0.16$.

(b) This looks like $f(x, y) = x^3 + y^3 - 6xy$ at $(1, 2)$. $f(1, 2) = 1 + 8 - 12 = -3$. $f_x = 3x^2 - 6y$, so $f_x(1, 2) = 3 - 12 = -9$. $f_y = 3y^2 - 6x$, so $f_y(1, 2) = 12 - 6 = 6$. This gives the linear approximation $f(x, y) = -3 - 9(x - 1) + 6(y - 2)$. Then $f(0.99, 2.01)$ is approximated as $-3 + 0.09 + 0.06$.

(c) This looks like $f(x, y, z) = \sqrt{x^2 + y^2 + z^2}$ at $(x, y, z) = (4, 4, 2)$. Then $f(4, 4, 2) = \sqrt{16 + 16 + 4} = \sqrt{36} = 6$. $f_x = x/\sqrt{x^2 + y^2 + z^2}$ and the same for y and z. Then $f_x(4, 4, 2) = 2/3$, $f_y(4, 4, 2) = 2/3$ and $f_z(4, 4, 2) = 1/3$. This gives the linear approximation $f(x, y, z) = 6 + 2/3(x - 4) + 2/3(y - 4) + 1/3(z - 2)$. Then $f(4.01, 3.98, 2.02) = 6 + 2/3(0.1) - 2/3(0.02) + 1/3(0.02)$.

That a function f has partial derivatives at some point is insufficient to say f is differentiable at that point. We also require that the affine approximation is a good approximation. In single variable calculus, this constraint is the same as saying that the linear approximation is a good approximation, which is actually implied by the existence of the derivative. Here, we are saying that not only are single variable changes are good approximations, but combinations of the variables too.

2.3.3 Derivative Matrix

Since partial derivatives are only partial, we discuss a notion of a function's derivative that is "entire", and holds all the information of the derivative, by means of a matrix.

Definition 2.13: Derivative Matrix

For a given function $f : \mathbb{R}^n \mapsto \mathbb{R}^m$, $f = (f_1, \dots, f_m)$, we define the derivative matrix of f, denoted $\mathbf{D}f$, as the $m \times n$ matrix

$$\mathbf{D}f = \left[\frac{\partial}{\partial x_j} f_i\right] = \begin{bmatrix} \frac{\partial}{\partial x_1} f_1 & \dots & \frac{\partial}{x_n} f_1 \\ \vdots & \ddots & \vdots \\ \frac{\partial}{\partial x_1} f_m & \dots & \frac{\partial}{\partial x_n} f_m \end{bmatrix}$$

Where for a given row, the function remains constant, and for a given column, the variable remains constant.

Example 2.15: MT Section 2.3 Problem 9

Compute the matrix of partial derivatives of the following functions:
(a) $f(x, y) = (x, y)$
(b) $f(x, y) = (xe^y + \cos y, x, x + e^y)$
(c) $f(x, y, z) = (x + e^z + y, yx^2)$
(d) $f(x, y) = (xye^{xy}, x\sin(y), 5xy^2)$

(a)

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

(b)

$$\begin{bmatrix} e^y & -\sin(y) \\ 1 & 0 \\ 1 & e^y \end{bmatrix}$$

(c)

$$\begin{bmatrix} 1 & 1 & e^z \\ 2xy & x^2 & 0 \end{bmatrix}$$

(d)

$$\begin{bmatrix} ye^{xy} + xy^2e^{xy} & xe^{xy} + x^2ye^{xy} \\ \sin(y) & x\cos(y) \\ 5y^2 & 10xy \end{bmatrix}$$

Example 2.16: MT Section 2.3 Problem 17

Let P be the tangent plane to the graph of $g(x, y) = 8 - 2x^2 - 3y^2$ at the point $(1, 2, -6)$. Let $f(x, y) = 4 - x^2 - y^2$. Find the point on the graph of f which has tangent plane parallel to P.

Since the tangent plane is of form $z = f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0)$, if you multiply things through, you're left with $f_x(x_0, y_0)x + f_y(x_0, y_0)y - z + D = 0$, which gives a normal vector of $(f_x(x_0, y_0), f_y(x_0, y_0), -1)$.

$g_x = -4x$ and $g_y = -6y$, so $g_x(1,2) = -4$ and $g_y(1,2) = -12$, so P has normal vector $(-4, -12, -1)$. $f_x = -2x$ and $f_y = -2y$, so the tangent plane at a given x, y will have the normal vector $(-2x, -2y, -1)$, of which the two normal vectors are equal at $x = 2$ and $y = 6$.

2.3.4 Introduction to Gradients

For the special case of $f : \mathbb{R} \mapsto \mathbb{R}^m$, we define the *gradient*, which we will revisit later with the del operator:

Definition 2.14: Gradient

Let f be a function $f : \mathbb{R} \mapsto \mathbb{R}^m$. We define the *gradient* of f, denoted grad f or ∇f, as

$$\mathbf{D}f = \left(\frac{\partial}{\partial x_1} f, \ldots, \frac{\partial}{\partial x_m} f\right).$$

Really, this is a column vector, so I should put a transpose there, but since literally every vector is actually a column vector, so I'll lazily say that parantheses notes a column vector, but written as a row vector.

2.3.5 Differentiability and continuity

Theorem 2.1: Differentiability implies continuity

Let $f : U \subset \mathbb{R}^n \mapsto \mathbb{R}^m$ be differentiable at $x_0 \in U$. Then f is continuous at x_0.

Theorem 2.2: Existence and continuity of partials implies differentiable

Let $f : U \subset \mathbb{R}^n \mapsto \mathbb{R}^m$, $f = (f_1, \ldots, f_m)$. Suppose each and every partial $\frac{\partial f_i}{x_j}$ exists and are continuous in a neighborhood around a point $x \in U$. Then f is differentiable at x.

2.4 "Properties of Derivatives" (MT 2.5)

Instead of $\frac{d}{dx}$, using $\mathbf{D}$, all of the original derivative rules still hold. That is,

- $\mathbf{D}(cf(x)) = c\mathbf{D}f(x)$
- $\mathbf{D}(f(x) + g(x)) = \mathbf{D}f(x) + \mathbf{D}g(x)$
- $\mathbf{D}(f(x)g(x)) = g(x)\mathbf{D}f(x) + f(x)\mathbf{D}g(x)$
- $\mathbf{D}(f(x)/g(x)) = (g(x)\mathbf{D}f(x) - f(x)\mathbf{D}g(x))/g(x)^2$

2.4.1 Chain rule

The only tricky one is the chain rule.

Theorem 2.3: Chain Rule

Let $U \subset \mathbb{R}^n$ and $V \subset \mathbb{R}^m$ be open sets. Let $g : U \subset \mathbb{R}^n \mapsto \mathbb{R}^m$ and $f : V \subset \mathbb{R}^m \mapsto \mathbb{R}^p$ be given functions such that g maps U to V so that $f \circ g$ is defined.
Suppose g is differentiable at x_0 and f is differentiable at $y_0 = g(x_0)$. Then $f \circ g$ is differentiable at x_0 and

$$\mathbf{D}(f \circ g)(x_0) = \mathbf{D}f(y_0)\mathbf{D}g(x_0).$$

The conditions for this are specific. If f is not (totally) differentiable, the chain rule is not applicable. Consider the following example:

Example 2.17: MT Section 2.5 Problem 22

Consider the function

$$f(x,y) = \begin{cases} \frac{xy^2}{x^2+y^2} & (x,y) \neq (0,0) \\ 0 & (x,y) = (0,0). \end{cases}$$

Show that

(a) f_x and f_y both exist at $(0,0)$

(b) If $g(t) = (at, bt)$ for constants a and b, then $f \circ g$ is differentiable and $(f \circ g)'(0) = (ab^2)/(a^2+b^2)$ but $\nabla f(0,0) \cdot g'(0) = 0$.

(a) On first glance, we might try to take partial derivatives, which give $f_x = \frac{(x^2+y^2)(y^2)-xy^2(2x)}{(x^2+y^2)^2}$ and $f_y = \frac{(x^2+y^2)(2xy)-xy^2(2y)}{(x^2+y^2)^2}$. However, these give indeterminate forms at $(0,0)$, so we instead use the limit quotient definitions for these:

$$\begin{aligned} \frac{\partial f}{\partial x}(0,0) &= \lim_{h\to 0} \frac{f(0+h,0) - f(0,0)}{h} \\ &= \lim_{h\to 0} \frac{h0^2/(h^2+0^2) - 0}{h} \\ &= 0. \end{aligned}$$

and

$$\begin{aligned} \frac{\partial f}{\partial y}(0,0) &= \lim_{h\to 0} \frac{f(0,0+h) - f(0,0)}{h} \\ &= \frac{0h^2/(0+h^2) - 0}{h} \\ &= 0. \end{aligned}$$

(b) Though not actually pertinent to the mechanics of the problem, we will first require that $g(t)$ is actually a line, so at least one of a, b must not be zero. To make the "moral" of the problem true, neither can be zero.

$$\begin{aligned} f(g(t)) &= \frac{at(bt)^2}{(at)^2 + (bt)^2} \\ &= \frac{ab^2t^3}{a^2t^2 + b^2t^2} \\ &= \frac{abt}{a^2 + b^2} \end{aligned}$$

Differentiating,

$$\begin{aligned} \frac{d}{dt} f(g(t)) &= \frac{d}{dt} \frac{abt}{a^2 + b^2} \\ &= \frac{ab}{a^2 + b^2}. \end{aligned}$$

So $f(g(t))$ is $ab/(a^2 + b^2)$ at $t = 0$.

$\nabla f = (f_x, f_y)$, so $\nabla f(0,0) = (f_x(0,0), f_y(0,0)) = (0,0)$. $g'(0) = (a,b)$, so $\nabla f(0,0) \cdot g'(0) = (0,0) \cdot (a,b) = 0$.

Example 2.18: MT Section 2.5 Problem 8

Let

$$f(u,v,w) = (e^{u-w}, \cos(u+v) + \sin(u+v+w))$$

and

$$g(x,y) = (e^x, \cos(y-x), e^{-y}).$$

Calculate $f \circ g$ and $\mathbf{D}(f \circ g)(0,0)$

$$f \circ g = e^{e^x - e^{-y}}, \cos(e^x + \cos(y-x)) + \sin(e^x + \cos(y-x) + e^{-y})$$

If there were ever a reason to want to use the chain rule as opposed to evaluating the function and then differentiating it, this would be it.

$$\mathbf{D}g = \begin{bmatrix} e^x & \sin(y-x) & 0 \\ 0 & -\sin(y-x) & -e^{-y} \end{bmatrix}$$

and

$$\mathbf{D}f = \begin{bmatrix} e^{u-w} & -\sin(u+v) + \cos(u+v+w) \\ 0 & -\sin(u+v) + \cos(u+v+w) \\ -e^{u-w} & \cos(u+v+w) \end{bmatrix}$$

$g(0,0) = (1,1,1)$.

$$\mathbf{D}g(0,0) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

$$\mathbf{D}f(1,1,1) = \begin{bmatrix} 1 & -\sin(2)+\cos(3) \\ 0 & -\sin(2)+\cos(3) \\ -1 & \cos(3) \end{bmatrix}$$

Multiplying the two gives

$$\mathbf{D}f(g(x,y)) = \begin{bmatrix} 1 & 0 & \sin(2)-\cos(3) \\ 0 & 0 & \sin(2)-\cos(3) \\ -1 & 0 & -\cos(3) \end{bmatrix}$$

2.4.2 First special case of the chain rule

Theorem 2.4: First special case of the chain rule

Let $c(t)$ be a path in $\mathbb{R}^3$. Then $h(t) = f(c(t))$ is given by

$$\frac{dh}{dt} = \mathbf{D}f(c(t))\mathbf{D}c(t) = \nabla f(c(t)) \cdot c'(t) = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} + \frac{\partial f}{\partial z}\frac{dz}{dt}.$$

Example 2.19: MT Section 2.5 Problem 3a

Verify the first special case of the chain rule for the composition $f \circ c$ for $f(x,y) = xy$ and $c(t) = (e^t, \cos(t))$.

By evaluating, $f(c(t)) = e^t\cos(t)$. Then $\frac{\partial}{\partial t}f(c(t)) = e^t\cos(t) - e^t\sin(t)$.

By the first special case, we have that

$$\mathbf{D}f = \begin{bmatrix} y & x \end{bmatrix}$$

and so $\mathbf{D}f(c(t)) = (\cos(t), e^t)$. Then

$$\mathbf{D}c = \begin{bmatrix} e^t \\ -\sin(t) \end{bmatrix}$$

The product of these two matrices is then $e^t\cos(t) - e^t\sin(t)$, which agrees.

Example 2.20: MT Section 2.5 Problem 10

Suppose that the temperature at the point (x,y,z) in space is $T(x,y,z) = x^2+y^2+z^2$. Let a particle follow the right-circular helix $\sigma(t) = (\cos(t), \sin(t), t)$ and let $T(t)$ be its temperature at time t.

(a) What is $T'(t)$?

(b) Find an approximate value for the temperature at $t = (\pi/2) + 0.01$.

This example is simple enough to not warrant the chain rule, but we'll use it anyways.

(a) $\mathbf{D}f = (2x, 2y, 2z)$, and so $\mathbf{D}f(\sigma(t)) = (2\cos(t), 2\sin(t), 2t)$. $\mathbf{D}\sigma = (-\sin(t), \cos(t), 1)$. By the first special case of the chain rule, $\mathbf{D}(f\circ\sigma)(t) = -2\sin(t)\cos(t) + 2\sin(t)\cos(t) + 2t = 2t$.

(b) Using our previous result, $T'(\pi/2) = \pi$, and $T(\pi/2) = \pi^2/4$. This gives a linear approximation of $\pi^2/4+(t-\pi/2)\pi$, so we approximate the temperature at $t = (\pi/2)+0.01$ as $\pi^2/4 + 0.01\pi$.

Example 2.21: MT Section 2.5 Problem 13

Suppose that a duck is swimming in the circle $x = \cos(t), y = \sin(t)$ and that the water temperature is given by the formula

$$T = x^2e^y - xy^3.$$

Find dT/dt, the rate of change in temperature the duck might feel:
(a) by the chain rule
(b) by expressing T in terms of t and differentiating.

Call the position of the duck $r(t) = (\cos(t), \sin(t))$.

(a) $\mathbf{D}T = (2xe^y - y^3, x^2e^y - 3xy^2)$, so $\mathbf{D}T(r) = (2\cos(t)e^{\sin(t)} - \sin^3(t), \cos^2(t)e^{\sin(t)} - 3\cos(t)\sin^2(t))$. $\mathbf{D}r = (-\sin(t), \cos(t))$. Then

$$\mathbf{D}T(r(t)) = \sin^4(t) - 2\sin(t)\cos(t)e^{\sin(t)} + \cos^3(t)e^{\sin(t)} - 3\cos^2(t)\sin^2(t).$$

(b) $T(r(t)) = \cos(t)^2e^{\sin(t)} - \cos(t)\sin^3(t)$. Then

$$\frac{d}{dt}T(r(t)) = -2\sin(t)\cos(t)e^{\sin(t)} + \cos^3(t)e^{\sin(t)} + \sin^4(t) - 3\sin^2(t)\cos^2(t)$$

2.4.3 Second special case of the chain rule

Theorem 2.5: Second special case of the chain rule

et $f : \mathbb{R}^3 \mapsto \mathbb{R}$ and let $g : \mathbb{R}^3 \mapsto \mathbb{R}^3$. Then let $g(x, y, z) = (u(x, y, z), v(x, y, z), w(x, y, z))$ and let $h(x, y, z) = f(g(x, y, z))$. Then

$$\begin{bmatrix} \frac{\partial h}{\partial x} & \frac{\partial h}{\partial y} & \frac{\partial h}{\partial z} \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial u} & \frac{\partial f}{\partial v} & \frac{\partial f}{\partial w} \end{bmatrix} \begin{bmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} & \frac{\partial u}{\partial z} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} & \frac{\partial v}{\partial z} \\ \frac{\partial w}{\partial x} & \frac{\partial w}{\partial y} & \frac{\partial w}{\partial z} \end{bmatrix}$$

Example 2.22: MT Section 2.5 Problem 17

Write out the chain rule for $\partial h/\partial x$ each of the following functions and justify your answer in each case using the chain rule.
(a) $h(x, y) = f(x, u(x, y))$
(b) $h(x) = f(x, u(x), v(x))$
(c) $h(x, y, z) = f(u(x, y, z), v(x, y), w(x))$.

I'll list the expressions, and then their justifications.

(a) $\frac{\partial h}{\partial x} = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}\frac{\partial u}{\partial x}$

(b) $\frac{\partial h}{\partial x} = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}\frac{\partial u}{\partial x} + \frac{\partial f}{\partial z}\frac{\partial v}{\partial x}$

(c) $\frac{\partial h}{\partial x} = \frac{\partial f}{\partial x}\frac{\partial u}{\partial x} + \frac{\partial f}{\partial y}\frac{\partial v}{\partial x} + \frac{\partial f}{\partial z}\frac{\partial w}{\partial x}$

The structure of the second special case of the chain rule is repeatedly used here.

$$\begin{aligned} \mathbf{D}h &= \begin{bmatrix} \frac{\partial h}{\partial x} & \frac{\partial h}{\partial y} \end{bmatrix} \\ &= \begin{bmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \end{bmatrix} \begin{bmatrix} \frac{\partial x}{\partial x} & \frac{\partial x}{\partial y} \\ \frac{\partial u}{x} & \frac{\partial u}{\partial y} \end{bmatrix} \\ &= \begin{bmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ \frac{\partial u}{x} & \frac{\partial u}{\partial y} \end{bmatrix} \\ &= \begin{bmatrix} \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}\frac{\partial u}{\partial x} & \frac{\partial f}{\partial y}\frac{\partial u}{\partial y} \end{bmatrix} \\ \implies \frac{\partial h}{\partial x} &= \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}\frac{\partial u}{\partial x} \end{aligned} \tag{a}$$

Similarly,

$$\begin{aligned} \mathbf{D}h &= \begin{bmatrix} \frac{\partial h}{\partial x} \end{bmatrix} \\ &= \begin{bmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} & \frac{\partial f}{\partial z} \end{bmatrix} \begin{bmatrix} \frac{\partial x}{\partial x} \\ \frac{\partial u}{\partial x} \\ \frac{\partial v}{\partial x} \end{bmatrix} \\ &= \begin{bmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} & \frac{\partial f}{\partial z} \end{bmatrix} \begin{bmatrix} 1 \\ \frac{\partial u}{\partial x} \\ \frac{\partial v}{\partial x} \end{bmatrix} \end{aligned} \tag{2}$$

and

$$\begin{aligned} \mathbf{D}h &= \begin{bmatrix} \frac{\partial h}{\partial x} & \frac{\partial h}{\partial y} & \frac{\partial h}{\partial z} \end{bmatrix} \\ &= \begin{bmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} & \frac{\partial f}{\partial z} \end{bmatrix} \begin{bmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} & \frac{\partial u}{\partial z} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} & \frac{\partial v}{\partial z} \\ \frac{\partial w}{\partial x} & \frac{\partial w}{\partial y} & \frac{\partial w}{\partial z} \end{bmatrix} \\ &= \begin{bmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} & \frac{\partial f}{\partial z} \end{bmatrix} \begin{bmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} & \frac{\partial u}{\partial z} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} & 0 \\ \frac{\partial w}{\partial x} & 0 & 0 \end{bmatrix} \end{aligned} \tag{c}$$

2.4.4 A weak implicit function theorem

Theorem 2.6: The weak implicit function theorem

Suppose we have a relation of $F(x, y) = 0$. Then we have that

$$\frac{\partial y}{\partial x} = -\frac{F_x}{F_y}$$

if both F_x and F_y are well-defined at the desired point.

This is proved in the following example:

Example 2.23: MT Section 2.5 Problem 19a, c

(a) Let $y(x)$ be implicitly defined by $G(x, y(x)) = 0$, where G is a given function of two variables. Prove that if $y(x)$ and G are differentiable, then the weak implicit function theorem holds.

(c) Let y be defined implicitly by $x^2 + y^3 + e^y = 0$. Compute $\frac{dy}{dx}$.

(a) Differentiating both sides with respect to x using the chain rule gives

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

Solving for $\frac{\partial y}{\partial x}$ gives the desired statement.

(c) Using the weak implicit function theorem, $F_x = 2x$ and $F_y = 3y^2$, so $\frac{\partial y}{\partial x} = -2x/3y^2$.

2.5 "Introduction to Paths and Curves" (MT 2.4)

Paths are a rigorous way to view the idea of "getting from point a to point b" in any arbitrary way. To this end, we have an idea of a position as a function of time, moving through space. Drawing an arbitrarily swiggly line from point a to point b is the image of that path. However, as we have discussed before, things are not unique, so even though we have a position function for time, to be complete, we must also equip it with a time interval in which it is defined for.

Definition 2.15: Paths and Curves

A path in $\mathbb{R}^n$ is a map $c : [a, b] \mapsto \mathbb{R}^n$.
The collection of points $c(t)$ for $t \in [a, b]$, denoted C, is called a curve, and we may call $c(a)$ and $c(b)$ the endpoints.

2.5.1 Velocity and tangent vectors

As we have discussed before, one principle way objects can vary are how fast something is traced out, with respect to unit time.

Definition 2.16: Velocity vector

If c is a path and is differentiable, then c is a differentiable path, and the velocity of c at time t, denoted $c'(t)$ is defined by the limit quotient

$$c'(t) = \lim_{h \to 0} \frac{c(t+h) - c(t)}{h}.$$

When drawing, we put the tail of c' at the point $c(t)$. We can then define the speed as $s = \|c'(t)\|$.

Example 2.24: MT Section 2.4 Problem 7

Determine the velocity vector of the path

$$c(t) = 6t\mathbf{i} + 3t^2\mathbf{j} + t^3\mathbf{k}.$$

In vector form, this is $c(t) = (6t, 3t^2, t^3)$ which has velocity vector $c'(t) = (6, 6t, 3t^2)$.

Example 2.25: MT Section 2.4 Problem 16

If the position of a particle in space is $(6t, 3t^2, t^3)$ at time t, what is its velocity vector at $t = 0$?

Let the path be $c(t) = (6t, 3t^2 + t^3)$, then $c'(t) = (6, 6t, 3t^2)$ and then $c'(0) = (6, 0, 0)$.

Theorem 2.7: Velocity is tangent to the path

he velocity vector $c'(t)$ is tangent to the path $c(t)$ at time t.
If C is the curve traced out by c and if $c'(t) \neq 0$, then c' is a vector tangent to C at the point $c(t)$.

Example 2.26: MT Section 2.4 Problem 13

Compute the tangent vector to the path

$$c(t) = (t\sin(t), 4t).$$

$$c'(t) = (\sin(t) + t\cos(t), 4).$$

2.5.2 Tangent Lines for Paths

Of course, whenever we define a notion of a derivative, it is accompanied by a linear approximation and tangent line:

Definition 2.17: Tangent line to a path

If $c(t)$ is a path and $c'(t_0) \neq 0$, the equation of the tangent line at the point $t(c_0)$, denoted as $l(t)$ is given by

$$l(t) = c(t_0) + (t - t_0)c'(t_0).$$

Example 2.27: MT Section 2.4 Problem 17

Determine the equation of the tangent line to the given path

$$(\sin 3t, \cos 3t, 2t^{5/2})$$

at $t = 1$.

$c(1) = (\sin(3), \cos(3), 2)$. $c'(t) = (3\cos(3t), -3\sin(3t), 5t^{3/2})$, so $c'(1) = (3\cos(3), -3\sin(3), 5)$. Then the tangent line is given by

$$(\sin(3), \cos(3), 2) + (t - 1)(3\cos(3), -3\sin(3), 5).$$

Example 2.28: MT Section 2.4 Problem 19

Suppose that a particle following the given path

$$c(t) = (t^2, t^3 - 4t, 0)$$

flies off on a tangent at $t = 2$. Compute the position of the particle at $t = 3$.

$c(2) = (4, 0, 0)$. $c'(t) = (2t, 3t^2 - 4, 0)$, so $c'(2) = (4, 8, 0)$. This gives the tangent LINE

$$(4, 0, 0) + (t - 2)(4, 8, 0)$$

which when evaluated at $t = 3$ gives $(8, 8, 0)$.

Example 2.29: MT Section 2.4 Problem 23

The position vector for a particle moving on a helix is

$$c(t) = (\cos(t), \sin(t), t^2).$$

(a) Find the speed of the particle at time $t_0 = 4\pi$.
(b) Is $c'(t)$ ever orthogonal to $c(t)$?
(c) Find a parameterization for the tangent line to $c(t)$ at $t_0 = 4\pi$.
(d) Where will this line intersect the xy-plane?

(a) $c'(t) = (-\sin(t), \cos(t), 2t)$, and $c'(4\pi) = (0, 1, 8\pi)$, so the speed is given as $s = \|c'(t)\|$ at $t = 4\pi$, so $s = \sqrt{1 + 64\pi}$.

(b) If $c'(t)$ and $c(t)$ are non-zero (which they always are) and are orthogonal, then $c(t) \cdot c'(t) = 0$.

$$\begin{aligned} c(t) \cdot c'(t) &= (\cos(t), \sin(t), t^2) \cdot (-\sin(t), \cos(t), 2t) \\ &= -\sin(t)\cos(t) + \sin(t)\cos(t) + 2t^3 \\ &= 2t^3 \end{aligned}$$

This is zero when $t = 0$.

(c) $c(4\pi) = (1, 0, 16\pi^2)$.

$$l(t) = (1, 0, 16\pi^2) + (t - 4\pi)(0, 1, 8\pi).$$

(d) This intersects the xy-plane when $z = 0$. So we are interested in finding the t such that $16\pi^2 + (t - 4\pi)8\pi = 0$. We want the $(4 - 4\pi)$ term to be -2π, so $t = 2\pi$. Evaluating l then gives $(1, -2\pi, 0)$.

2.6 "Gradients and Directional Derivatives" (MT 2.6)

Partial derivatives give a notion of how a function changes at a given point in the x or y direction. Though useful, we may want to know the rate of change in some other direction, perhaps something more angled. We may want to know the rate of change if one walks diagonally on the line $y = x$ in the positive direction, or any other angle combination. One approach is to construct a path that walks in the desired direction, and then take that derivative, which is more or less how this idea is defined.

Definition 2.18: Directional Derivative

The directional derivative of f at x along the vector v is given by

$$\frac{d}{dt} f(x + tv)\bigg|_{t=0}.$$

where v is generally enforced to be a unit vector.

This is generally not the most useful formula to compute these. In contrast,

Theorem 2.8: Directional Derivatives by Gradients

If $f : \mathbb{R}^m \mapsto \mathbb{R}$ is differentiable, then all directional derivatives exist and are equal to

$$\mathbf{D}f(x)v = \text{grad } f(x) \cdot v = \nabla f(x) \cdot v$$

I find that this (and the chain rule) are the instances in vector calculus where the machinery of linear algebra render the ideas extremely clear, although in neither case are they really presented nor important. You can ignore the derivation without recourse. I always remember this by noting that if v is a unit vector used for direction, the change is the sum of changes in the x and y directions – the partial derivatives – scaled by angle.

Example 2.30: MT Section 2.6 Problem 3

Compute the directional derivatives of the following functions along unit vectors at the indicated points in directions parallel to the given vector:

(a) $f(x,y) = x^y$, $(x_0, y_0) = (e, e)$, $d = 5\mathbf{i} + 12\mathbf{j}$

(b) $f(x,y,z) = e^x + yz$, $(x_0, y_0, z_0) = (1, 1, 1)$, $d = (1, -1, 1)$.

(c) $f(x,y,z) = xyz$, $(x_0, y_0, z_0) = (1, 0, 1)$, $d = (1, 0, -1)$

Let's first find the partial derivatives.

(a) $(f_x, f_y) = (yx^{y-1}, x^y \ln(x))$

(b) $(f_x, f_y, f_z) = (e^x, z, y)$

(c) $(f_x, f_y, f_z) = (yz, xz, xy)$

Now, evaluate,

(a) $\operatorname{grad} f(e,e) = (ee^{e-1}, e^e \ln(e)) = (e^e, e^e)$

(b) $\operatorname{grad} f(1,1,1) = (e, 1, 1)$

(c) $\operatorname{grad} f(1,0,1) = (0, 1, 0)$

Now we normalize the vectors,

(a) $n = 1/13(5, 12)$

(b) $n = 1/\sqrt{3}(1, -1, 1)$

(c) $n = 1/\sqrt{2}(1, 0, -1)$

Now we can take the dot product of the partial derivatives and the directions,

(a) $e^e/13(5, 12)$

(b) $1/\sqrt{3}(e, -1, 1)$

(c) 0

2.6.1 Direction and Magnitude of Fastest Increase

Theorem 2.9: The gradient is in the direction of fastest increase

If $\nabla f(x) \neq 0$, then $\nabla f(x)$ points in the direction in which f is increasing the fastest.

Theorem 2.10: The fastest increase is the directional derivative's magnitude

Let $f : \mathbb{R}^n \mapsto \mathbb{R}$, and $x_0 \in \mathbb{R}^n$. If v is a unit vector in $\mathbb{R}^n$, the maximum value of the directional derivative of f at x_0 along v is

$$\|\nabla f(x_0)\|.$$

Both of these are proved in the following example.

Example 2.31: Section 2.6 Problem 5

(a) Prove that the fastest increase is the directional derivative's magnitude.
(b) Let $f(x, y, z) = x^3 - y^3 + z^3$. Find the maximum value for the directional derivative of f at the point $(1, 2, 3)$.

(a) Since the directional derivative is equivalent to $\nabla f(x) \cdot v$, which is a dot product expression, we can use the dot product-angle/cosine similarity formula:

$$\nabla f(x) \cdot v = \|\nabla f(x)\| \|v\| \cos(\theta)$$

where θ is the inner angle between $\nabla f(x)$ and v.

The key observation is that in trying to maximize the right hand side (and thus the left hand side), $\|\nabla f(x)\|$ is constant, since we cannot affect $\nabla f(x)$, and $\|v\|$ is also constant, since v is a unit vector, so $\|v\| = 1$. The only degree of freedom is the θ in $\cos(\theta)$. This is maximized when $\cos(\theta) = 1$, when $\theta = 0$, mandating that the directional derivative is maximized when parallel to the gradient vector.

Cleaning things up, we have

$$\nabla f(x) \cdot v \leq \|\nabla f(x)\|.$$

Taking the norms, we have

$$\|\nabla f(x) \cdot v\| \leq \|\nabla f(x)\|$$

with equality (and maximizing) when $\theta = 0$.

(b) Taking the gradient gives $\nabla f = (3x^2, -3y^2, 3z^2)$. Evaluating gives $\nabla f(1, 2, 3) = (3, -12, 27)$. Then the maximum value for the directional derivative is given by $\|(3, -12, 27)\| = \sqrt{81 + 144 + 729} = \sqrt{954} = 3\sqrt{106}$.

Example 2.32: MT Section 2.6 Problem 22a

Captain Ralph is in trouble near the sunny side of Mercury. The temperature of the ship's hull when he is at location (x, y, z) will be given by

$$T(x, y, z) = e^{-x^2-2y^2-3z^2}$$

where x, y and z are measured in meters. He is currently at $(1, 1, 1)$.
In what direction should he proceed in order to decrease the temperature most rapidly?

Computing the gradient gives

$$\begin{aligned}\nabla T &= \nabla e^{-x^2-2y^2-3z^2} \\ &= (-2xe^{-x^2-2y^2-3z^2}, -4ye^{-x^2-2y^2-3z^2}, -6ze^{-x^2-2y^2-3z^2})\end{aligned}$$

$$= 2e^{-x^2-2y^2-3z^2}(-x, -2y, -3z).$$

The $e^{-x^2-2y^2-3z^2}$ is a scalar, and does not affect the direction of $(-x, -2y, -3z)$. Then the direction of the fastest increase is in the direction of $(-x, -2y, -3z)$. Then the direction of fastest decrease is the opposite, the direction of $(x, 2y, 3z)$. As he is currently at $(1, 1, 1)$, this is the direction of $\nabla f(1, 1, 1) = (1, 2, 3)$.

2.6.2 Gradients and Tangent Planes to Level Sets

Theorem 2.11: The gradient is normal to the level surface

Let f be a differentiable map and let (x_0, y_0, z_0) lie on the level surface S defined by $f(, x, y, z) = k$ for some constant k. Then $\nabla f(x_0, y_0, z_0)$ is normal to the level surface in that if v is the tangent vector of a path $c(t)$ on S with $c(0) = (x_0, y_0, z_0)$, then $\nabla f(x_0, y_0, z_0) \cdot v = 0$.

Example 2.33: MT Section 2.6 Problem 4

You are walking on the graph $f(x, y) = y\cos(\pi x) - x\cos(\pi y) + 10$, standing at the point $(2, 1, 13)$. Find an x, y-direction you should walk in to stay at the same level.

Taking the gradient,

$$\nabla f = (-\pi y\sin(\pi x) - \cos(\pi y), \cos(\pi x) + \pi x\sin(\pi y))$$

Evaluating,

$$\begin{aligned}\nabla f(2, 1) &= (-\pi 1\sin(2\pi) - \cos(\pi 1), \cos(\pi 2) + \pi 2\sin(\pi 1)) \\ &= (1, 1)\end{aligned}$$

I can handwave and say that we wish to be orthogonal the direction of fastest increase, and so we have $(-1, 1)$ and $(1, -1)$, Instead, note that we are interested in some vector v such that $\nabla f(2, 1) \cdot v = 0$; the directional derivative of the "height" is 0 and hence we stay the same level. Let $v = (x, y)$, then $\nabla f(2, 1) \cdot v = (1, 1) \cdot (x, y) = x + y$, which we wish to be zero, so $x = -y$, as desired.

Example 2.34: Section 2.6 Problem 7

Find the rate of change of $f(x, y, z) = xyz$ in the direction normal to the surface $yx^2 + xy^2 + yz^2 = 3$ at $(1, 1, 1)$.

We first find the gradient of f, $\nabla f = (yz, xz, xy)$. Evaluating, $\nabla f(1, 1, 1) = (1, 1, 1)$.

We find the gradient of the surface, given as $(2xy + y^2, x^2 + 2xy + z^2, 2yz)$. At $(1, 1, 1)$, this is $(3, 4, 2)$.

Then, $\nabla f \cdot (3, 4, 2)/\sqrt{9 + 16 + 4} = 9/\sqrt{29}$

Example 2.35: MT Section 2.6 Problem 31

Suppose that a particle is ejected from the surface $x^2 + y^2 - z^2 = -1$ at the point $(1, 1, \sqrt{3})$ along the normal directed toward the xy-plane to the surface at time $t = 0$ with a speed of 10 units per second. When and where does it cross the xy-plane?

Finding the normal, we compute the gradient of the surface, $\nabla(x^2 + y^2 - z^2) = (2x, 2y, -2z)$. At the point $(1, 1, \sqrt{3})$, this gives $(2, 2, -2\sqrt{3})$. Since $(1, 1, \sqrt{3})$ is above the xy-plane and the z-component of our normal is negative, these two are consistent – going in the direction of the normal goes towards the xy-plane.

Normalizing gives $1/\sqrt{4+4+12}(2, 2, -2\sqrt{3}) = 1/2\sqrt{5}(2, 2, -2\sqrt{3}) = 1/\sqrt{5}(1, 1, -\sqrt{3})$. As we travel at 10 units per second, our velocity vector is $2\sqrt{5}(1, 1, -\sqrt{3})$. In tandem with our initial position of $(1, 1, \sqrt{3})$, the z position is given as $z(t) = \sqrt{3} - 2\sqrt{3}\sqrt{5}t$. The xy-plane, $z = 0$, is crossed when $t = 1/2\sqrt{5} = \sqrt{5}/10$.

Similarly, we have that $x(t) = y(t) = 1 + 2\sqrt{5}t$, which gives that $x(\sqrt{5}/10) = y(\sqrt{5}/10) = 2$.

Alternatively, ignore the velocity vector business, use a regular t, explicitly noting its not time, find the intersection with the xy-plane, compute the distance, and then divide by the velocity.

Gradients provide yet another means to construct a tangent plane.

Definition 2.19: Tangent Plane to a Level Surface

Let S be the level surface given by $f(x, y, z) = k$. The tangent plane of S at a point (x_0, y_0, z_0) is defined by the equation

$$\nabla f(x_0, y_0, z_0) \cdot (x - x_0, y - y_0, z - z_0)$$

if the gradient is defined there.

We prove that this is consistent with previous discussions of tangent planes in a special case in an example below.

Example 2.36: MT Section 2.6 Problem 15

Show that the tangent plane to a level surface definition is consistent with the formula for the plane tangent to the graph of $f(x, y)$ by regarding the graph as a level surface of $F(x, y, z) = f(x, y) - z$.

The regular formula is

$$f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0) - z + f(x_0, y_0) = 0.$$

Hence, our goal is to use the tangent plane of a level surface formula and obtain the regular formula.

Taking S to be the level surface defined as $F(x, y, z) = k$ and applying the gradient gives

$$\begin{aligned}\nabla F &= \nabla f(x, y) - z \\ &= (f_x, f_y, -1).\end{aligned}$$

Evaluating gives $\nabla F(x_0, y_0, z_0) = (f_x(x_0, y_0), f_y(x_0, y_0), -1)$. Then, we have that the tangent plane to the level surface is given as

$$f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0) - 1(z - z_0) = 0.$$

Noting that (x_0, y_0, z_0) is given by $(x_0, y_0, f(x_0, y_0))$, the two are equal.

Example 2.37: MT Section 2.6 Problem 8a

Find the plane tangent to the surface $x^2 + 2y^2 + 3xz = 10$ at the point $(1, 2, 1/3)$.

Computing the gradient gives $\nabla f = (2x + 3z, 4y, 3x)$. Evaluating gives $\nabla f(1, 2, 1/3) = (2 + 1, 8, 1) = (3, 8, 3)$. The plane is then given as

$$3(x - 1) + 8(y - 2) + 3(z - 1/3) = 0.$$

3 "Higher-Order Derivatives and Extrema" (MT 3)

In single variable calculus, we learn an application of differentiation–the identification of relative extrema rather quickly. In particular, we utilize the second derivative test; if $f'(x) = 0$, then there is a possibility for extrema, which we can (usually) differentiate by investigating the second derivative. Yet another method is the candidates test, which uses the extreme value theorem to enable us to find extrema on closed intervals. We will see extensions of all of these.

3.1 "Iterated Partial Derivatives" (MT 3.1)

Recall that taking the partial derivative of a function gives another function. We can take the partial derivative again, and get another function, which we call the *second order partial derivative* of the function. Note that we do not have to take the partial derivative of a function with respect to same variable twice, only that we take two partial derivatives. So for a function $f(x, y)$, the following are all meaningful second order partial derivatives:

$$\frac{\partial^2 f}{\partial x^2} \quad \frac{\partial^2 f}{\partial y^2} \quad \frac{\partial^2 f}{\partial x \partial y} \quad \frac{\partial^2 f}{\partial y \partial x}$$

Of course, you could also take the partial derivative of f with respect to say, u, or something, but that wouldn't be meaningful. We call the second-order derivatives that arise from differentiating to the same variable twice *iterated partial derivatives* and otherwise *mixed partial derivatives.*

We describe functions with partial derivatives that are continuous as of class C^1. We similarly define C^2 with second order partial derivatives.

This process can also be repeated for third order derivatives.

Theorem 3.1: Equality of Mixed Partial Derivatives

Let $f(x, y)$ be of class C^2. Then

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}.$$

Example 3.1: MT Secton 3.1 Problem 4

Compute all second partial derivatives of

$$f(x, y) = e^{-xy^2} + y^3x^4.$$

We first compute the first oder partial derivatives:

$$\begin{aligned}\frac{\partial f}{\partial x} &= \frac{\partial}{\partial x} f \\ &= \frac{\partial}{\partial x} e^{-xy^2} + y^3x^4\end{aligned}$$

$$\begin{aligned}
&= -y^2 e^{-xy^2} + 4xy^3x^3 \\
\frac{\partial f}{\partial y} &= \frac{\partial}{\partial y} f \\
&= \frac{\partial}{\partial y} e^{-xy^2} + y^3x^4 \\
&= -2xye^{-xy^2} + 3y^2x^4
\end{aligned}$$

Now we can compute the second order derivatives, both ways:

$$\begin{aligned}
\frac{\partial^2 f}{\partial x^2} &= \frac{\partial}{\partial x}\frac{\partial f}{\partial x} \\
&= \frac{\partial}{\partial x}\left[-y^2 e^{-xy^2} + 3x^4y^2\right] \\
&= -y^2 e^{-xy^2} \cdot \frac{\partial}{\partial x}(-xy^2) + 3x^4y^2 \\
&= -y^2 e^{-xy^2} \cdot -y^2 + 3x^4y^2 \\
&= y^4 e^{-xy^2} + 3x^4y^2 \\
\frac{\partial^2 f}{\partial y^2} &= \frac{\partial}{\partial y}\frac{\partial f}{\partial y} \\
&= \frac{\partial}{\partial y}\left[-2xye^{-xy^2} + 3x^4y^2\right] \\
&= -2xe^{-xy^2} + 4x^2y^2e^{-xy^2} + 6x^4y \\
\frac{\partial^2 f}{\partial x \partial y} &= \frac{\partial}{\partial x}\frac{\partial f}{\partial y} \\
&= \frac{\partial}{\partial x}\left[-2xye^{-xy^2} + 3y^2x^4\right] \\
&= -2ye^{-xy^2} + 4xy^3e^{-xy^2} + 12x^3y^2 \\
\frac{\partial^2 f}{\partial y \partial x} &= \frac{\partial}{\partial y}\frac{\partial f}{\partial x} \\
&= \frac{\partial}{\partial y}\left[-y^2 e^{-xy^2} + 4x^3y^3\right] \\
&= -2ye^{-xy^2} + 2xy^3e^{-xy^3} + 12x^3y^2.
\end{aligned}$$

3.1.1 Verifying PDEs

There's really nothing much to say here (and the textbook doesn't say anything either!), so here's just a list of problems.

Example 3.2: MT Section 3.1 Problem 11

Show that the following functions satisfy the one-dimensional wave equation

$$\frac{\partial^2 f}{\partial x^2} = \frac{1}{c^2}\frac{\partial^2 f}{\partial t^2}$$

(a) $f(x,t) = \sin(x - ct)$
(b) $f(x,t) = \sin(x)\sin(ct)$
(c) $f(x,t) = (x - ct)^6 + (x + ct)^6$.

We compute the second partial derivatives, then consider the wave equation constraint:

(a)

$$\begin{aligned}\frac{\partial^2 f}{\partial x^2} &= \frac{\partial}{\partial x}\frac{\partial}{\partial x}\sin(x - ct)\\ &= \frac{\partial}{\partial x}\cos(x - ct)\\ &= -\sin(x - ct)\\ \frac{\partial^2 f}{\partial t^2} &= \sin(x - ct)\\ &= -c\cos(x - ct)\\ &= c^2\sin(x - ct).\end{aligned}$$

$$\begin{aligned}\frac{\partial^2 f}{\partial x^2} &\overset{?}{=} \frac{1}{c^2}\frac{\partial^2 f}{\partial t^2}\\ -\sin(x - ct) &\overset{?}{=} \frac{1}{c^2}(-c^2\sin(x - ct))\\ -\sin(x - ct) &\overset{\checkmark}{=} -\sin(x - ct)\end{aligned}$$

(b)

$$\begin{aligned}\frac{\partial^2 f}{\partial x^2} &= \frac{\partial}{\partial x}\frac{\partial}{\partial x}\sin(x)\sin(ct)\\ &= \frac{\partial}{\partial x}\cos(x)\sin(ct)\\ &= -\sin(x)\sin(ct)\\ \frac{\partial^2 f}{\partial t^2} &= \frac{\partial}{\partial t}\frac{\partial}{\partial t}\sin(x)\sin(ct)\\ &= \frac{\partial}{\partial t}c\sin(x)\cos(ct)\\ &= -c^2\sin(x)\sin(ct)\end{aligned}$$

$$\frac{\partial^2 f}{\partial x^2} \overset{?}{=} \frac{1}{c^2}\frac{\partial^2 f}{\partial t^2}$$

$$-\sin(x)\sin(ct) \stackrel{?}{=} \frac{1}{c^2}(-c^2\sin(x)\sin(ct))$$
$$-\sin(x)\sin(ct) \stackrel{\checkmark}{=} -\sin(x)\sin(ct).$$

(c)

$$\begin{aligned}
\frac{\partial^2 f}{\partial x^2} &= \frac{\partial}{\partial x}\frac{\partial}{\partial x}(x-ct)^6 + (x+ct)^6 \\
&= \frac{\partial}{\partial x}6(x-ct)^5 + 6(x+ct)^5 \\
&= 30(x-ct)^4 + 30(x+ct)^4 \\
\frac{\partial^2 f}{\partial t^2} &= \frac{\partial}{\partial t}\frac{\partial}{\partial t}(x-ct)^6 + (x+ct)^6 \\
&= \frac{\partial}{\partial t} - 6c(x-ct)^5 + 6c(x+ct)^5 \\
&= 30c^2(x-ct)^4 + 30c^2(x+ct)^4
\end{aligned}$$

$$\begin{aligned}
\frac{\partial^2 f}{\partial x^2} &\stackrel{?}{=} \frac{1}{c^2}\frac{\partial^2 f}{\partial t^2} \\
30(x-ct)^4 + 30(x+ct)^4 &\stackrel{?}{=} \frac{1}{c^2}(30c^2(x-ct)^4 + 30c^2(x+ct)^4) \\
30(x-ct)^4 + 30(x+ct)^4 &\stackrel{\checkmark}{=} 30(x-ct)^4 + 30(x+ct)^4.
\end{aligned}$$

Example 3.3: MT Section 3.1 Problem 11

(a) Show that $T(x,t) = e^{-kt}\cos(x)$ satisfies the one-dimensional heat equation

$$k\frac{\partial^2 T}{\partial x^2} = \frac{\partial T}{\partial t}.$$

(b) Show that $T(x,y,t) = e^{-kt}(\cos(x)+\cos(y))$ satisfies the two-dimensional heat equation

$$k\left(\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2}\right) = \frac{\partial T}{\partial t}.$$

(c) Show that $T(x,y,z,t) = e^{-kt}(\cos(x)+\cos(y)+\cos(z))$ satisfies the three-dimensional equation

$$k\left(\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} + \frac{\partial^2 T}{\partial z^2}\right) = \frac{\partial T}{\partial t}.$$

(a)

$$\frac{\partial^2 T}{\partial x^2} = \frac{\partial}{\partial x}\frac{\partial}{\partial x}e^{-kt}\cos(x)$$

$$\begin{aligned} &= \frac{\partial}{\partial x} - e^{-kt}\sin(x) \\ &= -e^{-kt}\cos(x) \\ \frac{\partial T}{\partial t} &= \frac{\partial}{\partial t} e^{-kt}\cos(x) \\ &= -ke^{-kt}\cos(x) \end{aligned}$$

$$\begin{aligned} k\frac{\partial^2 T}{\partial x^2} &\stackrel{?}{=} \frac{\partial T}{\partial t} \\ k(-e^{-kt}\cos(x)) &\stackrel{?}{=} -ke^{-kt}\cos(x) \\ -ke^{-kt}\cos(x) &\stackrel{\checkmark}{=} -ke^{-kt}\cos(x). \end{aligned}$$

(b) Since T is symmetric across $x = y$, we can find the second order partial derivative of T with respect to x, and obtain the second order partial derivative of T with respect to y by swapping.

$$\begin{aligned} \frac{\partial T}{\partial x^2} &= \frac{\partial}{\partial x}\frac{\partial}{\partial x} e^{-kt}(\cos(x) + \cos(y)) \\ &= \frac{\partial}{\partial x} - e^{-kt}\sin(x) \\ &= -e^{-kt}\cos(x) \\ \therefore \frac{\partial}{\partial y} &= -e^{-kt}\cos(y) \\ \frac{\partial T}{\partial t} &= \frac{\partial}{\partial t} e^{-kt}(\cos(x) + \cos(y)) \\ &= -ke^{-kt}(\cos(x) + \cos(y)) \end{aligned}$$

$$\begin{aligned} k\left(\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2}\right) &\stackrel{?}{=} \frac{\partial T}{\partial t} \\ k(-e^{-kt}\cos(x) - e^{-kt}\cos(y)) &\stackrel{?}{=} -ke^{-kt}(\cos(x) + \cos(y)) \\ -ke^{-kt}(\cos(x) + \cos(y)) &\stackrel{\checkmark}{=} -ke^{-kt}(\cos(x) + \cos(y)). \end{aligned}$$

(c) Since distributing gives $e^{-kt}\cos(x) + C$ where C is a constant with respect to x, as are the previous two functions, and since T is symmetric across $x = y = z$ the second order partial derivatives are the same, and we immediately conclude that

$$\begin{aligned} \frac{\partial^2 T}{\partial x^2} &= -e^{-kt}\cos(x) \\ \frac{\partial^2 T}{\partial y^2} &= -e^{-kt}\cos(y) \\ \frac{\partial^2 T}{\partial z^2} &= -e^{-kt}\cos(z). \end{aligned}$$

$$\begin{aligned}\frac{\partial T}{\partial t} &= \frac{\partial}{\partial t} e^{-kt}(\cos(x) + \cos(y) + \cos(z)) \\ &= -ke^{-kt}(\cos(x) + \cos(y) + \cos(z))\end{aligned}$$

$$\begin{aligned}k\left(\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} + \frac{\partial^2 T}{\partial z^2}\right) &\stackrel{?}{=} \frac{\partial T}{\partial t} \\ k(-e^{-kt}\cos(x) - e^{-kt}\cos(y) - e^{-kt}\cos(z)) &\stackrel{?}{=} -ke^{-kt}(\cos(x) + \cos(y) + \cos(z)) \\ -ke^{-kt}(\cos(x) + \cos(y) + \cos(z)) &\stackrel{\checkmark}{=} -ke^{-kt}(\cos(x) + \cos(y) + \cos(z)).\end{aligned}$$

3.1.2 Partial derivatives on compositions

The chain rule rears its ugly head for higher order partial derivatives as well. In general, keep in mind that mixed partials are equal, so we can "move" the function around the various differentials.

Example 3.4: MT Section 3.1 Problem 22

Let $w = f(x, y)$ be a function of two variables and let $x = u + v$ and $y = u - v$. Show that

$$\frac{\partial^2 w}{\partial u \partial v} = \frac{\partial^2 w}{\partial x^2} - \frac{\partial^2 w}{\partial y^2}$$

This problem is actually rather difficult on first glance and on second glance, and can often be time consuming to replicate. On my third time doing this problem, I can now do it rather quickly.

The key idea is that you can factor out partial derivatives and that mixed partial derivatives are equal, and to repeatedly use the chain rule. Of any problem, I imagine this to be the one where each individual step is relevant, rather than other problems, where there are only a few key ideas.

Note that since $x = u + v$, that $\frac{\partial x}{\partial u} = \frac{\partial x}{\partial v} = 1$, and similarly, since $y = u - v$, $\frac{\partial y}{\partial u} = 1$ and $\frac{\partial y}{\partial v} = -1$.

$$\begin{aligned}\frac{\partial^2 w}{\partial u \partial v} &= \frac{\partial}{\partial u}\frac{\partial}{\partial v} f(x(u,v), y(u,v)) \\ &= \frac{\partial}{\partial u}\left[\frac{\partial f}{\partial x}\frac{\partial x}{\partial v} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial v}\right] && \text{(Chain rule)} \\ &= \frac{\partial}{\partial u}\left[\frac{\partial f}{\partial x} - \frac{\partial f}{\partial y}\right] && (x_v = 1 \text{ and } x_v = -1) \\ &= \frac{\partial}{\partial u}\frac{\partial f}{\partial x} - \frac{\partial}{\partial u}\frac{\partial f}{\partial y} && \text{(Linearity of partial differentiation)} \\ &= \frac{\partial^2 f}{\partial u \partial x} - \frac{\partial^2 f}{\partial u \partial y}\end{aligned}$$

$$
\begin{aligned}
&= \frac{\partial^2 f}{\partial x \partial u} - \frac{\partial^2 f}{\partial y \partial u} && \text{(Equality of mixed partials)} \\
&= \frac{\partial}{\partial x}\frac{\partial}{\partial u} f(x(u,v), y(u,v)) - \frac{\partial}{\partial y}\frac{\partial}{\partial u} f(x(u,v), y(u,v)) && \text{(Expanding)} \\
&= \frac{\partial}{\partial x}\left[\frac{\partial f}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial u}\right] - \frac{\partial}{\partial y}\left[\frac{\partial f}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial u}\right] && \text{(Chain rule)} \\
&= \frac{\partial}{\partial x}\left[\frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}\right] - \frac{\partial}{\partial y}\left[\frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}\right] && (x_u = 1 \text{ and } y_u = 1) \\
&= \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial x \partial y} - \frac{\partial^2 f}{\partial y \partial x} - \frac{\partial^2 f}{\partial y^2} && \text{(Linearity of partial derivatives)} \\
&= \frac{\partial^2 f}{\partial x^2} - \frac{\partial^2 f}{\partial y^2} && \text{(Equality of mixed partials)}
\end{aligned}
$$

I show every single step in a separate line to make this understandable, but if you reproduce this or encounter a similar problem, you can realistically do multiple steps at once.

3.2 "Taylor's Theorem" (MT 3.2)

Perhaps differently from single variable calculus, which defines Taylor's theorem for an analytic function as

$$f(x) \approx \sum_{i=0}^{n} \frac{f^{(i)}(h)}{i!}(x-h)^i,$$

we express this differently, using a difference from h explicitly instead.

Let x_0 be the center of the Taylor expansion and h be a vector. We instead are interested in expressing

$$f(x_0 + h)$$

instead of just $f(x)$.

Theorem 3.2: Taylor's Theorem

Let $f : U \subset \mathbb{R}^n \mapsto \mathbb{R}$ have continuous partial derivatives up to order n. Denote $\binom{S}{n}$ as the set of all n-length permutations of S.

$$f(x_0 + h) = f(x_0) + \sum_{i=1}^{m} \frac{1}{i!} \sum_{\forall (a_i) \in \binom{N}{m}} \prod (h_{a_i}) \frac{\partial^m f}{\prod \partial (x_{a_i})}(x_0)$$

That is, from $i = 1$ to m, we take m of the variables, take the partial derivatives of them, and multiply all of these by the m components of h.

It may be better to consider the first and second order ones instead, and try to keep an abstract generalization process in your head:

$$f(x_0 + h) = f(x_0) + \sum_{i=1}^{n} h_i \frac{\partial f}{\partial x_i}(x_0) + R_1(x_0, h)$$

or the second order

$$f(x_0 + h) = f(x_0) + \sum_{i=1}^{n} h_i \frac{\partial f}{\partial x_i}(x_0) + \frac{1}{2}\sum_{i,j=1}^{n} h_i h_j \frac{\partial^2 f}{\partial x_i x_j}(x_0) + R_2(x_0, h).$$

Theorem 3.3: Taylor's Theorem Remainder

$$\begin{aligned} R_1(x_0, h) &= \sum_{i,j=1}^{n} \int_0^1 (1-t)\frac{\partial^2 f}{\partial x_i \partial x_j} f(x_0 + th) h_i h_j \, dt \\ &= \sum_{i,j=1}^{n} \frac{1}{2}\frac{\partial^2 f}{\partial x_i \partial x_j}(c_{i,j}) h_i h_j \\ R_2(x_0, h) &= \sum_{i,j,k=1}^{n} \int_0^1 \frac{(t-1)^2}{2} \frac{\partial^3 f}{\partial x_i \partial x_j \partial x_k}(x_0 + th) h_i h_j h_k \, dt \\ &= \sum_{i,j,k=1}^{n} \frac{1}{3!}\frac{\partial^3 f}{\partial x_i \partial x_j \partial x_k}(c_{i,j,k}) h_i h_j h_k \end{aligned}$$

where $c_{(\cdot)}$ is a point on the line from x_0 to $x_0 + h$.

Example 3.5: MT Section 3.2 Problem 2

Suppose $L : \mathbb{R}^2 \mapsto \mathbb{R}$ is linear, so that L has the form

$$L(x, y) = ax + by.$$

(a) Find the first-order Taylor approximation for L.
(b) Find the second-order Taylor approximation for L.
(c) What will higher-order approximations look like?

(a) We need all first order partial derivatives. $f_x = a$ and $f_y = b$. To parameterize the remainder term, we need the second order partial derivatives, all of which are 0 which entails that the remainder term is 0. This gives the approximation

$$L(X_0 + h) = ax_0 + by_0 + ah_1 + bh_2.$$

(b) As the second order partial derivatives are 0, the additional terms are also 0. Further, this entails the third order partial derivatives are 0, so the remainder term is also 0, which gives the approximation

$$L(X_0 + h) = ax_0 + by_0 + ah_1 + bh_2.$$

(c) They'll look the same. Since the second order partial derivatives are 0, so will higher order partial derivatives, so no additional higher order terms can be used, and the remainder is similarly zero.

Example 3.6: MT Section 3.2 Problem 3

Determine the second-order Taylor formula for

$$f(x, y) = (x + y)^2$$

where $(x_0, y_0) = (0, 0)$.

Let's first compute all partial derivatives, up to the third order:

$$\begin{aligned} f_x &= 2(x + y) \\ f_y &= 2(x + y) \\ f_{xx} &= 2 \\ f_{yy} &= 2 \\ f_{xy} &= 2 \\ f_{xxx} = f_{xxy} = f_{xyy} = f_{yyy} &= 0. \end{aligned}$$

Evaluating,

$$\begin{aligned} f(0, 0) &= 0 \\ f_x(0, 0) = f_y(0, 0) &= 0 \\ f_{xx}(0, 0) = f_{xy}(0, 0) = f_{yy}(0, 0) &= 2. \end{aligned}$$

Since the third order partial derivatives are 0, so is the remainder. This gives the approximation

$$\begin{aligned} f(x_0 + h_1, y_0 + h_2) &= 0 + 0h_1 + 0h_2 + \frac{2h_1^2 + 4h_1h_2 + 2h_2^2}{2} \\ &= h_1^2 + 2h_1h_2 + h_2^2. \end{aligned}$$

Example 3.7: MT Section 3.2 Problem 8

Determine the second-order Taylor formula for

$$f(x, y) = e^{(x-1)^2} \cos(y)$$

where $(x_0, y_0) = (1, 0)$.

Computing the partial derivatives up to second order,

$$\begin{aligned} f_x &= 2(x - 1)e^{(x-1)^2} \cos(y) \\ f_y &= -e^{(x-1)^2} \sin(y) \\ f_{xx} &= 2e^{(x-1)^2} \cos(y) + 4(x - 1)^2 e^{(x-1)^2} \cos(y) \\ f_{xy} &= -2(x - 1)e^{(x-1)^2} \sin(y) \\ f_{yy} &= -e^{(x-1)^2} \cos(y). \end{aligned}$$

Evaluating,

$$\begin{aligned} f(1,0) &= 1 \\ f_x(1,0) &= 0 \\ f_y(1,0) &= 0 \\ f_{xx}(1,0) &= 2 \\ f_{xy}(1,0) &= 0 \\ f_{yy}(1,0) &= -1. \end{aligned}$$

This gives the approximation

$$\begin{aligned} f(x_0+h_1, y_0+h2) &= 1 + 0h_1 + 0h_2 + \frac{2h_1^2 + 0h_1h_2 - 1h_2^2}{2} + R_2((x_0,y_0),(h_1,h_2)) \\ &= 1 + h_1^2 - \frac{h_2^2}{2} + R_2((x_0,y_0),(h_1,h_2)). \end{aligned}$$

3.3 "Extrema of Real-Valued Functions" (MT 3.3)

We borrow the familiar ideas from single-variable calculus, but bringing the ideas into neighborhoods instead of intervals. Since intervals are the 2-dimensional case of neighborhoods, all we're really doing is just restating things that continue to work in higher dimensions.

Definition 3.1: Critical Point

If $f : U \subset \mathbb{R}^n \mapsto \mathbb{R}$ is a scalar function, a point x_0 is a critical point of f if f is either non-differentiable at x_0 or if $\mathbf{D}f(x_0) = 0$.

Definition 3.2: Relative minimum

If $f : U \subset \mathbb{R}^n \mapsto \mathbb{R}$ is a scalar function, a point $x_0 \in U$ is called a local minimum of f if there is some neighborhood V of x_0 such that for all points x in V, $f(x) \geq f(x_0)$.

We can flip the inequality for a local maximum.

If a critical point is not a local minimum or maximum, it is called a *saddle point*.

Theorem 3.4: First Derivative Test

If $U \subset \mathbb{R}^n$ is open, the function $f : U \subset \mathbb{R}^n \mapsto \mathbb{R}$ is differentiable and $x_0 \in U$ is a local extremum, then x_0 is a critical point of f, so

$$\mathbf{D}f(x_0) = 0.$$

This gives us a criterion to identify candidates for extrema–by identifying all points x_0 where $\mathbf{D}f(x_0) = 0$, these are the possibilities for extrema, and we only have to focus on these points.

As we do in single variable calculus, we have a second derivative test as well, but we must first define the Hessian matrix:

Definition 3.3: Hessian matrix

Suppose that $f : U \subset \mathbb{R}^n \mapsto \mathbb{R}$ has second-order continuous derivatives for all x_i, x_j at a point $x_0 \in U$, the Hessian of f at x_0, denoted $Hf(x_0)(h)$ is the quadratic function

$$\begin{aligned} Hf(x_0)(h) &= \frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(x_0) h_i h_j \\ &= \frac{1}{2} \begin{bmatrix} h_1, \dots, h_n \end{bmatrix} \begin{bmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_n} \end{bmatrix} \begin{bmatrix} h_1 \\ \vdots \\ h_n \end{bmatrix} \end{aligned}$$

At a critical point, *the Hessian is the first non-constant term in the Taylor series of f.*

We call a quadratic function $g : \mathbb{R}^n \mapsto \mathbb{R}$ as positive definite if $g(h) \geq 0$ for all h and negative definite if we flip the inequality. We allow $g(h) = 0$ when $h = 0$ only.

Theorem 3.5: Second-Derivative Test

If $f : U \subset \mathbb{R}^n \mapsto \mathbb{R}$ is of class C^3, $x_0 \in U$ is a critical point of f, then

- If $Hf(x_0)$ is positive definite, then x_0 is a relative minimum.
- If $Hf(x_0)$ is negative definite, then x_0 is a relative maximum.

For the case of a two variable function, we have a special case that the expression

$$f_{xx}(x_0, y_0) f_{yy}(x_0, y_0) - f_{xy}^2(x_0, y_0),$$

known as the discriminant of the Hessian, is positive, then (x_0, y_0) is an extremum.

Example 3.8: MT Section 3.3 Problem 4

Find the critical points of

$$f(x, y) = x^2 + y^2 + 3xy$$

and then determine whether they are local maxima, local minima or saddle points.

Let's first find the first and second order partial derivatives:

$$\begin{aligned} f_x &= 2x + 3y \\ f_y &= 2y + 3x \\ f_{xx} &= 2 \\ f_{xy} &= 3 \\ f_{yy} &= 2 \end{aligned}$$

The first derivative test says all critical points are (x, y) such that $(f_x, f_y) = (0, 0)$, which is when $(x, y) = (0, 0)$.

The discriminant of the Hessian is then $2 \cdot 2 - 3^2 = 4 - 9 = -5$ which is negative, so $(0, 0)$ is a saddle point.

Example 3.9: MT Section 3.3 Problem 7

Find the critical points of

$$f(x,y) = 3x^2 + 2xy + 2x + y^2 + y + 4$$

and then determine whether they are local maxima, local minima or saddle points.

The first and second order partial derivatives are

$$\begin{aligned} f_x &= 6x + 2y + 2 \\ f_y &= 2x + 2y + 1 \\ f_{xx} &= 6 \\ f_{xy} &= 2 \\ f_{yy} &= 2 \end{aligned}$$

$(f_x, f_y) = (0,0)$ gives $(x,y) = (-1/4, -1/4)$.

The discriminant of the Hessian, $6 \cdot 2 - 2^2 = 12 - 4 = 8$, is positive, so this is an extremum.

Since f_{xx} is positive, the point is a local minimum.

Example 3.10: MT Section 3.3 Problem 13

Find the critical points of

$$f(x,y) = xy + \frac{1}{x} + \frac{1}{y}$$

and then determine whether they are local maxima, local minima or saddle points.

The first and second order partial derivatives are

$$\begin{aligned} f_x &= y - \frac{1}{x^2} \\ f_y &= x - \frac{1}{y^2} \\ f_{xx} &= \frac{2}{x^3} \\ f_{xy} &= 1 \\ f_{yy} &= \frac{2}{y^3} \end{aligned}$$

$(f_x, f_y) = (0,0)$ gives $(x,y) = (1,1)$. Evaluating the second order partial derivatives yields

$$\begin{aligned} f_{xx} &= 2 \\ f_{xy} &= 1 \\ f_{yy} &= 2. \end{aligned}$$

The discriminant of the Hessian, $2 \cdot 2 - 1^2 = 4 - 1 = 3$ is positive, so $(1,1)$ is an extrema. Since f_{xx} is positive, the point is a local minimum.

Example 3.11: MT Section 3.3 Problem 17

Find all local extrema for

$$f(x,y) = 8y^3 + 12x^2 - 24xy.$$

The first and second order partial derivatives are

$$\begin{aligned} f_x &= 24x - 24y \\ f_y &= 24y^2 - 24x \\ f_{xx} &= 24 \\ f_{xy} &= -24 \\ f_{yy} &= 48y. \end{aligned}$$

$(f_x, f_y) = (0,0)$ implies $(x,y) = (0,0)$ or $(x,y) = (1,1)$.

Evaluating, $f_{yy}(0,0) = 0$ and $f_{yy}(1,1) = 48$. Then the discriminant of the Hessian at $(0,0)$ is $24 \cdot 0 - (-24)^2 < 0$ and at $(1,1)$ is $24 \cdot 48 - (-24)^2 > 0$, so $(0,0)$ is a saddle point and $(1,1)$ is an extrema.

Since $f_{xx}(1,1) = 24$, $(1,1)$ is a local minimum.

Example 3.12: MT Section 3.3 Problem 23

An examination of the function

$$f : \mathbb{R}^2 \mapsto \mathbb{R}, (x,y) \mapsto (y - 3x^2)(y - x^2)$$

will give an idea of the difficulty of finding conditions that guarantee that a critical point is a relative extremum when the second derivative test fails. Show that

(a) the origin is a critical point of f,

(b) f has a relative minimum at $(0,0)$ on every straight line through $(0,0)$; that is, if $g(t) = (at, bt)$, then $f \circ g : \mathbb{R} \mapsto \mathbb{R}$ has a relative minimum at 0, for every choice of a and b,

(c) the origin is not a relative minimum of f.

Analysis of the function is actually easier when expanded. Expanding,

$$\begin{aligned} f(x,y) &= (y - 3x^2)(y - x^2) \\ &= y^2 - 4x^2 y + x^4. \end{aligned}$$

(a) The first order partial derivatives are $f_x = -8xy + 4x^3$ and $f_y = 2y - 4x^2$. $f_x(0,0) = f_y(0,0) = 0$, so the origin is a critical point.

(b) Though the chain rule is an option, it's simpler to directly express the composition:

$$\begin{aligned} f(g(t)) &= (bt)^2 - 4(at)^2(bt) + (at)^4 \\ &= b^2t^2 - 4a^2t^2bt + a^4t^4 \\ &= b^2t^2 - 4a^2bt^3 + a^4t^4. \end{aligned}$$

Then the first and second order partial derivatives are

$$\begin{aligned} f_t &= 2b^2t - 12a^2bt^2 + 4a^4t^3 \\ f_{tt} &= 2b^2 - 24a^2bt + 12a^2t^2. \end{aligned}$$

The origin, $(0,0)$ is achieved when $t = 0$. Evaluating, we have that $f_t = 0$ and $f_{tt}(0) = 2b^2$. We retroactively cover up that we allow for any value a, b, and now require that $b \neq 0$. In this case, $2b^2$ is positive, and by the single variable second derivative test, the origin is a local minimum.

(c) Perhaps it is best to take the ideas from (b) and consider the origin by some path through $(0,0)$. We have already found that all straight lines (except degenerate ones, which give no leads) fail, so the natural next step is to consider parabolas. Alternatively, note that x is always quadratic, so we reason that trying to make y quadratic might help. Let $y = cx^2$ for some x. This gives

$$\begin{aligned} f(x, cx^2) &= (cx^2 - 3x^2)(cx^2 - x^2) \\ &= (x^2(c-3))(x^2(c-1)) \\ &= x^2(c-1)(c-3). \end{aligned}$$

Observe that if $1 < c < 3$, then $f(x, cx^2)$ is always negative, and the origin is actually a local maximum, which contradicts that the origin is a local minimum.

Example 3.13: MT Section 3.2 Problem 41

Find the absolute maximum and minimum values for

$$f(x, y) = \sin(x) + \cos(y)$$

on the rectangle R defined by $0 \leq x \leq 2\pi$, $0 \leq y \leq 2\pi$.

On first glance, this might *seem* like a constrained optimization problem, which we will discuss soon, it's not. This just means rather than having an infinite number of solutions, we only pick those in the rectangle.

$$\begin{aligned} f_x &= \cos(x) \\ f_y &= -\sin(y) \\ f_{xx} &= -\sin(x) \\ f_{xy} &= 0 \\ f_{yy} &= -\cos(x) \end{aligned}$$

$(f_x, f_y) = (0,0)$ implies that $x = \pi/2$ or $3\pi/2$ and $y = 0$ or π. I'll aggregate the evaluations into a table:

		x	
		$\pi/2$	$3\pi/2$
y	0	2	0
	π	0	-2

Since sin and cos are both bounded by -1 and 1, the values of 2 and -2 are the maximum and minimum. Also I really don't want to do the Hessian. Trust me that the 0 are then saddle points.

3.3.1 A light treatment on constrained optimization

We are often interested in finding extrema with respect to some constraint. Though we will discuss a preferred method in the next section, we can still solve these problems if we can "impose" the constraint into the objective function we are interested in. For instance, if we are interesting in extremizing the function $f(x, y, z)$, and we are able to express the constraint in a way such as $x = g(y, z)$, we can then consider $f(g(y, z), y, z)$ and use our current methods. The next problems use this idea.

Example 3.14: MT Section 3.3 Problem 28

Find the point on the plane $2x - y + 2z = 20$ nearest the origin.

For a given point (x, y, z), we have that their distance to the origin is given as

$$d(x, y, z) = \sqrt{x^2 + y^2 + z^2}.$$

However, this is often messy to use, so since we wish to minimize the distance, it suffices to minimize the square distance,

$$d^2(x, y, z) = x^2 + y^2 + z^2$$

instead.

We still can't minimize the function just yet, as we must incorporate that the plane is on the constraint somewhere. Since the plane is a 2-dimensional surface, we actually only have 2 degrees of freedom, and we can express one variable entirely as a function of the other two (this is actually not entire what being 2-dimensional means, but in this case, it is true). Note that $y = 2x + 2z - 20$. Then, we can consider

$$d^2(x, 2x + 2z - 20, z) = x^2 + (2x + 2z - 20)^2 + z^2$$

instead.

There is no need to expand, as the chain rule is quite sufficient to keep the computations reasonable here:

$$\begin{aligned} f_x &= 2x + 2(2x + 2z - 20) \cdot 2 \\ &= 2x + 4(2x + 2z - 20) \\ &= 2x + 8x + 8z - 80 \end{aligned}$$

$$\begin{aligned} &= 10x + 8z - 80 \\ f_z &= 10z + 8x - 80. \end{aligned}$$

$(f_x, f_z) = (0, 0)$ implies that $(x, z) = (40/9, 40/9)$. Then $y = 20/9$.

Since the problem statement explicates the existence of a minimum, this must be it, but for posterity, the second order partial derivatives are

$$\begin{aligned} f_{xx} &= 10 \\ f_{xz} &= 8 \\ f_{zz} &= 10 \end{aligned}$$

and the discriminant of the Hessian, $10 \cdot 10 - 8^2 = 100 - 64 = 36 > 0$, so the point is an extremum. Since $f_{xx} > 0$, the point is a minimum.

Example 3.15: MT Section 3.3 Problem 31

Write the number 120 as a sum of three numbers so that the sum of the products taken two at a time is a maximum.

The description of the desired sum of products is rather unclear. It's

$$xy + yz + zx.$$

The first part, then, is the restriction that

$$x + y + z = 120.$$

Let's eliminate z. We have that $z = 120 - x - y$, and so we are interesting in maximizing

$$f(x, y) = xy + y(120 - x - y) + (120 - x - y)x.$$

Expanding here may be useful, but the product rule actually keeps things reasonable, if you're willing to see:

$$\begin{aligned} f_x &= y - y - x + 120 - x - y \\ &= 120 - 2x - y \\ f_y &= 120 - 2y - x. \end{aligned}$$

$(f_x, f_y) = (0, 0)$ has a solution at $(x, y) = (40, 40)$.

The second order derivatives are

$$\begin{aligned} f_{xx} &= -2 \\ f_{yy} &= -2 \\ f_{xy} &= -1. \end{aligned}$$

The discriminant of the Hessian is then $-2 \cdot -2 - (-1)^2 = 4 - 1 = 3$, which is positive, so the point is an extremum. Since $f_{xx} < 0$, the point is a relative maximum.

3.3.2 The Hessian for multiple variables

For a function of more than 2 variables, the Hessian is positive definite if all n square submatrices have positive determinant. The Hessian is negative definite if the n square matrices alternate signs.

3.3.3 Developing Global Extrema

We develop the theory for identifying global extrema, and discuss a technique for such in a later section.

Definition 3.4: Global extrema

Suppose $f : A \subset \mathbb{R}$ is a function defined on a set A. A point $x_0 \in A$ is said to be an absolute maximum point of f if

$$f(x) \leq f(x_0)$$

for all $x \in A$.
Flipping the inequality defines a global maximum.

Definition 3.5: Bounded sets

A set $D \in \mathbb{R}^n$ is said to be bounded if there exists a number $M > 0$ such that

$$\|x\| < M$$

for all $x \in D$. A set is closed if it contains all its boundary points.

Alternatively, a set is bounded if it can be contained in some ball.

Theorem 3.6: Global Existence Extrema Theorem

Let D be closed and bounded in $\mathbb{R}^n$ and let $f : D \mapsto \mathbb{R}$ be continuous. Then f assumes its absolute maximum and minimum values at some points x_0 and x_1 of D.

Hence a strategy to identify the global extrema is to identify the critical points of f on the interior of a region D, then identify the critical points of f on the boundary of the region ∂D by treating f as a function defined only on ∂D, and go from there.

3.4 "Constrained Extrema and Lagrange Multipliers" (MT 3.4)

This section describes a means to solve an optimization problem of maximizing f subject to some constraint $g = c$. This can be used for identifying global extrema, as we can look the describe a point laying on the boundary of a closed region by using the constraint to impose that a point is on the boundary.

Theorem 3.7: The Method of Lagrange Multipliers

Suppose that $f : U \subset \mathbb{R}^n \mapsto \mathbb{R}$ and $g : U \subset \mathbb{R}^n \mapsto \mathbb{R}$ are given C^1 real valued functions. Let $x_0 \in U$ and $g(x_0) = c$ and let S be the level set for g with value c and that $\nabla g(x_0) \neq 0$.
If $f \mid S$, denoting "f restricted to S", has a local maximum or minimum on S at x_0, then there is a real number λ such that

$$\nabla f(x_0) = \lambda \nabla g(x_0).$$

Theorem 3.8: The Gradient is Perpendicular to the Level Set

If f, when constrained to a surface S has a maximum or minimum at x_0, then $\nabla f(x_0)$ is perpendicular to S at x_0.

This is just a restatement of theorem 2.11.

3.4.1 Existence of an extrema

Identifying that a minimum/maximum exists is necessary and is its own idea. It generally suffices to show that the region is bounded. Do note that if we have multiple candidates of global extrema, some may be saddle points, but the minima/maxima of the candidates are the minima/maxima, if we have shown extrema exist.

3.4.2 Multiple constraints

Suppose we have multiple g_i and c_i, where we have $g_i = c_i$. Then the method becomes

$$\nabla f(x_0) = \sum_{i=1}^{k} \lambda_i \nabla g_i(x_0).$$

3.4.3 The Second Derivative Test for Constrained Extrema

We must again describe a second derivative test to identify if a critical point is an extrema.

Theorem 3.9: The Bordered Hessian

Let $f : U \subset \mathbb{R}^2 \mapsto \mathbb{R}$ and $g : U \subset \mathbb{R}^2 \mapsto \mathbb{R}$ be smooth (at least C^2) functions. Let $v_0 \in U$, $g(v_0) = c$ and S be the level curve for g with value c. Assume that $\nabla g(v_0) \neq 0$ and there there is a real number λ such that $\nabla f(v_0) = \lambda \nabla g(v_0)$, define

$$h = f - \lambda g$$

and we define the *Bordered Hessian* determinant

$$|\bar{H}| = \begin{vmatrix} 0 & -\frac{\partial g}{\partial x} & -\frac{\partial g}{\partial y} \\ -\frac{\partial g}{\partial x} & \frac{\partial^2 h}{\partial x^2} & \frac{\partial^2 h}{\partial x \partial y} \\ -\frac{\partial g}{\partial y} & \frac{\partial^2 h}{\partial x \partial y} & \frac{\partial^2 h}{\partial y^2} \end{vmatrix}$$

- If $|\bar{H}| > 0$ then v_0 is a local maximum for $f \mid S$
- If $|\bar{H}| < 0$ then v_0 is a local maximum for $f \mid S$
- If $|\bar{H}| = 0$, the test is inconclusive.

Example 3.16: MT Section 3.4 Problem 2

Consider all rectangles with a fixed perimeter p. Use Lagrange multipliers to show that the rectangle with maximal area is a square.

Let the rectangle have side values x and y. The perimeter constraint gives that

$$g : 2x + 2y = p.$$

The area is then

$$f(x, y) = xy.$$

First we have to show that a maximal area rectangle exists. Observe that we are only interested in areas that are non-negative, since lengths can't be negative, so we have that $x, y \geq 0$. Further, x and y, if we allow for a degenerate rectangle, achieve maximal value at $x, y = p/2$. Then we have a closed interval of possibilities as x and y range from 0 to $p/2$, also included the perimeter constraint. Finally, xy is bounded by $p^2/4$ on the overly permissive region $[0, p/2] \times [0, p/2]$, so we can proceed.

$$\nabla f = \lambda \nabla g$$
$$(y, x) = \lambda(2, 2)$$
$$(y, x) = (2\lambda, 2\lambda)$$

This gives that $y = 2\lambda = x$, which is a square.

Then, since we have that by the constraint $2x + 2y = p$, this yields $2x + 2x = p$ which implies $x = p/4$, and $y = p/4$. This is a square, since $x = y$.

Shouldn't the method of Lagrange multipliers also produce the global minimum as well, here? Well, after our discussion, we implicitly included that $x, y \geq 0$, which are not fully captured by the method of Lagrange multipliers. For completion, we should really compute the Hessian to check.

Fully solving gives $x = y = p/4$.

$$h = f - g = xy - 2x - 2y.$$

The partial derivatives are given as

$$\begin{aligned} h_x &= y - 2 \\ h_y &= x - 2 \\ h_{xx} &= 0 \\ h_{xy} &= 1 \\ h_{yy} &= 0 \end{aligned}$$

Evaluating gives $h_x(p/2, p/2) = h_y(p/2, p/2) = p/2 - 2$, and $g_x(p/2, p/2) = g_y(p/2, p/2) = p$ We then construct the bordered Hessian determinant,

$$\begin{aligned} |\bar{H}| &= \begin{vmatrix} 0 & -p & -p \\ -p & 0 & 1 \\ -p & 1 & 0 \end{vmatrix} \\ &= 0 - (-p)(p) + (-p)(-p) \\ &= 2p^2 \end{aligned}$$

which is positive, since $p > 0$.

Example 3.17: MT Section 3.4 Problem 3

Find the extrema of

$$f(x, y, z) = x - y + z$$

subject to

$$x^2 + y^2 + z^2 = 2.$$

Observe that the constraint describes the surface of a sphere of radius $\sqrt{2}$, a closed surface. Further, f is bounded on such a constraint, so a minimum/maximum must exist.

$$\begin{aligned} \nabla f &= \lambda \nabla g \\ (1, -1, 1) &= \lambda(2x, 2y, 2z) \\ (1, -1, 1) &= (2\lambda x, 2\lambda y, 2\lambda z) \end{aligned}$$

Since $2\lambda x = 1 = 2\lambda z$, we have that $x = z$. Similarly, since $2\lambda y = -1 = -2\lambda x$, $y = -x$. Inserting these into the the constraint yields

$$\begin{aligned} x^2 + (-x)^2 + x^2 &= 2 \\ x^2 + x^2 + x^2 &= 2 \end{aligned}$$

$$
\begin{aligned}
3x^2 &= 2 \\
x^2 &= \frac{2}{3} \\
x &= \pm\sqrt{\frac{2}{3}}.
\end{aligned}
$$

So our solutions are either $(\sqrt{2/3}, -\sqrt{2/3}, \sqrt{2/3})$ and $(-\sqrt{2/3}, \sqrt{2/3}, -\sqrt{2/3})$. Evaluating the first gives $\sqrt{2/3}$ and the second gives $-\sqrt{2/3}$, which are thus the maximum and minimum.

Example 3.18: MT Section 3.4 Problem 4

Find the extrema of

$$f(x, y) = x - y$$

subject to

$$x^2 - y^2 = 2.$$

This certainly doesn't look very closed, does it? We suspect that the method of Lagrange multipliers will fail here.

$$
\begin{aligned}
\nabla f &= \lambda \nabla g \\
(1, -1) &= \lambda(2x, -2y) \\
(1, -1) &= 2\lambda x, -2\lambda y)
\end{aligned}
$$

solving gives $2\lambda x = 1 = -(-1) = -(-2\lambda y) = 2\lambda y$, so $x = y$. However, when we consider the constraint, this gives

$$
\begin{aligned}
x^2 - y^2 &= x^2 - x^2 \\
&= 0,
\end{aligned}
$$

which does not satisfy the constraint.

From another perspective, we can express the constraint as $(x-y)(x+y) = 2$. So we can take any arbitrary value for $x - y$, and just identify an x, y that satisfies that and $x+y = 2/(x-y)$, so it's completely unbounded.

Example 3.19: MT Section 3.4 Problem 12

Use the method of Lagrange multipliers to find the absolute maximum and minimum values of

$$f(x, y) = x^2 + y^2 - x - y + 1$$

on the unit disk.

The unit disk can be expressed with the constraint

$$x^2 + y^2 \leq 1,$$

which isn't immediately usable with Lagrange multipliers, so we'll to decompose the constraint into two regions, one that is the interior,

$$x^2 + y^2 < 1$$

and one that is just the boundary,

$$x^2 + y^2 = 1,$$

where the boundary can considered by Lagrange multipliers.

We'll first consider the interior, which we do with a normal extrema approach:

$$\begin{aligned} f_x &= 2x - 1 \\ f_y &= 2y - 1 \\ f_{xx} &= 2 \\ f_{xy} &= 0 \\ f_{yy} &= 2 \end{aligned}$$

Then $(f_x, f_y) = (0, 0)$ gives $x, y = (1/2, 1/2)$. The discriminant of the Hessian here is $2 \cdot 2 - 0^2 = 4 > 0$, so this point is an extrema. As $f_{xx} > 0$, this is a local minimum. The actual value of which is 1/2.

Now we set our eyes on the boundary.

$$\begin{aligned} \nabla f &= \lambda \nabla g \\ (2x - 1, 2y - 1) &= \lambda(2x, 2y) \\ (2x - 1, 2y - 1) &= (2\lambda x, 2\lambda y) \end{aligned}$$

This is the first involved system. In these cases, usually solving for lambda gives a foothold.

$$\begin{aligned} 2x - 1 = 2\lambda x, \quad 2y - 1 = 2\lambda y &\implies \lambda = (2x - 1)/2x \quad \lambda = (2y - 1)/2y \\ &\implies \frac{2x - 1}{2x} = \frac{2y - 1}{2y} \\ &\implies (2x - 1)(2y) = (2y - 1)(2x) \\ &\implies 4xy - 2y = 4xy - 2x \\ &\implies x = y. \end{aligned}$$

The constraint then tells that

$$\begin{aligned} x^2 + x^2 &= 1 \\ 2x^2 &= 1 \\ x^2 &= \frac{1}{2} \\ x &= \pm\sqrt{\frac{1}{2}} \end{aligned}$$

so our possibilities are $(\sqrt{1/2}, \sqrt{1/2})$ and $(-\sqrt{1/2}, -\sqrt{1/2})$. Evaluating gives $2-\sqrt{2}$ for the first and $2+\sqrt{2}$ for the second. Since the unit disk's boundary is closed and the function is bounded, these are the minimum and maximum of f on the boundary.

Comparing, we have that $2+\sqrt{2}$ is the maximum and $1/2$ for the minimum.

Example 3.20: MT Section 3.4 Problem 13

Consider the function $f(x,y) = x^2 + xy + y^2$ defined on the unit disk, namely

$$D = \{(x,y) \mid x^2 + y^2 \leq 1\}.$$

Use the method of Lagrange multipliers to locate the maximum and minimum points on the unit circle. Use this to determine the absolute maximum and minimum values for f on D.

We firstly consider the interior of the unit disk.

$$\begin{aligned} f_x &= 2x + y \\ f_y &= 2y + x \\ f_{xx} &= 2 \\ f_{yy} &= 2 \\ f_{xy} &= 1. \end{aligned}$$

Then $(f_x, f_y) = (0,0)$ implies $(x,y) = (0,0)$. The discriminant of the Hessian, $2 \cdot 2 - 1^2 = 3 > 0$, so this is an extrema. Since $f_{xx} = 2 > 0$ this is a relative minimum. The value of which is 0.

We now inspect the boundary.

$$\begin{aligned} \nabla f &= \lambda \nabla g \\ (2x + y, 2y + x) &= \lambda(2x, 2y) \\ (2x + y, 2y + x) &= (2\lambda x, 2\lambda y) \end{aligned}$$

Again, solving for λ works here:

$$\begin{aligned} (2x + y, 2y + x) = (2\lambda x, 2\lambda y) &\implies (\lambda, \lambda) = ((2x + y)/2x, (2y + x)/2y) \\ &\implies (2x + y)(2y) = (2y + x)(2x) \\ &\implies 4xy + 2y^2 = 2xy + 2x^2 \\ &\implies 2y^2 = 2x^2 \\ &\implies x^2 = y^2. \end{aligned}$$

The constraint then implies

$$\begin{aligned} x^2 + x^2 &= 1 \\ 2x^2 &= 1 \\ x^2 &= \frac{1}{2} \end{aligned}$$

$$x = \pm\sqrt{\frac{1}{2}}$$

We now have four solutions, described by

$$(\pm\sqrt{1/2}, \pm\sqrt{1/2}).$$

- $f(-,-) = 3/2$
- $f(-,+) = 1/2$
- $f(+,-) = 1/2$
- $f(+,+) = 3/2$

Since the circumference is closed and f is bounded on the circumference, $1/2$ is the minimum and $3/2$ is the maximum on the circumference.

In total, 0 is the global minimum and $3/2$ is the maximum.

Example 3.21: MT Section 3.4 Problem 15

Find the extrema of

$$f(x,y) = 4x + 2y$$

subject to the constraint

$$2x^2 + 3y^2 = 21.$$

The constraint describes some kind of ellipse, which is closed, and f is probably bounded on it, so we may proceed.

$$\begin{aligned} \nabla f &= \lambda \nabla g \\ (4,2) &= \lambda(4x + 6y) \\ (4,2) &= (4\lambda x, 6\lambda y) \end{aligned}$$

Again, solving for λ suffices,

$$\begin{aligned} (\lambda, \lambda) &= (4/4x, 6/2y) \\ x &= 3y \end{aligned}$$

Then the constraint gives

$$\begin{aligned} 2(3y)^2 + 3y^2 &= 21 \\ 18y^2 + 3y^2 &= 21 \\ 21y^2 &= 21 \\ y^2 &= 1 \\ y &= \pm 1. \end{aligned}$$

Then $x = \pm 3$.

Evaluating gives 14 and -14, which are the maximum and minimum.

4 "Vector-Valued Functions" (MT 4)

We have spent much time discussing functions of multiple variables that map to a scalar, that is, functions

$$f : \mathbb{R}^n \mapsto \mathbb{R}$$

which are important as since they map to a scalar, there is a meaningful description of an extrema, since the scalars are comparable. Further, many functions are naturally workable with the idea of "take in many variables, and yield a number." We now turn our attention over to functions that result in multiple numbers, or really, vectors. The most common application of such functions are where vectors are the most present – physics. If we describe some multidimensional domain, usually space itself, and we associate to each point a vector, we can describe this vector as some quantity of force.

4.1 "Differentiation of Vector Functions" (MT 4.1)

Differentiation of vector based functions is the same, except the result is now a tuple of derivatives. We differentiate each component.

Do note that when we discuss the tangent line of vector functions, we no longer use the

$$L(t) + f(t_0) + (t - t_0)f'(t_0)$$

form, but

$$L(s) + f(t_0) + sf'(t_0)$$

instead. That is, we don't consider a time offset anymore.

Example 4.1: MT Section 4.1 Problem 3

Find the velocity and acceleration vectors and the equation of the tangent line for

$$r(t) = \sqrt{2}t\mathbf{i} + e^t\mathbf{j} + e^{-t}\mathbf{k}$$

at $t = 0$.

We first write this in component form,

$$r(t) = (\sqrt{2}t, e^t, e^{-t}).$$

Then, we obtain the velocity vector by differentiating each component with respect to t,

$$r'(t) = (\sqrt{2}, e^t, -e^{-t}$$

and same with the acceleration vector,

$$r''(t) = (0, e^t, e^{-t}).$$

Evaluating,

$$r(0) = (0, 1, 1)$$

$$r'(0) = (\sqrt{2}, 1, -1)$$
$$r''(0) = (0, 1, 1).$$

The tangent line is then given as

$$\begin{aligned} L(s) &= r(0) + sr'(0) \\ &= (0, 1, 1) + s(\sqrt{2}, 1, -1) \\ &= (s\sqrt{2}, 1 + s, 1 - s). \end{aligned}$$

Example 4.2: MT Section 4.1 Problem 4

Find the velocity and acceleration vectors and the equation of the tangent line for

$$c(t) + t\mathbf{i} + tj + \frac{2}{3}t^{3/2}\mathbf{k}$$

at $t = 9$.

In component form, this is $(t, t, 2/3t^{3/2})$. Differentiating, we have

$$c'(t) = (1, 1, t^{1/2})$$
$$c''(t) = (0, 0, 1/2t^{-1/2}).$$

Evaluating, we have

$$c(9) = (9, 9, 18)$$
$$c'(9) = (1, 1, 3)$$
$$c''(9) = (0, 0, 1/6).$$

The tangent line is then

$$\begin{aligned} L(s) &= (9, 9, 18) + s(1, 1, 3) \\ &= (9 + s, 9 + s, 18 + 3s). \end{aligned}$$

Example 4.3: MT Section 4.1 Problem 6

Let

$$c_1(t) = e^t\mathbf{i} + \sin(t)\mathbf{j} + t^3\mathbf{k}$$
$$c_2(t) = e^{-t}\mathbf{i} + \cos(t)\mathbf{j} - 2t^3\mathbf{k}.$$

Find

$$\frac{d}{dt}\left[c_1 \cdot c_2\right]$$

in two different ways to verify derivative rules.

In component form, these are

$$\begin{aligned} c_1(t) &= (e^t, \sin(t), t^3) \\ c_2(t) &= (e^{-t}, \cos(t), -2t^3). \end{aligned}$$

We can first proceed by going inside – the dot product – then out by differentiating the result:

$$\begin{aligned} \frac{d}{dt}[c_1 \cdot c_2] &= \frac{d}{dt}\left[(e^t, \sin(t), t^3) \cdot (e^{-t}, \cos(t), -2t^3)\right] \\ &= \frac{d}{dt}\left[e^0 + \sin(t)\cos(t) - 2t^6\right] \\ &= \cos^2(t) - \sin^(t) - 12t^5. \end{aligned}$$

Then we can use the product rule, which yields

$$\begin{aligned} \frac{d}{dt}[c_1 \cdot c_2] &= c_1' \cdot c_2 + c_1 \cdot c_2' \\ &= (e^{-t}, \cos(t), -2t^3) \cdot \frac{d}{dt}\left[(e^t, \sin(t), t^3)\right] + (e^t, \sin(t), t^3) \cdot \frac{d}{dt}(e^{-t}, \cos(t), -2t^3) \\ &= (e^{-t}, \cos(t), -2t^3) \cdot (e^t, \cos(t), 3t^2) + (e^t, \sin(t), t^3) \cdot (-e^{-t}, -\sin(t), -6t^2) \\ &= e^0 + \cos^2(t) - 6t^5 + -e^0 - \sin^2(t) - 6t^5 \\ &= \cos^2(t) - \sin^2(t) - 12t^5 \end{aligned}$$

which agrees!

Example 4.4: MT Section 4.1 Problem 9

Consider the helix given by

$$c(t) = (a\cos(t), a\sin(t), bt).$$

Show that the acceleration vector is always parallel to the xy-plane.

Differentiating gives

$$\begin{aligned} c'(t) &= (-a\sin(t), a\cos(t), b) \\ c''(t) &= (-a\cos(t), -a\sin(t), 0). \end{aligned}$$

The normal vector to the xy-plane is any vector of the form $(0, 0, k)$ for some constant k. Taking the dot product of this with the acceleration vector gives

$$\begin{aligned} (0, 0, k) \cdot c''(t) &= (0, 0, k) \cdot (-a\cos(t), -a\sin(t), 0) \\ &= 0 \end{aligned}$$

so the acceleration vector is orthogonal to the normal vector of the xy-plane and is hence parallel to the xy-plane.

Example 4.5: MT Section 4.1 Problem 10

Prove that

$$\frac{d}{dt}(c_1 \cdot c_2) = c_1' c_2 + c_1 c_2'$$

for any differentiable functions c_1, c_2.

By the dot product definition,

$$c_1 \cdot c_2 = \sum_{\forall i} (c_1)_i (c_2)_i.$$

Then, by linearity of differentiation and the product rule, we have that

$$\begin{aligned} \frac{d}{dt} \sum_{\forall i} (c_1)_i (c_2)_i &= \sum_{\forall i} \frac{d}{dt} \left[(c_1)_i (c_2)_i \right] \\ &= (c_1)_i' (c_2)_i + (c_1)_i (c_2)_i' \\ &= c_1' \cdot c_2 + c_1 \cdot c_2'. \end{aligned}$$

4.2 "Arc Length" (MT 4.2)

For some path $c(t)$, the magnitude of the velocity vector,

$$\|c'(t)\|$$

refers to the speed.

Then, since speed refers to an instantaneous distance, the sum aggregate of distances is the total distance traveled across the path, which we rigorously define as the *arc length of the path*.

Definition 4.1: Arc-Length of a Path

Let $c(t)$, where $t_0 \leq t \leq t_1$ be a path $f : \mathbb{R} \mapsto \mathbb{R}^n$ of class C^1. Then the arc length of the path is defined as

$$L(c) = \int_{t_0}^{t_1} \|c'(t)\| \, dt.$$

We describe the speed as an instantaneous distance, and since we sum over this instantaneous quantity, this is naturally an integral. For posterity when we cover line integrals later, we denote this arc-length differential explicitly:

Definition 4.2: Arc-Length Differential

An infinitesimal displacement of a particle following a C^1 path $c(t)$ is given by $c'(t)$ and its arc length

$$ds = \|c'(t)\|$$

is the differential of arc length.

This can be seen by considering the point $c(t)$, considering $c(t+\varepsilon)$, which leads to a c_i' displacement in the i-th component, of which the total displacement is then given by the inner product norm/distance formula. We can then super compactly define arc length as

$$\int_{t_0}^{t_1} ds.$$

Example 4.6: MT Section 4.2 Problem 3

Find the arc length of

$$(\sin 3t, \cos 3t, 2t^{3/2}$$

for $0 \le t \le 1$.

The derivative is $c'(t) = (3\cos 3t, -3\sin 3t, 3t^{1/2})$, of which the norm is

$$\begin{aligned} \|c'(t)\| &= \sqrt{9\cos^2 3t + 9\sin^2 3t + 9t} \\ &= \sqrt{9+9t} \\ &= 3\sqrt{1+t} \end{aligned}$$

Integrating, we have

$$\begin{aligned} L(c) &= \int_0^1 3\sqrt{1+t}\, dt \\ &= 3\int_0^1 \sqrt{t+1}\, dt \\ &= 3\left(\frac{2}{3}(t+1)^{3/2}\right)\Bigg|_{t=0}^{t=1} \\ &= 2(2^{3/2} - 1^{3/2}) \\ &= 4\sqrt{2} - 2. \end{aligned}$$

Example 4.7: MT Section 4.2 Problem 7

Find the arc length of

$$c(t) = (t, |t|)$$

for $-1 \le t \le 1$.

Normally this is done by using the piecewise characterization of the absolute value function to write $c(t)$ as

$$c(t) = \begin{cases} (t, -t) & t < 0 \\ (t, t) & t > 0 \end{cases}$$

but this is actually not necessary. Bear with me, let

$$\operatorname{sgn}(x) = \begin{cases} -1 & x < 0 \\ 0 & x = 0 \\ 1 & x > 1 \end{cases}.$$

Then the derivative of the absolute value function is compactly

$$\frac{d}{dt}|t| = \operatorname{sgn}(t)$$

except when $t \neq 0$, where $|t|$ fails to be differentiable.

Then

$$c'(t) = (1, \operatorname{sgn}(t)).$$

The norm, since sgn^2 is 1 except when $t = 1$, becomes

$$\|c'(t)\| = \sqrt{1+1} = \sqrt{2}.$$

Then the arc-length integral becomes

$$\begin{aligned} L(c) &= \int_{-1}^{1} \sqrt{2}\, dt \\ &= 2\sqrt{2}. \end{aligned}$$

I note that while this certainly looks like significantly more work than splitting the path, you wouldn't actually write down the definition of the signum function or the derivative fo the absolute value. In most problems, though, splitting the interval of the path is useful.

Example 4.8: MT Section 4.2 Problem 9

Let C be the line segment connecting the point $p = (1, 2, 0)$ to the point $q = (0, 1, -1)$.

(a) Find a curve $c(t) : [a, b] \subset \mathbb{R} \mapsto \mathbb{R}^3$ that traces out C.

(b) Find the arc length of $c(t)$

(c) Find $\|p - q\|$.

(a) This is given as $c(t) = p + t(q - p)$ for $0 \leq t \leq 1$, which gives

$$c(t) = (1, 2, 0) + t(-1, -1, -1)$$

for $0 \leq t \leq 1$.

(b) The derivative is $c'(t) = (-1, -1, -1)$. The arc length differential is then $\|c'(t)\| = \sqrt{3}$. The arc length integral is then

$$\begin{aligned} L(c) &= \int_{0}^{1} \sqrt{3}\, dt \\ &= \sqrt{3}. \end{aligned}$$

(c) $p - q = (1, 1, 1)$, so $\|p - q\| = \sqrt{3}$, which agrees with the arc length.

Example 4.9: MT Section 4.2 Problem 13

Let c be the path

$$c(t) = (2t, t^2, \log t)$$

defined for $t > 0$. Find the arc length of c between the points $(2, 1, 0)$ and $(4, 4, \log 2)$.

The points correspond to $t = 1$ and $t = 2$, since $\log 1 = 0$ and $\log 2 = \log 2$. The derivative is

$$c'(t) = (2, 2t, 1/t).$$

Then the arc length differential is given as

$$\begin{aligned} \|c'(t)\| &= \sqrt{2^2 + (2t)^2 + (1/t)^2} \\ &= \sqrt{4 + 4t^2 + (1/t)^2} \\ &= \sqrt{(2t + 1/t)^2} \\ &= |2t + 1/t| \\ &= 2t + 1/t \end{aligned}$$

since $t > 0$, so the inner expression is always positive.

Integrating,

$$\begin{aligned} L(c) &= \int_1^2 2t + 1/t \, dt \\ &= t^2 + \ln|t| \Big|_{t=1}^{t=2} \\ &= 4 - 1 + \ln 2 - \ln 1 \\ &= 3 + \ln 2. \end{aligned}$$

Example 4.10: MT Section 4.2 Problem 16

Let

$$c : [a, b] \subset \mathbb{R} \mapsto \mathbb{R}^3$$

be an infinitely differentiable path. Assume $c'(t) \neq 0$ for all t. The vector

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

is tangent to c at $c(t)$ and because $\|T(t)\|$, T is called the unit tangent vector to c.

(a) Show that $T'(T) \cdot T(t) = 0$
  - Hint: Differentiate $T(t) \cdot T(t) = 1$

(b) Write down a formula for $T'(t)$ in terms of c.

When I took multivariate calculus for the first time, this was a far more important concept. Too bad it's relegated to being defined in only a single problem.

(a) Differentiating $T(t) \cdot T(t) = 1$ by the dot product rule gives

$$\begin{aligned} \frac{d}{dt}\left[T(t) \cdot T(t)\right] &= \frac{d}{dt}\left[1\right[\\ T'(t) \cdot T(t) + T(t) \cdot T'(t) &= 0 \\ 2T'(t) \cdot T(t) &= 0 \\ T'(t) \cdot T(t) &= 0. \end{aligned}$$

(b) If we directly try to differentiate $T(t)$, by the quotient rule, we obtain

$$T'(t) = \frac{\|c'(t)\|c''(t) - c'(t)\frac{d}{dt}\|c'(t)\|}{\|c'(t)\|^2}$$

which is almost there, but we need a way to simplify the derivative of the norm of the tangent vector. I present two ways to do this, a direct way, and an indirect way that differentiates the square of the norm. In either case, we use that $\|c'(t)\| = \sqrt{c'(t) \cdot c'(t)}$.

- Differentiating the norm of the tangent vector,

$$\begin{aligned} \frac{d}{dt}\|c'(t)\| &= \frac{d}{dt}\sqrt{c'(t) \cdot c'(t)} \\ &= \frac{1}{2\sqrt{c'(t) \cdot c'(t)}}\frac{d}{dt}(c'(t) \cdot c'(t)) \\ &= \frac{c'(t) \cdot c''(t) + c''(t) \cdot c'(t)}{2\|c'(t)\|} \\ &= \frac{c'(t) \cdot c''(t)}{\|c'(t)\|} \end{aligned}$$

- Differentiating the square of the norm of the tangent vector,

$$\frac{d}{dt}\|c'(t)\|^2 = 2\|c'(t)\|\frac{d}{dt}\|c'(t)\|$$

Since $\|c'(t)\|^2 = c(t) \cdot c(t)$, $\frac{d}{dt}\|c'(t)\| = c'(t) \cdot c''(t) + c''(t) \cdot c'(t) = 2c'(t) \cdot c''(t)$. Then we have that

$$(2c'(t) \cdot c''(t)) = 2\|c'(t)\|\frac{d}{dt}\|c'(t)\|$$

and we can solve for $\|c'(t)\|$ as

$$\begin{aligned} \|c'(t)\| &= \frac{2c'(t) \cdot c''(t)}{2\|c'(t)\|} \\ &= \frac{c'(t) \cdot c''(t)}{\|c'(t)\|} \end{aligned}$$

Then, we have that

$$T'(t) = \frac{\|c'(t)\|c''(t) - c'(t)\frac{c'(t) \cdot c''(t)}{\|c'(t)\|}}{\|c'(t)\|^2}$$

$$= \frac{\|c'(t)\|^2 c''(t) - c'(t)(c'(t) \cdot c''(t))}{\|c'(t)\|^3}$$

We're more or less done here, but we can do more. Factoring out the inverse of the norm of the tangent vector, we have

$$\begin{aligned} T'(t) &= \frac{1}{\|c'(t)\|} \left(\frac{\|c'(t)\| c''(t)}{\|c'(t)\|} - \frac{c'(t) \cdot (c'(t) \cdot c''(t))}{\|c'(t)\|^2} \right) \\ &= \frac{1}{\|c'(t)\|} \left(c''(t) - \frac{c'(t) \cdot (c'(t) \cdot c''(t))}{\|c'(t)\|^2} \right) \end{aligned}$$

Note that the second term in the parentheses, when written as

$$\frac{c'(t) \cdot c''(t)}{\|c'(t)\|} c'(t)$$

is the orthogonal projection of $c''(t)$ onto $c'(t)$.

So we have a vector subtracted by the orthogonal projection of the same vector onto another vector. We claim that the aforementioned expression is itself orthogonal to the other vector. To make things simpler, this is equivalent to claiming the expression

$$w - \frac{v \cdot w}{\|v\|^2} v$$

is orthogonal to v, where $w = c''(t)$ and $v = c'(t)$. To see why, we dot the entire expression with respect to v, and get

$$\begin{aligned} v \cdot (w - \operatorname{proj}_v w) &= v \cdot w - \frac{(v \cdot w)(v \cdot v)}{\|v\|^2} \\ &= v \cdot w - \frac{(v \cdot w)\|v\|^2}{\|v\|^2} \\ &= v \cdot w - v \cdot w \\ &= 0. \end{aligned}$$

Since the vectors are orthogonal, their inner angle $\theta = 90°$, and importantly, $\sin(\theta) = 1$. Using the cross product sine formula on v and $w - \operatorname{proj}_v u$, we have that

$$\begin{aligned} \|v \times w\| &= \|v\|\|w\| \sin(\theta) \\ &= \|v\|\|w\| \\ &= \|v\|\|w - \operatorname{proj}_v u\| \\ \frac{\|v \times w\|}{\|v\|^2} &= \frac{1}{\|v\|} \|w - \operatorname{proj}_v w\| \end{aligned}$$

which, since $v = c'(t)$ and $w = c''(t)$, is the expression we have for $T'(t)$.

Thus we have that

$$T'(t) = \frac{\|c'(t) \times c''(t)\|}{\|c'(t)\|^2}.$$

This is the equation you may work with in a multi-variable calculus setting.

4.3 "Vector Fields" (MT 4.3)

Extrema dealt with functions of the form

$$f : \mathbb{R}^n \mapsto \mathbb{R}$$

and arc length dealt with functions of the form

$$f : \mathbb{R} \mapsto \mathbb{R}^n.$$

We now discuss the fully general functions of the form

$$f : \mathbb{R}^n \mapsto \mathbb{R}^m.$$

This is a pretty short section, as we just define a few ideas. We will talk about things to do with such mathematical objects when we talk about divergence, gradient and curl, which is important enough to be the name of a book.

Definition 4.3: Vector Fields

A vector field in $\mathbb{R}^n$ is a map

$$F : A \subset \mathbb{R}^n \mapsto \mathbb{R}^n$$

that assigns each point x in its domain A a vector $F(x)$ in the same dimension.

Unfortunately, that the dimensions are the same is critical for much of vector analysis, and we generally not discuss functions that do not much else for the rest of the text, except for compositions.

The most readily applicable concept for vector fields is that of fluid flow/fluid dynamics. For each point in space, we associate a vector that describes the movement of fluid per unit time at that point in space. This is explored rather beautifully in 3Blue1Brown's Divergence and curl: The Language of Maxwell's equations, fluid flow, and more. For visualizations, I highly recommend Andrei Kashcha's Field Play applet.

4.3.1 Gradient Vector Fields

There is a special important case of vector fields when the vector field can be understood as the gradient of a scalar function, that is,

$$F = \nabla f.$$

Definition 4.4: Gradient Vector Field

A vector field F is called a *gradient vector field* if there exists some scalar function f such that

$$F = \nabla f.$$

Generally, this arises when f is a force function (and hence F may be called a force field).

4.3.2 Flow Lines

Definition 4.5

For a given vector field F, a *flow line* for F is a path $c(t)$ such that

$$c'(t) = F(c(t)).$$

Example 4.11: MT Section 4.3 Problem 18

Show that the given curve

$$c(t) = \left(\frac{1}{t^3}, e^t, \frac{1}{t}\right)$$

is a flow line of the given velocity vector field

$$F(x, y, z) = (-3z^4, y, -z^2).$$

Let's first compose,

$$\begin{aligned} F(c(t)) &= (-3(1/t)^4, e^t, -(1/t)^2) \\ &= \left(\frac{3}{t^4}, e^t, -\frac{1}{t^2}\right). \end{aligned}$$

Now let's differentiate,

$$c'(t) = \left(-\frac{3}{t^4}, e^t, -\frac{1}{t^2}\right)$$

which are equal, so $c(t)$ is a flow line for F.

4.4 "Divergence and Curl" (MT 4.4)

We've seen ∇ a bunch of times. We can now give it a proper (albeit not completely rigorous) definition:

Definition 4.6: Del operator

In $\mathbb{R}^n$, the del operator, ∇, is defined as

$$\left(\frac{\partial}{\partial x_i}\right) = \left(\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_n}\right).$$

In $\mathbb{R}^3$, for example, the del operator is

$$\left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right).$$

We also define the Laplacian as the second-order analogue of the del operator

Definition 4.7: Laplacian operator

In $\mathbb{R}^n$, the Laplacian operator, ∇^2, is defined as

$$\left(\frac{\partial^2}{\partial x_i^2}\right).$$

4.4.1 Gradient

Using this, we can actually define the gradient of a vector function by operating on it–

$$\nabla f = \left(\frac{\partial f}{\partial x_i}\right).$$

Through an abuse of notation, we can define other important operators of vector calculus.

4.4.2 Divergence

Definition 4.8: Divergence of a Vector Field

For a given vector field F, the divergence of F is the *scalar* function

$$\operatorname{div} F = \nabla \cdot F = \sum_{\forall i} \frac{\partial}{\partial x_i} F_i.$$

The divergence represents, for a given region, the net quantity of fluid leaving the region. Another idea is, for a given region, the expansion of that region when considering where the fluid particles in that region after applying the vector field for a step of time.

Example 4.12: MT Section 4.4 Problem 1

Find the divergence of the vector field

$$V(x, y, z) = e^{xy}\mathbf{i} - e^{xy}\mathbf{j} + e^{yz}\mathbf{k}.$$

In component form, this is $V(x, y, z) = (e^{xy}, -e^{xy}, e^{yz})$. Then, the divergence is given by

$$\begin{aligned}\operatorname{div} V &= \frac{\partial}{\partial x}\left[e^{xy}\right] + \frac{\partial}{\partial y}\left[-e^{xy}\right] + \frac{\partial}{\partial z}\left[e^{yz}\right] \\ &= ye^{xy} - xe^{xy} + ye^{yz}\end{aligned}$$

Example 4.13: MT Section 4.4 Problem 4

Find the divergence of the vector field

$$V(x, y, z) = x^2\mathbf{i} + (x + y)^2\mathbf{j} + (x + y + z)^2\mathbf{k}.$$

In component form, this is $V(x, y, z) = (x^2, (x+y)^2, (x+y+z)^2)$. Then, the divergence is given by

$$\begin{aligned} \operatorname{div} V &= 2x + 2(x+y) + 2(x+y+z) \\ &= 6x + 4y + 2z. \end{aligned}$$

Example 4.14: MT Section 4.4 Problem 27

Suppose $f, g, h : \mathbb{R}^2 \mapsto \mathbb{R}$ are differentiable. Show that the vector field

$$F(x, y, z) = (f(y, z), g(x, z), h(x, y))$$

has zero divergence.

Applying the definition of divergence, we obtain

$$\begin{aligned} \operatorname{div} F &= \frac{\partial}{\partial x} f(y, z) + \frac{\partial}{\partial y} g(x, z) + \frac{\partial}{\partial z} h(x, y) \\ &= 0 + 0 + 0 \\ &= 0 \end{aligned}$$

Since f, g and h are not functions of x, y and z and hence are constants with respect to x, y and z respectively.

4.4.3 Curl

Definition 4.9: Curl of a Vector Field

For a given vector field $F : \mathbb{R}^3 \mapsto \mathbb{R}^3$, the curl F is the vector field

$$\operatorname{curl} F = \nabla \times F = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ F_1 & F_2 & F_3 \end{vmatrix} = ((F_3)_y - (F_2)_z, (F_1)_z - (F_3)_x, (F_2)_x - (F_1)_y).$$

Note that, since in this context, the cross product is only defined in 3 dimensions (or fewer dimensions when viewed as a 3 dimensional object with other components 0), the curl is only defined in 3 dimensions (or lower). It is really this limitation and that divergence is only meaningful when the dimensions are the same why we generally only consider vector fields of the form

$$F : \mathbb{R}^3 \mapsto \mathbb{R}^3.$$

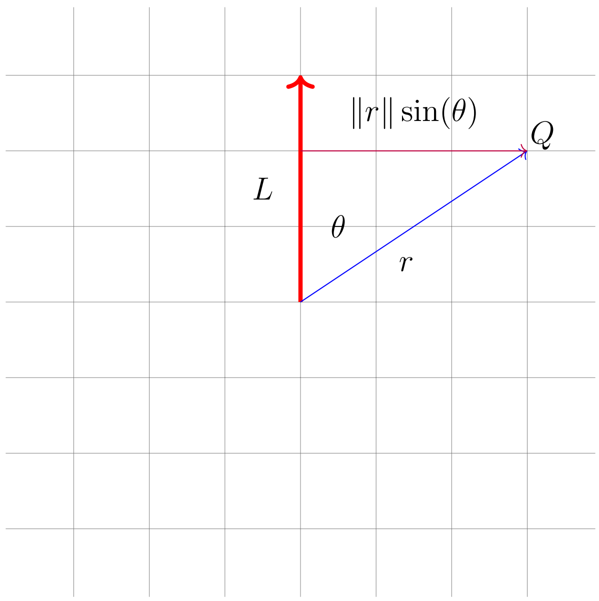
Curl describes the rotation of a vector field; if we drop a twig into a point on the vector field, the curl of the vector field at the point describes roughly how much the twig ends up rotating relative to its original position.

Theorem 4.1: The Curl is Twice the Angular Velocity

Consider a solid rigid body B rotating about an axis L. The rotational motion of the body can be described by a vector ω along the axis of rotation, the direction chose so that the body rotates about ω; that is, ω is on the axis of rotation. The length $\|\omega\|$ is the angular speed of B, that is, the speed of any point on B divided by its distance from the axis of rotation.
The motion of points in the rotating body is described by a vector field v whose value at each point is the velocity at each point.
Then

$$\operatorname{curl} v = 2\omega.$$

To see this, let's first identify v. Let Q be any point in B and let α be the distance from Q to L. If we let r be the vector from the origin to Q and θ be the angle between r and the axis of rotation L. Then the distance is $\alpha = \|r\| \sin\theta$.



The tangential velocity is then the product of the rotational rate of the entire body, $\|\omega\|$ and the distance to L, or

$$\|v\| = \|\omega\|\|r\|\sin(\theta)$$

which then implies $v = \omega \times r$.

If we construct a coordinate system where L is on the z-axis, then $\omega = (0, 0, \|\omega\|)$ and let the position vector r be $r = (x, y, z)$. Then

$$\begin{aligned} v &= \omega \times r \\ &= (0, 0, \|\omega\|) \times (x, y, z) \\ &= \|\omega\|(-y, x, 0). \end{aligned}$$

Then computing the curl gives

$$\begin{aligned} \operatorname{curl} v &= \|\omega\| \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ dx & dy & dz \\ -y & x & 0 \end{vmatrix} \\ &= \|\omega\|(0,0,2) \\ &= 2\omega. \end{aligned}$$

Example 4.15: MT Section 4.4 Problem 14

Compute the curl of

$$F(x,y,z) = yz\mathbf{i} + xz\mathbf{j} + xy\mathbf{k}.$$

In component form, this is (yz, xz, xy). I find that when computing curl, it may actually be preferable to memorize the special case cross product for simpler functions.

$$\begin{aligned} \operatorname{curl} F &= \left(\frac{\partial}{\partial y}xy - \frac{\partial}{\partial z}xz, \frac{\partial}{\partial z}yz - \frac{\partial}{\partial x}xy, \frac{\partial}{\partial x}xz - \frac{\partial}{\partial y}yz\right) \\ &= (x-x, y-y, z-z) \\ &= (0,0,0) \end{aligned}$$

Do note that when using the special case, you ought to not write the first step down, otherwise I imagine the writing-processing trade-off to no longer be worth it.

Example 4.16: MT Section 4.4 Problem 15

Compute the curl of

$$F(x,y,z) = (x^2+y^2+z^2)(3\mathbf{i}+4\mathbf{j}+5\mathbf{k}).$$

In component form, this can be expressed in one of two ways:

$$F(x,y,z) = (x^2+y^2+z^2)(3,4,5) = (3x^2+3y^2+3z^2, 4x^2+4y^2+4z^2, 5x^2+5y^2+5z^2).$$

Amazingly, since these are all linear combinations of purely x, y and zs with no mixed terms, the curl is rather simple to compute.

$$\operatorname{curl} F = (10y-8z, 6z-10x, 8x-6y).$$

Definition 4.10: Irrotational vector fields

A vector field F is called *irrotational* if

$$\operatorname{curl} F = 0.$$

Note that irrotational means a solid body does not rotate when acted upon by the vector field. The solid body can still be displaced in a rotational path, such as a circle.

4.4.4 Scalar Curl

Though I've alluded to using curl in lower dimensions by representing it in higher dimensions with component values zero, this special case is worthwhile enough to have a special name.

Definition 4.11: Scalar curl

For a vector field $F : \mathbb{R}^3 \mapsto \mathbb{R}^3$ with z-component 0 (or alternatively, a function $F : \mathbb{R}^2 \mapsto \mathbb{R}^3$ represented in $\mathbb{R}^3$), we define the scalar curl of F as

$$\operatorname{curl} F \cdot \mathbf{k}.$$

If

$$F = P\mathbf{i} + Q\mathbf{j},$$

then the scalar curl of F is given as

$$Q_x - P_y.$$

Keep this in mind when we encounter Green's theorem (and Stoke's theorem) later.

Example 4.17: MT Section 4.4 Problem 20

Compute the scalar curl of the vector field

$$F(x, y) = x\mathbf{i} + y\mathbf{j}$$

This is just

$$\begin{aligned} \frac{\partial}{\partial x} y - \frac{\partial}{\partial y} x &= 0 - 0 \\ &= 0. \end{aligned}$$

4.4.5 Relationships of Vector Field Operators

Example 4.18: MT Section 4.4 Problem 25

Suppose $F : \mathbb{R}^3 \mapsto \mathbb{R}^3$ is a C^2 vector field. Which of the following expressions are meaningful, and which are nonsense? For those which are meaningful, decide whether the expression defines a scalar function or a vector field.

(a) $\operatorname{curl} \operatorname{grad} F$
(b) $\operatorname{grad} \operatorname{curl} F$
(c) $\operatorname{div} \operatorname{grad} F$
(d) $\operatorname{grad} \operatorname{div} F$
(e) $\operatorname{curl} \operatorname{div} F$
(f) $\operatorname{div} \operatorname{curl} F$

To make this quick and painless, let me review the domains and codomains of our vector function operators:

- The gradient takes a scalar function and maps it to a vector function.
- The divergence takes a vector function and maps it to a scalar function.
- The curl takes a vector function and maps it to a vector function.

(a) Nonsense. The gradient of a vector field is ill-defined.

(b) Nonsense. The curl of a vector field is a vector field, but the gradient of a vector field is ill-defined.

(c) Nonsense. The gradient of a vector field is ill-defined.

(d) Meaningful. The divergence of a vector field is a scalar function, and the gradient of a scalar function is a vector field.

(e) Nonsense. The divergence of a vector field is a scalar function, but the curl of a scalar function is ill-defined.

(f) Meaningful. The curl of a vector field is a vector field, and the divergence of a vector field is a scalar function.

Theorem 4.2: Gradients are Curl Free

For any C^2 function f,

$$\operatorname{curl}(\operatorname{grad} F) = \nabla \times (\nabla F) = 0.$$

The converse is also something important, but is more restrictive. We will discuss that when we cover the fundamental theorem of line integrals.

Example 4.19: MT Section 4.4 Problem 22a

Which of the vector fields could be gradient fields?

- $F(x, y, z) = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$
- $F(x, y, z) = yz\mathbf{i} + xz\mathbf{j} + xy\mathbf{k}$
- $F(x, y, z) = (x^2 + y^2 + z^2)(3\mathbf{i} + 4\mathbf{j} + 5\mathbf{k})$
- $F(x, y, z) = \frac{yz\mathbf{i} - xz\mathbf{j} + xy\mathbf{k}}{x^2+y^2+z^2}$

If F is a gradient field, then $F = \nabla f$ for some scalar function f. Since gradients are curl free, then $\operatorname{curl} F$ must be 0, or else we have some function f where $\operatorname{curl}\operatorname{grad} f \neq 0$, which is a contradiction.

- $\operatorname{curl} F = (0, 0, 0)$, so F could be a gradient field.
- We've computed this before; $\operatorname{curl} F = (0, 0, 0)$, so F could be a gradient field.
- We've computed this before; $\operatorname{curl} F = (10y - 8z, 6z - 10x, 8x - 6y)$, which is not zero, so F cannot be a gradient field.
- Just trust me it's non-zero, so F could not be a gradient field.

Theorem 4.3: Curls are Divergence Free

For any C^2 vector field F,

$$\text{div}(\text{curl}\, F) = \nabla \cdot (\nabla \times F) = 0.$$

Example 4.20: MT Section 4.4 Problem 22b

Which of the vector fields could be the curl of some vector field $V : \mathbb{R}^3 \mapsto \mathbb{R}^3$?

- $F(x, y) = x^3\mathbf{i} - x\sin(xy)\mathbf{j}$
- $F(x, y) = y\mathbf{i} - x\mathbf{j}$
- $F(x, y) = \sin(xy)\mathbf{i} - \cos(x^2y)\mathbf{j}$
- $F(x, y) = xe^y\mathbf{i} - (y/(x + y))\mathbf{j}$

If F is the curl of some vector field, then $\text{div}\, F$ must be equal to 0, as otherwise we would have that $\text{div}\,\text{curl}\, F = 0$ which is a contradiction.

- $\text{div}\, F = 3x^2 - x^2\cos(xy)$ which is not zero, so F cannot be the curl of a vector field.
- $\text{div}\, F = 0 + 0 = 0$ which is zero, so F can be the curl of a vector field.
- $\text{div}\, F = y\cos(xy) + x^2\sin(x^2y)$ which not zero, so F cannot be the curl of a vector field.
- $\text{div}\, F = e^y - \frac{x}{(x+y)^2}$ which is not zero, so F cannot be the curl of a vector field.

Theorem 4.4: Vector Analysis Identities

Let f and g be scalar functions, F and G be vector functions, and c be a constant.

- $\nabla(f + g) = \nabla f + \nabla g$ (del distributes over addition)
- $\nabla(cf) = c\nabla f$ (del distributes over scalar multiplication
- $\nabla(fg) = f\nabla g + g\nabla f$ (del product rule)
- $\nabla(f/g) = (g\nabla f - f\nabla g)/g^2$, when $g \neq 0$ (del quotient rule)
- $\text{div}(F + G) = \text{div}\, F + \text{div}\, G$ (divergence distributes over addition)
- $\text{curl}(F + G) = \text{curl}\, F + \text{curl}\, G$ (curl distributes over addition)
- $\text{div}(fF) = f\,\text{div}\, F + F \cdot \nabla f$ (divergence scalar product rule)
- $\text{div}(F \times G) = G \cdot \text{curl}\, F - F \cdot \text{curl}\, G$ (divergence cross product rule)
- $\text{div}\,\text{curl}\, F = 0$ (curls are divergence free)
- $\text{curl}(fF) = f\,\text{curl}\, F + \nabla f \times F$ (curl scalar product rule)
- $\text{curl}\,\nabla f = 0$ (gradients are curl free)
- $\nabla^2(fg) = f\nabla^2 g + g\nabla^2 f + 2(\nabla f \cdot \nabla g)$ (Laplacian product rule)
- $\text{div}(\nabla f \times \nabla g) = 0$ (the cross product of two gradient fields is a curl)
- $\text{div}(f\nabla g - g\nabla f) = f\nabla^2 g - g\nabla^2 f$ (I have no clue)

Example 4.21: MT Section 4.4 Problem 28

Prove

$$\text{div}(\nabla f \times \nabla g) = 0.$$

Let $F = \nabla f$ and $G = \nabla g$. Then applying the divergence cross product rule, this gives

$$\operatorname{div}(F \times G) = G \cdot \operatorname{curl} F - F \cdot \operatorname{curl} G.$$

Since $F = \nabla F$ and $G = \nabla g$, this becomes

$$\operatorname{div}(\nabla f \times \nabla g) = G \cdot \operatorname{curl} \nabla f - F \cdot \operatorname{curl} \nabla g.$$

However, since gradients are curl free, $\operatorname{curl} \nabla f = \operatorname{curl} \nabla g = 0$, so the entire expression is 0.

5 "Double and Triple Integrals" (MT 5)

We have discussed much of defining derivatives for multi-variable functions. Now we turn to defining integrals for multi-variable functions. As we have done before when first discussing differentiation, we will first only consider the case of

$$f : \mathbb{R}^n \mapsto \mathbb{R}$$

and will discuss integration over vector fields in the next sections.

When I first took multi-variable calculus, this proved to be the most difficult section, primarily because drawing in multiple dimensions is hard, and it's generally necessary to do most of these problems when one first encounters them.

5.1 "Introduction" (MT 5.1)

First, let's define a useful notation:

Definition 5.1: Cartesian Product

The Cartesian product of two sets X and Y, denoted $X \times Y$ is the set of all tuples with the first element from X and second element from Y,

$$X \times Y = \{(x, y) \mid x \in X, \quad y \in Y\}.$$

So when I say $[0, 1] \times [2, 3]$, this is the same as the simultaneous set of constraints $0 \leq x \leq 1$ and $2 \leq y \leq 3$.

Definition 5.2: Double Integrals

The volume of the region above R and under the graph of a non-negative function f is called the double integral of f (sometimes double is omitted) over R and is denoted by

$$\iint_R f(x, y)\, dA.$$

dA here is generally going to mean something like $dx\, dy$, but since, as we will see later, that the order is sometimes super important, I will generally prefer to leave that part ambiguous.

To actually compute something, if $R = [a, b] \times [c, d]$, then we have that

$$\iint_R f(x, y)\, dA = \int_a^b \left[\int_b^c f(x, y)\, dy\right] dx = \int_c^d \left[\int_a^b f(x, y)\, dx\right] dy.$$

This is actually Fubini's theorem which we cover in the next subsection.

Example 5.1: MT Section 5.1 Problem 1a

Evaluate the following iterated integral:

$$\int_0^1 \int_0^1 (1 - x^3 + xy)\, dx\, dy$$

We just proceed directly, remembering that x and y are constants with respect to each other:

$$\begin{aligned}
\int_0^1 \int_0^1 (1 - x^3 + xy)\, dx\, dy &= \int_0^1 \left[x - \frac{1}{4}x^4 + \frac{1}{2}x^2 y \right]_{x=0}^{x=1} dy \\
&= \int_0^1 (1-0) - (1/4 - 0) + (1/2y - 0)\, dy \\
&= \int_0^1 \frac{3}{4} + \frac{1}{2} y\, dy \\
&= \left[\frac{3}{4} y + \frac{1}{4} y^2 \right]_{y=0}^{y=1} \\
&= \frac{3}{4} + \frac{1}{4} \\
&= 1.
\end{aligned}$$

Example 5.2: MT Section 5.1 Problem 14

Find the volume bounded by the graph of $f(x,y) = 1 + 2x + 3y$, the rectangle $[1,2] \times [0,1]$ and the four vertical sides of the rectangle R.

We need to check what the bounds are. Since $f(x,y) > 0$ when $x, y > 0$, f is always above the rectangle, so the bounds are just the rectangle. Then the volume is given by integrating

$$\begin{aligned}
\int_1^2 \int_0^1 1 + 2x + 3y\, dy\, dx &= \int_1^2 \left[y + 2xy + \frac{3}{2} y^2 \right]_{y=0}^{y=1} dx \\
&= \int_1^2 1 + 2x + \frac{3}{2}\, dx \\
&= \int_1^2 \frac{5}{2} + 2x\, dx \\
&= \left[\frac{5}{2} x + x^2 \right]_{x=1}^{x=2} \\
&= 10 - \frac{5}{2} + 4 - 1 \\
&= \frac{21}{2}.
\end{aligned}$$

5.1.1 Cavalieri's Principle

For extremely weird cases where integrating normally is difficult but the geometric ideas are extremely clear, we can consider the integral of area slices of an object.

Theorem 5.1: Cavalieri's Principle

Let S be a solid, and for $a \leq x \leq b$, let P_x be a family of parallel planes such that S lies between P_a and P_b, the area of the slice cut by P_x is $A(x)$. Then the volume of S is equal to

$$\int_a^b A(x)\,dx.$$

Example 5.3: MT Section 5.3 Problem 15

Find the volume of the region inside the surface

$$z = x^2 + y^2$$

between $z = 0$ and $z = 10$.

Let's slice by z. For a given value z, the cross sectional area is the area of the region $x^2 + y^2 = z$. That is, the area of a circle with radius squared $r^2 = z$. As the area of a circle is πr^2, the cross sectional area for a particular z is

$$A(z) = \pi z.$$

Then, integrating gives

$$\begin{aligned}
\int_0^{10} A(z)\,dz &= \int_0^{10} \pi z\,dz \\
&= \pi \int_0^{10} z\,dz \\
&= \frac{\pi}{2}\left[z^2\right]_{z=0}^{z=10} \\
&= 50\pi.
\end{aligned}$$

5.2 "The Double Integral Over a Rectangle" (MT 5.2)

Theorem 5.2: Integrability is super lax

Any continuous function defined on a closed rectangle R is integrable.

This doesn't say anything about if the actual result is expressible in closed form, just that the integral has a value that exists.

We can get even more lax about whether the function needs to be defined:

Theorem 5.3: Integrability of Bounded Functions

Let $f : \mathbb{R} \mapsto \mathbb{R}$ be a bounded real-value function on the rectangle R and suppose that the set of points where f is discontinuous lies on a finite union of graphs of continuous functions. Then f is integrable over R.

In single variable calculus, we had the notion that if a function only had a countable number of points where it failed to be defined, the function is still integrable, since the contribution of an individual point to an integral is 0. We have a similar idea, but we can go a little further. Here, we require that the set of discontinuities must be over a finite *union of graphs*. I like to think that the region R has dimension 2, so if our discontinuities have dimension 0 (just points) or 1 (a path of discontinuities), it's not enough to matter.

Theorem 5.4: Fubini's Theorem

Let f be a continuous function with a rectangular domain $R = [a, b] \times [c, d]$. Then

$$\int_a^b \int_c^d f(x, y)\, dy\, dx = \int_c^d \int_a^b f(x, y)\, dx\, dy = \iint_R f(x, y)\, dA$$

Example 5.4: MT Section 5.2 Problem 1

Evaluate the following integral over $R = [0, 1] \times [0, 1]$:

$$\iint_R x^3 + y^2\, dA$$

$$\begin{aligned}
\int_0^1 \int_0^1 x^3 + y^2\, dx\, dy &= \int_0^1 \left[\frac{1}{4}x^4 + xy^2\right]_{x=0}^{x=1} dy \\
&= \int_0^1 \frac{1}{4} - 0 + y^2 - 0\, dy \\
&= \int_0^1 \frac{1}{4} + y^2\, dy \\
&= \left[\frac{1}{4}y + \frac{1}{3}y^3\right]_{y=0}^{y=1} \\
&= \frac{1}{4} - 0 + \frac{1}{3} - 0 \\
&= \frac{7}{12}.
\end{aligned}$$

Example 5.5: MT Section 5.2 Problem 1b

Evaluate the following integral over $R = [0,1] \times [0,1]$:

$$\iint_R ye^{xy}\,dA$$

This is the first time where choosing the order of integration matters. Integrating with respect to y is surely worse, as we have to do integration by parts right off the bat, which is definitely my least favorite integration technique, so let's integrate first by x and hope for the best.

$$\begin{aligned}
\int_0^1 \int_0^1 ye^{xy}\,dx\,dy &= \int_0^1 \left[e^{xy}\right]_{x=0}^{x=1} dy \\
&= \int_0^1 e^y - 1\,dy \\
&= \left[e^y - y\right]_{y=0}^{y=1} \\
&= (e-1) - (1-0) \\
&= e - 2.
\end{aligned}$$

Example 5.6: MT Section 5.2 Problem 4

Evaluate over the region $R : [0,1] \times [-2,2]$

$$\iint_R \frac{y}{1+x^2}\,dx\,dy$$

$$\begin{aligned}
\int_0^1 \int_{-2}^2 \frac{y}{1+x^2}\,dy\,dx &= \int_0^1 \frac{1}{1+x^2} \int_{-2}^2 y\,dy\,dx \\
&= \int_0^1 \frac{1}{1+x^2} \left[\frac{1}{2}y^2\right]_{y=-2}^{y=2} dx \\
&= \int_0^1 \frac{1}{1+x^2}(2-2)\,dx \\
&= 0.
\end{aligned}$$

5.3 "The Double Integral Over More General Regions" (MT 5.3)

Definition 5.3: Simple Regions

We say that a region is y-simple if there exist two continuous real valued functions $\phi_1 : [a,b] \mapsto \mathbb{R}$ and $\phi_2 : [a,b] \mapsto \mathbb{R}$ that satisfy $\phi_1(x) \leq \phi_2(x)$ and for all $(x,y) \in D$, $x \in [a,b]$ and $\phi_1(x) \leq y \leq \phi_2(x)$.

We define an x-simple region the same way. A region that is both x-simple and y-simple is a simple region with no qualifier.

A region that is simple in some way we often call an elementary region.

As more of a novelty theorem rather than one we actually wrestle with, we can extend the idea of rectangles to elementary regions (or really the other way around) to integrate arbitrary simple regions.

Definition 5.4: Integration over elementary regions

For a given elementary region D in the plane, pick an arbitrary rectangle R that contains D. Given $f : D \mapsto \mathbb{R}$ where f is continuous and bounded, we define $\iint_D f(x, y)\, dA$ as the integral of f over the set D by extending f to f^* over R:

$$f^* = \begin{cases} f(x,y) & (x,y) \in D \\ 0 & (x,y) \notin (x,y) \end{cases}$$

and we can then define

$$\iint_D f(x,y)\, dA = \iint_R f^*(x,y)\, dA.$$

When $f(x, y) \geq 0$ on D, the integral is the volume of the three dimensional region between the graph of f and D.

Theorem 5.5: Integrating y-simple regions

If D is a y-simple region, then

$$\iint_D f(x,y)\, dA = \int_a^b \int_{\phi_1(x)}^{\phi_2(x)} f(x,y)\, dy\, dx.$$

and similarly for x-simple regions.

Example 5.7: MT Section 5.3 Problem 4a

Evaluate

$$\int_{-3}^{2} \int_{0}^{y^2} x^2 + y\, dx\, dy.$$

$$\begin{aligned} \int_{-3}^{2} \int_{0}^{y^2} x^2 + y\, dx\, dy &= \int_{-3}^{2} \left[\frac{1}{3}x^3 + xy\right]_{x=0}^{x=y^2} dy \\ &= \int_{-3}^{2} \frac{1}{3}y^6 + y^3\, dy \end{aligned}$$

$$= \left[\frac{1}{21}y^7 + \frac{1}{4}y^4\right]_{y=-3}^{y=2}$$
$$= \frac{2059}{21} - \frac{65}{4}$$

Example 5.8: MT Section 5.3 Problem 4b

Evaluate
$$\int_{-1}^{1}\int_{-2|x|}^{|x|} e^{x+y}\,dy\,dx.$$

As with most absolute value bounds, it's best to split this into two integrals, based on the sign of x.

$$\begin{aligned}
\int_{-1}^{1}\int_{-2|x|}^{|x|} e^{x+y}\,dy\,dx &= \int_{-1}^{0}\int_{2x}^{-x} e^{x+y}\,dy\,dx + \int_{0}^{1}\int_{-2x}^{x} e^{x+y}\,dy\,dx \\
&= \int_{-1}^{0} \left[e^{x+y}\right]_{y=2x}^{y=-x}\,dy\,dx + \int_{0}^{1} \left[e^{x+y}\right]_{y=-2x}^{y=x}\,dx \\
&= \int_{-1}^{0} 1 - e^{3x}\,dx + \int_{0}^{1} e^{2x} - e^{-x}\,dx \\
&= \left[x - \frac{1}{3}e^{3x}\right]_{x=-1}^{x=0} + \left[\frac{1}{2}e^{2x} + e^{-x}\right]_{x=0}^{x=1} \\
&= (0-(-1)) - 1/3(1-e^{-3}) + 1/2(e^2-1) + e^{-1} - 1 \\
&= \frac{e^2}{2} + \frac{1}{e} + \frac{1}{3e^3} - \frac{5}{6}.
\end{aligned}$$

Example 5.9: MT Section 5.3 Problem 4c

Evaluate
$$\int_{0}^{1}\int_{0}^{\sqrt{1-x^2}} dy\,dx$$

Note that the bounds on the integral with respect to y imply that
$$0 \le y \le \sqrt{1-x^2}.$$

Squaring the right inequality gives $y^2 \le 1 - x^2$ which then becomes
$$x^2 + y^2 \le 1.$$

However, $y \ge 0$ because of the first bound. Further, $x \ge 0$ because of the integral bounds with respect to x. So this is a circle, where x and y are both restricted be in the first quadrant, which is a quarter circle with area
$$\frac{\pi}{4}.$$

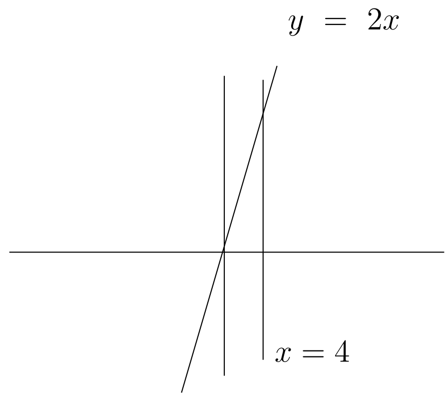
5.4 "Changing the Order of Integration" (MT 5.4)

It is often easier to evaluate some integrals by integrating first by y and then by x and vice-versa. For rectangles, we can completely swap the bounds without issue, but for more complex regions, more care must be taken.

Example 5.10: MT Section 5.4 Problem 1(a-c)

Change the order of integration of the following integrals:
(a) $\int_0^8 \int_{1/2y}^4 dx\, dy$
(b) $\int_0^9 \int_0^{\sqrt{y}} dx\, dy$
(c) $\int_0^4 \int_{-\sqrt{16-y^2}}^{\sqrt{16-y^2}} dx\, dy$

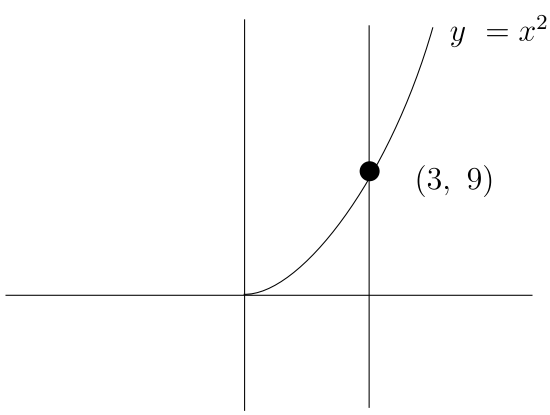
(a) $0 \leq y \leq 8$ and $1/2y \leq x \leq 4$. Note that $(1/2)y \leq x$ is just $y \leq 2x$.



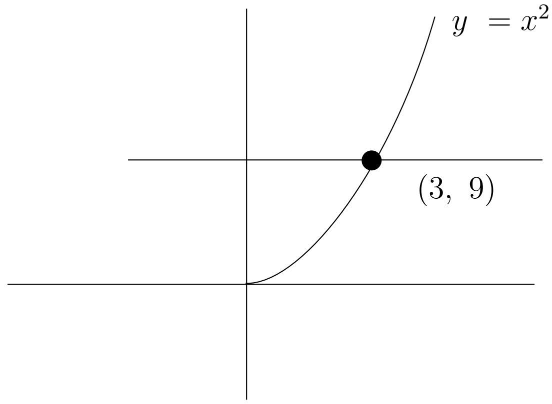
is roughly what the graph looks like. We have to first parameterize the bounds of y first; note that y is always bounded above by $2x$ and below by 0. Then note that x is goes from 0 to 4. This yields

$$\int_0^4 \int_0^{2x} dy\, dx.$$

(b) The bounds note that $0 \leq y \leq 9$ and $0 \leq x \leq \sqrt{y}$. Similar to before, we note that $x \leq \sqrt{y}$ is $y \geq x^2$. Graphing, we have that



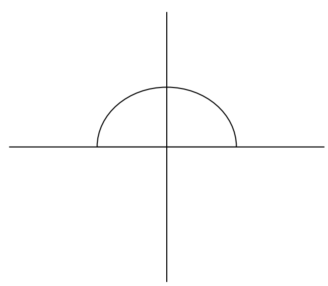
however, note that because $x \leq \sqrt{y}$, the region we are interested in is *not* below $y = x^2$, but actually to the *left* of it. The boundary, then, is actually by the top, when $y = 9$. More accurately, our region is



Then, if we take an arbitrary point, we note that $x^2 \leq y \leq 9$, and that x ranges from 0 to 3. This gives

$$\int_0^3 \int_{x^2}^9 dy\, dx.$$

(c) The bounds on the integral with respect to x gives the inequality $-\sqrt{16-y^2} \leq x \leq \sqrt{16-y^2}$. Squaring gives $x^2 \leq 16 - y^2$, or $x^2 + y^2 \leq 16$, which is a circle of radius 4. Since we have the negative bound, x may be negative. However, y ranges from 0 to 4, so we only have the portion of the circle that has positive y-value. Graphing, this looks like



x ranges from -4 to 4 here. y is bounded below by 0; by the same mechanism that bounded x in the original integral, y is bounded by $\sqrt{16-x^2}$. This gives

$$\int_{-4}^4 \int_0^{\sqrt{16-x^2}} dy\, dx.$$

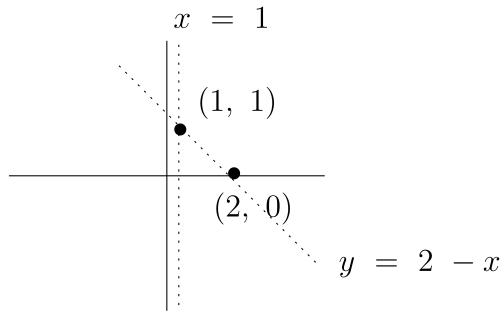
Example 5.11: MT Section 5.4 Problem 3c

Change the order of integration, sketch the corresponding region, and evaluate the integral

$$\int_0^1 \int_1^{2-y} (x+y)^2 \, dx \, dy$$

both ways.

The region looks like this:



Integrating normally,

$$\begin{aligned}\int_0^1 \int_1^{2-y} (x+y)^2 \, dx \, dy &= \frac{1}{3} \int_0^1 \left[(x+y)^3\right]_{x=1}^{x=2-y} \, dy \\ &= \frac{1}{3} \int_0^1 8 - (1+y)^3 \, dy \\ &= \frac{1}{3} \left[8y - \frac{1}{4}(1+y)^4\right]_{y=0}^{y=1} \\ &= \frac{8 - 4 + \frac{1}{4}}{3} \\ &= \frac{17}{12}.\end{aligned}$$

Changing the order of integration,

$$\begin{aligned}\int_1^2 \int_0^{2-x} (x+y)^2 \, dy \, dx &= \frac{1}{3} \int_1^2 \left[(x+y)^3\right]_{y=0}^{y=2-x} \, dx \\ &= \frac{1}{3} \int_1^2 8 - x^3 \, dx \\ &= \frac{1}{3} \left[8x - \frac{1}{4}x^4\right]_{x=1}^{x=2} \\ &= \frac{16 - 8 - 4 + 1/4}{3} \\ &= \frac{17}{12}.\end{aligned}$$

5.4.1 Mean-Value Stuff

Theorem 5.6: Mean-Value Inequality

Suppose there exists numbers m and M such that $m \leq f(x,y) \leq M$ for all $(x,y) \in D$. Then, letting $A(D)$ be the area of the region, we have that

$$mA(D) \leq \int_D f(x,y)\,dA \leq MA(D).$$

Theorem 5.7: Mean-Value Theorem

Suppose $f : D \mapsto \mathbb{R}$ is continuous and D is an elementary region. Then for some point $(x_0, y_0) \in D$ we have

$$\iint_D f(x,y)\,dA = f(x_0,y_0)A(D).$$

5.5 "The Triple Integral" (MT 5.5)

There is not much to add, since the ideas of double integrals transfer over, more or less entirely.

Example 5.12: MT Section 5.5 Problem 6

Perform

$$\iiint_B ze^{x+y}\,dx\,dy\,dz$$

over the given box $B = [0,1] \times [0,1] \times [0,1]$

$$\begin{aligned}
\iiint_B ze^{x+y}\,dx\,dy\,dz &= \int_0^1 \int_0^1 \int_0^1 ze^{x+y}\,dx\,dy\,dz \\
&= \int_0^1 z \int_0^1 \left[e^{x+y}\right]_{x=0}^{x=1}\,dy\,dz \\
&= \int_0^1 z \int_0^1 e^{1+y} - e^y\,dy\,dz \\
&= \int_0^1 z\left[e^{y+1} - e^y\right]_{y=0}^{y=1}\,dz \\
&= \int_0^1 z(e^2 - e^1 - e^1 + e^0)\,dz \\
&= \int_0^1 z(e^2 - 2e + 1)\,dz \\
&= (e^2 - 2e + 1)\int_0^1 z\,dz
\end{aligned}$$

$$= \frac{e^2 - 2e + 1}{2} \left[z^2\right]_{z=0}^{z=1}$$
$$= \frac{e^2 - 2e + 1}{2}.$$

Example 5.13: MT Section 5.5 Problem 8

Describe the region cut out of the ball $x^2 + y^2 + z^2 \leq 4$ by the elliptic cylinder $2x^2 + z^2 = 1$ (that is, the region inside the cylinder and the ball) as an elementary region.

The elliptic cylinder opens on the y-axis, so the only conditions that bound the y-value are induced by the ball, so we'll start from there.

$$-\sqrt{4 - x^2 - z^2} \leq y \leq \sqrt{4 - x^2 - z^2}.$$

The ellipse, for any value of x and z, is wholly contained in the cylinder. Thus our constraints for x and z are induced by the ellipse. I'll pick z first.

$$-\sqrt{1 - 2x^2} \leq z \leq \sqrt{1 - 2x^2}$$

Then

$$-\sqrt{\frac{1}{2}} \leq x \leq \sqrt{\frac{1}{2}}.$$

Example 5.14: MT Section 5.5 Problem 13

Find the volume of the solid bounded by $x = y$, $z = 0$, $y = 0$, $x = 1$ and $x + y + z = 0$.

If we solve for z, we get that $z = -x - y$. As x is bounded by y which is bounded by 0 and x is bounded by 1, this gives the chained inequality

$$0 \leq y \leq x \leq 1$$

and so since x and y are greater than 0 in this constraint, z is negative and is then bounded above by 0.

$$\int_0^1 \int_0^x \int_{-x-y}^0 dz\, dy\, dx$$
$$= \int_0^1 \int_0^x [z]_{z=-x-y}^0\, dy\, dx$$
$$= \int_0^1 \int_0^x x + y\, dy\, dx$$
$$= \int_0^1 \left[xy + \frac{1}{2}y^2\right]_{y=0}^{y=x} dx$$

$$\begin{aligned}
&= \int_0^1 x^2 + \frac{1}{2}x^2\,dx \\
&= \frac{3}{2}\int_0^1 x^2\,dx \\
&= \frac{1}{2}\left[x^3\right]_{x=0}^{x=1} \\
&= \frac{1}{2}.
\end{aligned}$$

Example 5.15: MT Section 5.5 Problem 24(b)

Write the integral

$$\int_0^1 \int_0^x \int_0^y f(x,y,z)\,dz\,dy\,dx$$

with the integration order $dx\ dy\ dz$.

Combining inequalities, we have that $0 \le z \le y \le x \le 1$. For each variable, we pick the "closest" bounds that haven't yet been chosen yet:

$$\int_0^1 \int_z^1 \int_y^1 f(x,y,z)\,dx\,dy\,dz$$

5.5.1 One cylindrical coordinate example

Example 5.16: MT Section 5.5 Problem 18

Evaluate the integral

$$\iiint_W z\,dx\,dy\,dz$$

where W is the region bound by planes $x = 0$, $y = 0$, $z = 0$, $z = 1$, and the cylinder $x^2 + y^2 = 1$ with $x, y \ge 0$.

The cylindrical condition makes us work from 0 to $\pi/2$ for θ. r then ranges from 0 to 1, and z ranges from 0 to 1.

$$\begin{aligned}
\int_0^{\pi/2}\int_0^1\int_0^1 zr\,dz\,dr\,d\theta &= \frac{\pi}{4}\int_0^1 r\int_0^1 z\,dz\,dr \\
&= \frac{\pi}{4}\int_0^1 r\left[z^2\right]_{z=0}^{z=1}\,dr \\
&= \frac{\pi}{4}\int_0^1 r\,dr \\
&= \frac{\pi}{8}\left[r^2\right]_{r=0}^{r=1} \\
&= \frac{\pi}{8}.
\end{aligned}$$

6 "The Change of Variables Formula and Applications of Integration" (MT 6)

In single variable calculus, we learned about a powerful technique that undoes the chain-rule, integration by substitution or u-sub. While still usable in multi-variable calculus contexts, sometimes we wish to do multiple substitutions simultaneously, for both x and y, for example. These substitutions can also be more involved than their counterparts in single variable calculus, as we might wish to make a substitution of the form $x = f(u, v)$ and $y = g(u, v)$, which intertwines the variables we integrate over.

6.1 "The Geometry of Maps from $\mathbb{R}^2$ to $\mathbb{R}^2$ (MT 6.1)

This section lays out the definitions and mechanics for later sections. Though there are problems here, I won't show them, as they're more tedious than anything.

Definition 6.1: Image of a map

Let $f : \mathbb{R}^n \mapsto \mathbb{R}^m$ and let $A \subset \mathbb{R}^n$. The image of A under f, denoted

$$f(A) = \{f(a) \mid a \in A.\}$$

is the set of all points in A evaluated by f.

Let D^* be a subset of $\mathbb{R}^2$, a region were are interested in integrating over. We are interested in a continuously differentiable map $T : D^* \mapsto \mathbb{R}^2$, so T takes points in D^* to $\mathbb{R}^2$.

It is the transformation T that generally underlay the integration techniques of polar or spherical or cylindrical coordinates. For instance the circle $x^2 + y^2 \leq 1$ can be understood as the region $0 \leq r \leq 1$ and $0 \leq \theta \leq 2\pi$. In the θr plane, this is a rectangle, which we may call D. But under the transformation $T(r, \theta) = (r\cos(\theta), r\sin(\theta))$, the rectangle is transformed into a circle in the xy-plane, so $T(D)$ is our original circle.

Theorem 6.1: Matrix Transformations are Linear

Let A be an $n \times n$ matrix with $\det A \neq 0$ and let T be the linear mapping of $\mathbb{R}^n$ to $\mathbb{R}^n$ given by matrix multiplication

$$T(x) = Ax.$$

Then T transforms parallelograms to parallelograms and vertices into vertices. Moreover, if $T(D^*)$ is a parallelogram, so is D^*.

Definition 6.2: One-to-one Maps

A mapping T is one-to-one (or injective) on D^* if for (u, v) and $(u', v') \in D*$,

$$T(u, v) = T(u', v') \implies (u, v) = (u', v').$$

Definition 6.3: Onto Maps

A mapping T is onto D (or surjective) if for every point $(x, y) \in D$ there exists a point (u, v) in the domain of T such that $T(u, v) = (x, y)$.

6.2 "The Change of Variables Theorem" (MT 6.2)

We can now finally investigate the mathematics behind variable transformations. Recall that our target is that given two regions D and D^* in $\mathbb{R}^2$, we equip a differentiable map T on D^* such that $T(D^*) = D$. This way for any real valued function $f : D \mapsto \mathbb{R}$, we would like to express the integral

$$\iint_D f(x, y)\, dA$$

in terms of $f \circ T$. T is our mechanism for transforming variables, and our goal is that expressing the integral in such a way gives easier bounds to work with.

For simplicity, let

$$T(u, v) = (x(u, v), y(u, v))$$

where $(u, v) \in D^*$. Unfortunately, we can't just jam T into f and integrate $f(T(u, v))$. This gets into the idea of how different paramereterizations have different "rates." Here, it is distortion of the area we must regularize.

6.2.1 Jacobian Determinants

The Jacobian determinant describes, for a given variable transformation, how much the area distorts.

Definition 6.4: 2-Dimensional Jacobian Determinant

Let $T : D^* \subset \mathbb{R}^2 \mapsto \mathbb{R}^2$ be a C^1 transformation given by $x = x(u, v)$ and $y = y(u, v)$. The Jacobian determinant of T is the determinant of the derivative matrix $\mathbf{D}T(u, v)$ of T:

$$\frac{\partial(x, y)}{\partial(u, v)} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix}.$$

Higher dimensional analogues follow the same idea.

6.2.2 Change of Variables Formula

Theorem 6.2: Change of Variables: Double Integrals

Let D and D^* be elementary regions in the plane and let $T : D^* \mapsto D$ be of class C^1. Let T be one-to-one on D^* and $D = T(D^*)$. Then for any integrable function $f : D \mapsto \mathbb{R}$, we have

$$\iint_D f(x,y)\,dx\,dy = \iint_{D^*} f(x(u,v), y(u,v)) \left|\frac{\partial(x,y)}{\partial(u,v)}\right| du\,dv.$$

Example 6.1: MT Section 6.2 Problem 4

Let D be the region $0 \le y \le x$ and $0 \le x \le 1$. Evaluate

$$\iint_D x + y\,dx\,dy$$

by making the change of variables $x = u + v$, $y = u - v$. Check your answer by evaluating the integral directly by using an iterated integral.

First, normally,

$$\begin{aligned}
\int_0^1 \int_0^x x + y\,dy\,dx &= \int_0^1 \left[xy + \frac{1}{2}y^2\right]_{y=0}^{y=x} dx \\
&= \int_0^1 x^2 + \frac{1}{2}x^2\,dx \\
&= \frac{3}{2}\int_0^1 x^2\,dx \\
&= \frac{1}{2}\left[x^3\right]_{x=0}^{x=1} \\
&= \frac{1}{2}
\end{aligned}$$

Now we consider the change of variables. We have the transformation $T(u,v) = (u+v, u-v)$, so it remains to compute the Jacobian determinant for T and actually identify the region D^* such that $T(D^*) = D$.

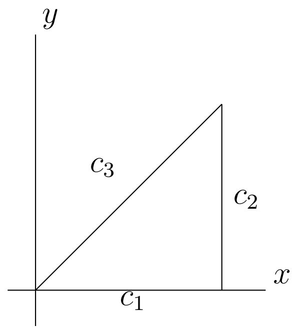
Since $x = u + v$ and $y = u - v$, we have that

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

which gives us a way to transform our information in the xy-plane into the uv-plane.

Normally, you would just cite that since T is a linear transformations, triangles are still triangles, and we can just identify the vertices in the uv-plane, and then construct the region from there, I'll show you a way that generalizes to stranger boundaries.

We consider the region D, and denote the paths that bound the region:



We can then parameterize the bounding paths as

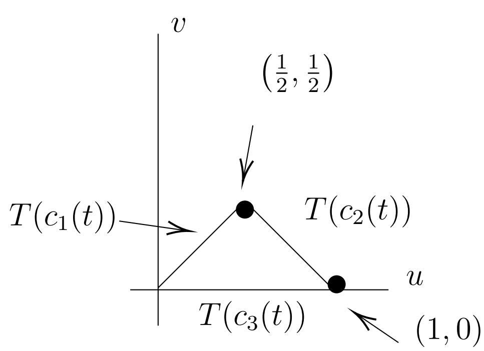
$$c_1(t) = (t, 0)$$
$$c_2(t) = (1, t)$$
$$c_3(t) = (t, t)$$

where all range from $0 \leq t \leq 1$.

We now consider the image of these paths under T:

$$c_1(t) = \left(\frac{t}{2}, \frac{t}{2}\right)$$
$$c_2(t) = \left(\frac{t+1}{2}, \frac{1-t}{2}\right)$$
$$c_3(t) = (t, 0)$$

Plotting this in the uv-plane yields



It seems promising to then just integrate over v then u over two intervals, but we can also do it in one. The first interval is $v = u$, the second interval is $v = 1 - u$. In terms of u, these are then $u = v$ and $u = 1 - v$. Integrating over u then v, we consider

$$\int_0^1 \int_v^{v-2} (u + v + u - v) \left| \frac{\partial(x, y)}{\partial(u, v)} \right| dv$$

Computing the Jacobian,

$$\left| \frac{\partial(x, y)}{\partial(u, v)} \right| = \left| \begin{vmatrix} 1 & 1 \\ 1 & -1 \end{vmatrix} \right|$$

$$\begin{aligned} &= |-1-1| \\ &= 2. \end{aligned}$$

Integrating,

$$\begin{aligned} \int_0^{1/2}\int_v^{v-2} 2u\cdot 2\,du\,dv &= 4\int_0^1\int_v^{v-1} u\,du\,dv \\ &= 2\int_0^{1/2}\left[u^2\right]_{u=v}^{u=1-v}\,dv \\ &= 2\int_0^{1/2} v^2-2v+1-v^2\,dv \\ &= 2\int_0^{1/2} 1-2v\,dv \\ &= 2\left[v-v^2\right]_{v=0}^{v=1/2} \\ &= 2(1/2-1/4) \\ &= \frac{1}{2}. \end{aligned}$$

Example 6.2: MT Section 6.2 Problem 7

Evaluate

$$\iint_D \frac{dx\,dy}{\sqrt{1+x+2y}}$$

where $D=[0,1]\times[0,1]$ by setting $T(u,v)=(u,v/2)$ and evaluating an integral over D^*, where $T(D^*)=D$.

We are presented with the substitution $x(u,v)=u$ and $y(u,v)=v/2$. Finding the Jacobian determinant right away, we have

$$\begin{aligned} \left|\frac{\partial(x,y)}{\partial(u,v)}\right| &= \left|\begin{vmatrix} 1 & 0 \\ 0 & \frac{1}{2}\end{vmatrix}\right| \\ &= \frac{1}{2}. \end{aligned}$$

We now look to find D^*. Note that since T suggests we replace y with $v/2$, and y ranges from $[0,1]$, then $v=2y$ and so v ranges from $[0,2]$. Then D^* is $[0,1]\times[0,2]$ on the uv-plane.

Integrating,

$$\begin{aligned} \int_0^1\int_0^2 (1+u+v)^{-1/2}\cdot 1/2\,du\,dv &= \frac{1}{2}\int_0^1\int_0^2 (1+u+v)^{-1/2}\,du\,dv \\ &= \int_0^1 \left[(1+u+v)^{1/2}\right]_{u=0}^{u=2}\,dv \end{aligned}$$

$$
\begin{aligned}
&= \int_0^1 (v+3)^{1/2} - (v+1)^{1/2}\,dv \\
&= \frac{2}{3}\left[(v+3)^{3/2} - (v+1)^{3/2}\right]_{v=0}^{v=1} \\
&= \frac{2}{3}(4^{3/2} - 3^{3/2} - 2^{3/2} + 1^{3/2}) \\
&= \frac{2}{3}(8 - 3\sqrt{3} - 2\sqrt{2} + 1) \\
&= \frac{2}{3}(9 - 3\sqrt{3} - 2\sqrt{2}) \\
&= 6 - 2\sqrt{3} - \frac{4}{3}\sqrt{2}
\end{aligned}
$$

6.2.3 Polar, cylindrical, and spherical integration examples

The Jacobian determinant and change of variables theorem in general precipitates the area differentials for all of the variable substitutions we are familiar with:

- Polar – r
- Cylindrical – r
- Spherical – $\rho^2 \sin\phi$

Example 6.3: MT Section 6.2 Problem 3

Let D be the unit disk: $x^2 + y^2 \leq 1$. Evaluate

$$\iint_D \exp(x^2 + y^2)\,dx\,dy.$$

The unit disk, in polar coordinates, corresponds to the constraints that $0 \leq \theta \leq 2\pi$ and $0 \leq r \leq 1$. Further, since we use the transformation $T(r,\theta) = (r\cos(\theta), r\sin(\theta))$, our integrand becomes

$$
\begin{aligned}
\exp(x^2 + y^2) &= \exp((r\cos(\theta))^2 + (r\sin(\theta))^2) \\
&= \exp(r^2\cos^2(\theta) + r^2\sin^2(\theta)) \\
&= \exp(r^2).
\end{aligned}
$$

Remembering to include the polar area differential, we integrate:

$$
\begin{aligned}
\int_0^{2\pi}\int_0^1 \exp(r^2) r\,dr\,d\theta &= 2\pi\int_0^1 r\exp(r^2)dr \\
&= \pi\left[\exp(r^2)\right]_{r=0}^{r=1} \\
&= \pi(e-1).
\end{aligned}
$$

Example 6.4: MT Section 6.2 Problem 25

Evaluate

$$\iiint_W \frac{dx\,dy\,dz}{(x^2+y^2+z^2)^{3/2}}$$

where W is the solid bounded by the two spheres $x^2+y^2+z^2=a^2$ and $x^2+y^2+z^2=b^2$ where $0<b<a$.

These are spheres with radius a and b, which in spherical coordinates are $\rho=a$ and $\rho=b$. $\rho^2=x^2+y^2+z^2$. Then, integrating, we have

$$\begin{aligned}
\iiint_W \frac{dx\,dy\,dz}{(x^2+y^2+z^2)^{3/2}} &= \int_0^{2\pi}\int_0^{\pi}\int_b^{a}(\rho^2)^{-3/2}\cdot\rho^2\sin\phi\,d\rho\,d\phi,d\theta \\
&= 2\pi\int_0^{\pi}\int_b^{a}\frac{1}{\rho}\sin\phi\,d\rho\,d\phi \\
&= 2\pi\int_0^{\pi}\sin\phi\,[\ln|\rho|]_{\rho=b}^{\rho=a}\,d\phi \\
&= 2\pi\ln\left(\frac{a}{b}\right)\int_0^{\pi}\sin\phi\,d\phi \\
&= 2\pi\ln\left(\frac{a}{b}\right)[-\cos\phi]_{\phi=0}^{\phi=\pi} \\
&= 4\pi\ln\left(\frac{a}{b}\right).
\end{aligned}$$

Example 6.5: MT Section 6.2 Problem 26

Use spherical coordinates to evaluate

$$\int_0^3\int_0^{\sqrt{9-x^2}}\int_0^{\sqrt{9-x^2-y^2}}\frac{\sqrt{x^2+y^2+z^2}}{1+(x^2+y^2+z^2)^2}\,dz\,dy\,dx$$

We again inspect the bounds. x ranges from 0 to 3. y ranges from 0 to $\sqrt{9-x^2}$, which when squared gives $y^2\le 9-x^2$ and that $x^2+y^2\le 9$, but y is still greater than 0. Similarly, we have that $x^2+y^2+z^2\le 9$ with $z>0$. $x,y,z>0$ describes the first octant, and hence $0<\theta<\pi/2$ and $0<\phi<\pi/2$. This is also a sphere of radius 3. Integrating,

$$\begin{aligned}
I &= \int_0^3\int_0^{\sqrt{9-x^2}}\int_0^{\sqrt{9-x^2-y^2}}\frac{\sqrt{x^2+y^2+z^2}}{1+(x^2+y^2+z^2)^2}\,dz\,dy\,dx \\
&= \int_0^{\pi/2}\int_0^{\pi/2}\int_0^3\frac{\sqrt{\rho^2}}{1+(\rho^2)^2}\cdot\rho^2\sin\phi\,d\rho\,d\phi\,d\theta \\
&= \frac{\pi}{2}\int_0^{\pi/2}\sin\phi\int_0^3\frac{\rho^3}{1+\rho^4}d\,\rho\,d\phi \\
&= \frac{\pi}{2}\int_0^{\pi/2}\sin\phi\int_1^{82}\frac{\rho^3}{u}\cdot\frac{1}{4\rho^3}du\,d\phi
\end{aligned}$$

$$= \frac{\pi}{8}\int_0^{\pi/2} \sin\phi \int_1^{82} \frac{1}{u} du\, d\phi$$
$$= \frac{\pi}{8}\ln(82)\int_0^{\pi/2} \sin(\phi)$$
$$= \frac{\pi}{8}\ln(82).$$

Example 6.6: MT Section 6.2 Problem 30

Evaluate the following by using cylindrical coordinates.

(a) $\iiint_B z\, dx\, dy\, dz$ where B is the region within the cylinder $x^2+y^2=1$ above the xy-plane and below the cone $z=\sqrt{x^2+y^2}$

(b) $\iiint_W (x^2+y^2+z^2)^{-1/2}\, dx\, dy\, dz$ where W is the region determined by the conditions $1/2 \le z \le 1$ and $x^2+y^2+z^2 \le 1$

(a) Containment by $x^2+y^2=1$ implies $0 \le r \le 1$. Since $r^2 = x^2+y^2$, the cone can be described as $z=\sqrt{r^2}=r$. As B is above the xy-plane, this gives the lower bound of z as 0. Integrating,

$$\int_0^{2\pi}\int_0^1\int_0^r z\cdot r\, dr\, dz\, d\theta = 2\pi\int_0^1 r\int_0^r z\, dz\, dr$$
$$= \pi\int_0^1 r\left[z^2\right]_{z=0}^{z=r}$$
$$= \pi\int_0^1 r(r^2-0)$$
$$= \pi\int_0^1 r^3$$
$$= \frac{\pi}{4}\left[r^4\right]_{r=0}^{r=1}$$
$$= \frac{\pi}{4}.$$

(b) Beyond the fact that the problem mandates we solve this by cylindrical coordinates, cylindrical coordinates are preferred over spherical coordinates for this problem as expressing the $1/2 \le z \le 1$ in spherical coordinates is likely quite messy. Instead, since $r^2=x^2+y^2$, the condition that $x^2+y^2+z^2 \le 1$ becomes $r^2+z^2 \le 1$. Since we have z's boundary, we solve for r, and this gives that $-\sqrt{1-z^2} \le r \le \sqrt{1-z^2}$. However, r is always positive by definition, so the lower bound is actually 0. Integrating,

$$\iiint_W \frac{1}{\sqrt{x^2+y^2+z^2}} = \int_0^{2\pi}\int_{1/2}^1\int_0^{\sqrt{1-z^2}} \frac{1}{\sqrt{r^2+z^2}}\cdot r\, dr\, dz\, d\theta$$
$$= 2\pi\int_{1/2}^1\int_0^{\sqrt{1-z^2}} \frac{1}{\sqrt{r^2+z^2}}\cdot r\, dr\, dz$$

$$
\begin{aligned}
&= 2\pi \int_{1/2}^{1} \left[\sqrt{(r^2+z^2)}\right]_{r=0}^{r=\sqrt{1-z^2}} dz \\
&= 2\pi \int_{1/2}^{1} 1 - z\, dz \\
&= 2\pi \left[z - \frac{1}{2}z^2\right]_{z=1/2}^{z=1} \\
&= 2\pi(1 - 1/2 - 1/2 + 1/8) \\
&= \frac{\pi}{4}.
\end{aligned}
$$

Example 6.7: MT Section 6.2 Problem 33

Let E be the ellipsoid

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} \leq 1,$$

where a, b and c are positive.

(a) Find the volume of E.

(b) Evaluate $\iiint_E (x^2/a^2) + (y^2/b^2) + (z^2/c^2)\, dx\, dy\, dz$

(Hint: Change variables and then use spherical coordinates).

To avoid issues with fractions, we propose the change of variables given by

$$
\begin{aligned}
x &= au \\
y &= bv \\
z &= cw
\end{aligned}
$$

We argue that

$$D^* : u^2 + v^2 + w^2 \leq 1$$

with the transformation $T(u, v, w) = (au, bv, cw)$ is a valid change of variables, and $T(D^*) = E$. Notice that for any (u, v, w), we have that when using the change of variables under T, the ellipsoid becomes

$$
\begin{aligned}
\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} \leq 1 &\overset{T}{\Rightarrow} \frac{a^2u^2}{a^2} + \frac{b^2v^2}{v^2} + \frac{c^2w^2}{c^2} \leq 1 \\
&\implies u^2 + v^2 + w^2 \leq 1
\end{aligned}
$$

which is precisely D^*.

Computing the Jacobian gives abc.

(a) It suffices to compute $\iiint_{D^*} abc\, dV$. Since that's just the volume of the sphere in the uvw-space, this gives

$$\frac{4abc}{3}\pi$$

(b) This becomes

$$abc \iiint_{D^*} u^2 + v^2 + w^2 \, dV$$

We then use spherical coordinates defined over the uvw-space. Integrating,

$$\begin{aligned}
abc \iiint_{D^*} u^2 + v^2 + w^2 \, dV &= abc \int_0^{2\pi} \int_0^{\pi} \int_0^1 \rho^2 \cdot \rho^2 \sin\phi \, d\rho \, d\phi \, d\theta \\
&= 2abc\pi \int_0^{\pi} \sin\phi \int_0^1 \rho^4 d\rho d\phi \\
&= 2abc\pi \int_0^{\pi} \sin\phi \left[\frac{1}{5}\rho^5\right]_{\rho=0}^{\rho=1} d\phi \\
&= \frac{2abc\pi}{5} \int_0^{\pi} \sin\phi \, d\phi \\
&= \frac{2abc\pi}{5} \left[-\cos\phi\right]_{\phi=0}^{\phi=\pi} \\
&= \frac{4abc\pi}{5}.
\end{aligned}$$

7 "Integrals Over Paths and Surfaces" (MT 7)

When we first investigate definite integrals in single variable calculus of the form

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

we can visualize this, geometrically, as the total area under the curve from the interval $x \in [a, b]$. We also discussed integration in $\mathbb{R}^2$ similarly as the signed volume under the graph of a function and the xy-plane, over some region in the xy-plane D.

We discuss two generalizations:

- One where we consider the integral over a path
- One where we consider the integral over a surface

7.1 "The Path Integral" (MT 7.1)

Now we discuss a different kind of generalization of the integral over an interval into $\mathbb{R}^2$. Instead of describing the volume under the graph of a function by considering the volume over some region D, we describe the area under the graph of a function over a path.

This naturally only works for when we can associate a "height", which means we will only discuss functions of the form

$$f : \mathbb{R}^n \mapsto \mathbb{R}.$$

We will discuss the case where f takes a vector codomain when we discuss line integrals.

Definition 7.1: Path Integral

We define the integral of a scalar function $f : \mathbb{R}^n \mapsto \mathbb{R}$ along a path $c : I = [a, b] \mapsto \mathbb{R}^n$ for a C^1 path c and when the composition $f \circ c$ is continuous on I as

$$\int_c f\,ds = \int_a^b f(c(t))\|c'(t)\|\,ds$$

In the case where $f(c(t))$ is piece-wise continuous, we sum over integrals of the continuous intervals.

Note that in the case where $f = 1$, this is the arc length integral.

Example 7.1: MT Section 7.1 Problem 10

Evaluate the following path integrals $\int_c f(x, y, z)\,ds$ where
(a) $f(x, y, z) = x + y + z$ and $c : t \mapsto (\sin t, \cos t, t)$, $t \in [0, 2\pi]$
(b) $f(x, y, z) = \cos z$, c as in part (a).

Since c is the same for both parts and we calculate $\|c'(t)\|$ for both, we'll calculate that now.

$$c'(t) = (\cos(t), -\sin(t), 1)$$

Then, the norm is

$$\begin{aligned}\|c'(t)\| &= \sqrt{\cos^2 t + \sin^2 t + 1} \\ &= \sqrt{1+1} \\ &= \sqrt{2}.\end{aligned}$$

(a) $f(c(t)) = \sin t + \cos t + t$. Integrating,

$$\begin{aligned}\int_c f\, ds &= \int_0^{2\pi} (\sin t + \cos t + t) \cdot \sqrt{2}\, dt \\ &= \sqrt{2} \int_0^{2\pi} \sin t + \cos t + t\, dt \\ &= \sqrt{2} \left[-\cos t + \sin t + \frac{1}{2} t^2 \right]_{t=0}^{t=2\pi} \\ &= 2\sqrt{2}\pi^2.\end{aligned}$$

(b) $f(c(t)) = \cos t$. Integrating,

$$\begin{aligned}\int_c f\, ds &= \int_0^{2\pi} (\cos t) \cdot \sqrt{2}\, dt \\ &= \sqrt{2} \int_0^{2\pi} \cos t\, dt \\ &= \sqrt{2}\, [\sin t]_{t=0}^{t=2} \\ &= 0.\end{aligned}$$

Example 7.2: MT Section 7.1 Problem 11

Evaluate the following path integrals $\int_c f(x,y,z)\, ds$ where
(a) $f(x,y,z) = \exp \sqrt{z}$ and $c : t \mapsto (1, 2, t^2)$, $t \in [0,1]$
(b) $f(x,y,z) = yz$ and $c : t \mapsto (t, 3t, 2t)$, $t \in [1,3]$

(a) $c'(t) = (0, 0, 2t)$ so $\|c'(t)\| = \sqrt{4t^2} = 2|t|$ which is $2t$ since we only consider $t > 0$.

$f(c(t)) = \exp(\sqrt{t^2}) = \exp|t| = \exp t$ since we only consider $t > 0$.

Integrating,

$$\begin{aligned}\int_0^1 \exp t \cdot 2t\, dt &= 2 \int_0^1 t \exp t\, dt \\ &= 2 \left[te^t - e^t \right]_{t=0}^{t=1} \\ &= 2(1e^1 - e^1 - 0e^0 + e^0) \\ &= 2(e - e - 0 + 1) \\ &= 2.\end{aligned}$$

(b) $c'(t) = (1, 3, 2)$, so $\|c'(t)\| = \sqrt{1+9+4} = \sqrt{14}$.

$f(c(t)) = 3t \cdot 2t = 6t^2$.

Integrating,

$$\begin{aligned}\int_1^3 6t^2 \cdot \sqrt{14}\, dt &= 6\sqrt{14} \int_1^3 t^2\, dt \\ &= 2\sqrt{14} \left[t^3\right]_{t=1}^{t=3} \\ &= 52\sqrt{14}.\end{aligned}$$

Example 7.3: MT Section 7.1 Problem 24

Compute the path integral of $f(x, y) = y^2$ over the graph $y = e^x$, $0 \le x \le 1$.

We're not given a path directly here, but we just derive one, letting t take the value of x and using the points on the graph $f(x) = e^x$ for the path. This gives

$$c(t) = (t, e^t), \quad 0 \le t \le 1.$$

Then $c'(t) = (1, e^t)$, and then $\|c'(t)\| = \sqrt{1 + e^{2t}}$.

$f(c(t)) = e^{2t}$. Integrating,

$$\begin{aligned}\int_0^1 e^{2t} \cdot \sqrt{1+e^{2t}}\, dt &= \frac{1}{2} \int_1^{1+e^2} \sqrt{u}\, du \\ &= \frac{1}{3} \left[u^{3/2}\right]_{u=2}^{u=1+e^2} \\ &= \frac{1}{3} \left((1+e^2)^{3/2} - 2^{3/2}\right).\end{aligned}$$

7.1.1 Average Value

Theorem 7.1: The Average Value of a Function on a Path

Let $L(c)$ denote the arc length of a path as we have defined before,

$$L(c) = \int_c \|c'(t)\|\, dt.$$

Then the average value of f along c is defined as

$$\frac{\int_c f(x, y, z)\, ds}{L(c)}.$$

Example 7.4: MT Section 7.1 Problem 17

Find the average y coordinate of the points on the semicircle parameterized by $c : [0, \pi] \mapsto \mathbb{R}^3$, $\theta \mapsto (0, a \sin\theta, a \cos\theta)$; $a > 0$.

It may not be immediately clear, but the path is

$$c(\theta) = (0, a\sin\theta, a\cos\theta), \quad 0 \leq \theta \leq \pi.$$

We're not provided a function, but since we want the average y coordinate, I claim that the scalar function of interest is

$$f(x, y, z) = y.$$

Let's first find the arc length of the path. $c'(t) = (0, a\cos\theta, -a\sin\theta)$, then

$$\begin{aligned}\|c'(t)\| &= \sqrt{0 + a^2\cos^2\theta + a^2\sin^2\theta} \\ &= \sqrt{a^2} \\ &= |a| \\ &= a\end{aligned}$$

since $a > 0$.

Integrating,

$$\begin{aligned}L(c) &= \int_0^\pi a\,d\theta \\ &= a\int_0^\pi 1\,d\theta \\ &= a\pi.\end{aligned}$$

Alternatively, recognize that this is a circle with radius a, so the circumference is $2a\pi$, so the circumference of the semicircle is half that, $a\pi$.

$f(c(t)) = a\sin\theta$. Integrating,

$$\begin{aligned}\frac{1}{a\pi}\int_0^\pi a\sin\theta\cdot a\,d\theta &= \frac{a}{\pi}\int_0^\pi \sin\theta\,d\theta \\ &= \frac{a}{\pi}\left[-\cos\theta\right]_{\theta=0}^{\theta=\pi} \\ &= \frac{2a}{\pi}.\end{aligned}$$

7.2 "Line Integrals" (MT 7.2)

We now discuss a generalization of path integrals for functions of the form

$$f : \mathbb{R}^n \mapsto \mathbb{R}^m.$$

As with much of vector calculus with vector functions, line integrals over vector functions have a particular connection to describing situations in physics. Suppose we have some vector field

$$F : \mathbb{R}^n \mapsto \mathbb{R}^n.$$

Then a particle at some position X will experience a force of $F(X)$. We are often interested in the total amount of work induced on a particle as it traces across a path c. Let's first consider a simple case, where F is a constant vector force, and c is a straight line, which we can then describe as a vector d from the beginning to the end. Then the total force induced, the work, is

$$F \cdot d$$

or the product of the force's magnitude and the displacement in the direction of force. The dot product is clearer when we consider the dot product cosine formula,

$$\|F\|\|d\|\cos(\theta).$$

Let's see if this makes sense. As d, the displacement vector, increases in length, we expect that the force induced on the particle should increase, since the force has been induced on a larger displacement. Naturally, the total force induced should increase if the constant force increases, as well. Finally, if the force is in the direction of the displacement, wherein $\theta = 0$ and hence $\cos\theta = 1$, the work induced on the particle should be positive. The idea is that if the force is in the same direction of the trajectory, the force drives the movement, and hence a positive amount of work is applied. On the other hand, if the force is opposite the direction of the displacement, we just flip the sign. And for things in between, only some of the force contributes to the trajectory, so it should scale.

I don't actually know physics, so hopefully the ideas are clear.

So what about if F and c aren't so simple? Well, let's first consider the case that c is represented as a union of straight lines. Then we can just calculate the work on each line, and then sum it up. We can also break up F that way. Doing this to infinity leads to an integral of such work, which motivates the line integral.

Definition 7.2: Line Integral

Let F be a vector field in $\mathbb{R}^n$ that is continuous on the C^1 path $c : [a, b] \mapsto \mathbb{R}^n$. We define the line integral of F along c, denoted by $\int_c F \cdot ds$ as

$$\int_c F \cdot ds = \int_a^b F(c(t)) \cdot c'(t)\, dt.$$

Note the $(\cdot)$ to differentiate a line integral from a path integral, although they're really the same, since a path integral is just a line integral when $n = 1$.

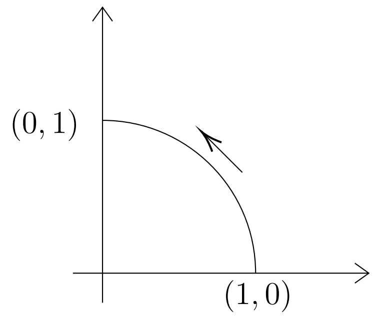
Example 7.5: MT Section 7.2 Problem 1

Evaluate the line integral

$$\int_C F \cdot ds$$

where $F(x, y) = y^2\mathbf{i} - xy\mathbf{j}$ and C is the part of the circle $x^2 + y^2 = 1$ that starts at $(1, 0)$ and ends at $(0, 1)$, oriented counterclockwise.

C looks like this:



We'll use the familiar polar coordinates for our path, that is,

$$c(\theta) = (\cos\theta, \sin\theta).$$

For boundaries, note that since this ends at $(0, 1)$, then $0 \le \theta \le \pi/2$.

$c'(\theta) = (-\sin\theta, \cos\theta)$.

$F(c(\theta)) = (\sin^2\theta, -\cos\theta\sin\theta)$.

Taking the dot product,

$$\begin{aligned} F(c(\theta)) \cdot c'(\theta) &= (\sin^2\theta, -\cos\theta\sin\theta) \cdot (-\sin\theta, \cos\theta) \\ &= -\sin^3\theta - \cos^2\theta\sin\theta \\ &= -\sin\theta(\sin^2\theta + \cos^2\theta) \\ &= -\sin\theta. \end{aligned}$$

Integrating,

$$\begin{aligned} \int_C F \cdot ds &= \int_0^{\pi/2} -\sin\theta \\ &= [\cos t]_{t=0}^{t=\pi/2} \\ &= -1. \end{aligned}$$

Example 7.6: MT Section 7.2 Problem 3

Let

$$F(x, y, z) = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}.$$

Evaluate the integral of F along each of the following paths:

(a) $c(t) = (t, t, t)$, $0 \le t \le 1$.
(b) $c(t) = (\cos t, \sin t, 0)$, $0 \le t \le 2\pi$
(c) $c(t) = (\sin t, 0, \cos t)$, $0 \le t \le 2\pi$
(d) $c(t) = (t^2, 3t, 2t^3)$, $-1 \le t \le 2$

This is almost inherently repetitive, so I'll do a bunch of steps at once:

	$c(t)$	$c'(t)$	$F(c(t))$	$F(c(t))\cdot c'(t)$
(a)	(t,t,t)	$(1,1,1)$	(t,t,t)	$3t$
(b)	$(\cos t,\sin t,0)$	$(-\sin t,\cos t,0)$	$(\cos t,\sin t,0)$	0
(c)	$(\sin t,0,\cos t)$	$(\cos t,0,-\sin t)$	$(\sin t,0,\cos t)$	0
(d)	$(t^2,3t,2t^3)$	$(2t,3,6t^2)$	$(t^2,3t,2t^3)$	$2t^3+9t+12t^5$

So then the line integrals for (b) and (c) are then going to be 0.

For (a) and (d),

$$\begin{aligned}(a):&\int_0^1 3t\,dt\\ &3\int_0^1 t\,dt\\ &\frac{3}{2}\left[t^2\right]_{t=0}^{t=1}\\ &\frac{3}{2}.\\ (b):&\int_{-1}^2 2t^3+9t+12t^5\,dt\\ &\left[\frac{1}{2}t^4+\frac{9}{2}t^2+2t^6\right]_{t=-1}^{t=2}\\ &\left[8-\frac{1}{2}+18-\frac{9}{2}+128-2\right]\\ &147.\end{aligned}$$

Example 7.7: MT Section 7.2 Problem 5

Consider the force field

$$F(x,y,z)=x\mathbf{i}+y\mathbf{j}+z\mathbf{k}.$$

Compute the work done in moving a particle along the parabola $y=x^2$, $z=0$, from $x=-1$ to $x=2$.

This is actually a cylindrical parabola, as z is constantly 0. Then, since $-1\le x\le 2$ and $y=x^2$, we can let x play the role of t which gives the path

$$c(t)=(t,t^2,0),\quad -1\le t\le 2.$$

Then $f(c(t))=(t,t^2,0)$, $c'(t)=(1,2t,0)$ and $f(c(t))\cdot c'(t)=t+2t^3$. Integrating,

$$\begin{aligned}\int_{-1}^2 t+2t^3\,dt&=\left[\frac{1}{2}t^2+\frac{1}{2}t^4\right]_{t=-1}^{t=2}\\ &=(2+8-\frac{1}{2}-\frac{1}{2})\\ &=9.\end{aligned}$$

Example 7.8: MT Section 7.2 Problem 11

Evaluate the integral of the vector field $F(x, y) = x\mathbf{i} + y\mathbf{j}$ around the curve

$$(\cos^3 t, \sin^3 t), \quad 0 \leq t \leq 2\pi.$$

$c'(t) = (-3\sin t\cos^2 t, 3\sin^2 t\cos t)$, $F(c(t)) = (\cos^3 t, \sin^3 t)$, so

$$\begin{aligned} F(c(t)) \cdot c'(t) &= (\cos^3 t, \sin^3 t) \cdot (-3\sin t\cos^2 t, 3\sin^2 t\cos t) \\ &= -3\sin t\cos^5 t + 3\sin^5 t\cos t \\ &= 3\sin t\cos t(-\cos^4 t + \sin^4 t) \\ &= 3\sin t\cos t(\sin^2 t - \cos^2 t)(\sin^2 t + \cos^2 t) \\ &= 3\sin t\cos t(\sin^2 t - \cos^2 t) \\ &= 3\sin^3 t\cos t - 3\sin t\cos^3 t. \end{aligned}$$

Integrating,

$$\begin{aligned} \int_0^{2\pi} 3\sin^3 t\cos t - 3\sin t\cos^3 t &= 3\left[\sin^3 t + \cos^3 t\right]_{t=0}^{t=2\pi} \\ &= 0. \end{aligned}$$

Theorem 7.2: Tangential Line Integral

For a path where $c(t) \neq 0$ and a function F such that F and c meaningfully describe a line integral, we denote the unit tangent vector

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

and we have that

$$\int F \cdot ds = \int_a^b [F(c(t)) \cdot T(t)]\|c'(t)\|\, dt.$$

Another common way of writing line integrals (that I am personally not a fan of) that makes the connection of line integrals to the integral theorems of vector calculus clearer is a differential form, which we do not cover in this text:

$$\int_c F \cdot ds = \int_c \sum_{\forall i} F_i\, dx_i.$$

This can be made clear by identifying $ds = (dx_i)$, but we must be clear that these differentials are with respect to the path c.

7.2.1 A Preview of Conservative Vector Fields

For gradient fields, that is, some vector field F that be

$$F = \nabla f$$

these are also described as conservative vector fields.

For this special case, computing a line integral is significantly simpler, we only need to evaluate f and the endpoints, similar to the fundamental theorem of calculus. This is then aptly named:

Theorem 7.3: Fundamental Theorem of Line Integrals

If $F = \nabla f$ for some scalar function f, over the path $c : [a, b] \mapsto \mathbb{R}^n$

$$\int_c \nabla F \cdot ds = f(c(b)) - f(c(a))$$

This leads to an important consequence in regards to work done on closed intervals

Theorem 7.4: Work done on closed paths

Given a gradient vector field $F = \nabla f$ for some scalar function f, for a closed path $c : [a, b] \mapsto \mathbb{R}^n$, that is, $c(b) = c(a)$, the line integral of F over c is 0.

Example 7.9: MT Section 7.2 Problem 19

Consider the gravitational force field (with $G = m = M = 1$) defined [for $(x, y, z) \neq (0, 0, 0)$] by

$$F(x, y, z) = -\frac{1}{(x^2 + y^2 + z^2)^{3/2}}(x\mathbf{i} + y\mathbf{j} + z\mathbf{k}).$$

Show that the work done by the gravitational force as a particle moves from (x_1, y_1, z_1) to (x_2, y_2, z_2) along any path depends only on the radii $R_1 = \sqrt{x_1^2 + y_1^2 + z_1^2}$ and $R_2 = \sqrt{x_2^2 + y_2^2 + z_2^2}$.

Beyond the suggestive placement, that the work done depends only on the end points should immediately suggest attempting to invoke the fundamental theorem of line integrals, as the desired proposition is the result of the theorem. Thus, it suffices to identify some f such that $F = \nabla f$.

There exists a proper technique for this, but I'll save this for later. Just believe me that

$$f = \frac{1}{\sqrt{x^2 + y^2 + z^2}}$$

suffices.

Then we have that for any path $C : [a, b]$, by the fundamental theorem of line integrals,

$$\int_a^b F(C(t)) \cdot C'(t) = f(c(b)) - f(c(a)).$$

Let C be an arbitrary path such that $C(a) = (x_1, y_1, z_1)$ and $C(b) = (x_2, y_2, z_2)$. Then

$$f(c(b)) = f(x_1, y_1, z_1)$$

$$= \frac{1}{\sqrt{x_1^2 + y_1^2 + z_1^2}}$$
$$= \frac{1}{R_1}$$

and similarly $f(c(b)) = R_2^{-1}$. By the fundamental theorem of line integrals, we have that for any path C, the total work done is

$$\frac{1}{R_2} - \frac{1}{R_1}.$$

7.3 "Parameterized Surfaces" (MT 7.3)

We quickly discuss how to describe surfaces before we can discuss integrating over them. Much of our study of surfaces in $\mathbb{R}^3$ have been graphs of functions $f(x, y)$, but we can't describe all surfaces this way, or even all surfaces that we might want.

As an analogy, much of mathematics leading into calculus implicitly describes most mathematical objects as functions $y = f(x)$, but this is insufficient. We can't describe, say, a circle this way, we need the implicit relationship

$$x^2 + y^2 = 1.$$

Principally, we can't do this because it would fail the vertical line test. Alternatively, we can still describe it explicitly, but we just have to describe x and y with separate relationships with an auxiliary relationship, which is where we get the familiar polar parameterization of

$$c(\theta) = (\cos\theta, \sin\theta).$$

Similarly, we have surfaces that we cannot describe because they too would fail the vertical line test. Perhaps the best example is the sphere. Again, we will use the approach of trying to describe relationships of x, y and z with respect to an auxiliary variables, but we only get a path if we use a single variable; we can have a line arbitrarily as squiggly as possible, but under the standard framework of vector calculus, this can't really be a plane or anything similar to that. So we use two variables.

Definition 7.3: Parameterized Surfaces

A parameterization of a surface is a function $\phi : D \subset \mathbb{R}^2 \mapsto \mathbb{R}^3$ where D is some domain in $\mathbb{R}^2$. The surface S corresponding to the function ϕ is its image $S = \phi(D)$. We can write

$$\Phi(u, v) = (x(u, v), y(u, v), z(u, v)).$$

If ϕ is differentiable or is of class C^1, we call S a differentiable surface.

Geometrically, we can imagine Φ as some deformation of D into the surface S.

7.3.1 Tangent Vectors and Regular Surfaces

Recall that when we discuss paths, we describe the path as the image of some map $c(t) : [a, b] \mapsto \mathbb{R}^n$, and we noted that the derivative of c with respect to t at some point t_0 is tangent to the curve at $c(t_0)$. We similarly describe this idea for parameterized surfaces.

Definition 7.4: Tangent Vectors to a Parameterized Surface

We can obtain two tangent vectors at a point $\Phi(u_0, v_0)$ by considering the derivative of Φ with respect to either u or v at that point. That is, both

- $T_u = \frac{\partial}{\partial u}\Phi(u_0, v_0)$
- $T_v = \frac{\partial}{\partial v}\Phi(u_0, v_0)$

are tangent vectors to Φ at $\Phi(u_0, v_0)$.

The expression

$$T_u \times T_v$$

is particularly important, especially when we discuss surface integrals later. For now, just observe that since T_u and T_v are tangent vectors. Borrowing the wisps of some ideas from linear algebra and calculus, the partial derivative with respect to a variable for a function represents the "ideal" linear approximation of a function at that point. This allows us to construct a tangent line. Consequently, with two partial derivatives, we can express the ideal linear approximation of a function at that point, which takes the form of a tangent line. As T_u and T_v are both tangent vectors to the the surface, $T_u \times T_v$ will be normal to the surface and hence characterizes the tangent plane completely with Φ.

Example 7.10: MT Section 7.3 Problem 9

Find an expression for a unit vector normal to the surface

$$(x, y, z) = (\cos v \sin u, \sin v \sin u, \cos u)$$

at the image of a point (u, v) for u in $[0, \pi]$ and v in $[0, 2\pi]$. Identify this surface.

This is the unit sphere.

$$\begin{aligned} x^2 + y^2 + z^2 &= \cos^2 v \sin^2 u + \sin^2 v \sin^2 u + \cos^2 u \\ &= \sin^2 u + \cos^2 u \\ &= 1. \end{aligned}$$

I'd actually still have to prove that it's surjective, and that there exists a (u, v) such that $\Phi(u, v) = (x, y, z)$ for any $x^2 + y^2 + z^1 = 1$, but I won't do it here.

Onto finding the unit normal vector.

$$\begin{aligned} T_u &= (\cos v \cos u, \sin v \cos u, -\sin u) \\ T_v &= (-\sin v \sin u, \cos v \sin u, 0) \\ T_u \times T_v &= (\sin^2 u \cos v, \sin^2 u \sin v, \cos^2 v \sin u \cos u + \sin^2 v \sin u \cos u) \end{aligned}$$

$$
\begin{aligned}
&= (\sin^2 u \cos v, \sin^2 u \sin v, \sin u \cos u) \\
&= \sin u(\sin u \cos v, \sin u \sin v, \cos u)
\end{aligned}
$$

Which is parallel to $(\sin u \cos v, \sin u \sin v, \cos u)$, which will be the normal vector of interest. Denote this vector N. Normalizing,

$$
\begin{aligned}
\|N\| &= \sin^2 u \cos^2 v + \sin^2 u \sin^2 v + \cos^2 u \\
&= \sin^2 u + \cos^2 u \\
&= 1
\end{aligned}
$$

so N is already a unit normal vector.

7.3.2 Regular Surfaces

If $T_u \times T_v = 0$, the tangent plane does not exist, and we claim without proof that this means there is a sharp corner.

Definition 7.5: Regular Surfaces

A surface S is regular at $\Phi(u_0, v_0)$ provided that

$$T_u \times T_v \neq 0 \quad \text{at} \quad (u_0, v_0).$$

A surface S is regular without qualifiers if for all points $\Phi(u_0, v_0)$, S is regular.

This is actually a tricky definition, and I must clarify something by discussion and then by example. In the footnote of the source text, the regularity of a surface depends on the parameterization Φ. So simply considering $T_u \times T_v$ is sufficient to identify whether or not a *given parameterization* is regular. However, this is insufficient to identify if a *given surface* is regular. A given parameterization can identify if certain points are regular, but cannot actually say if certain points are irregular. Thus,

> To identify if a point is regular, one must describe a singular parameterization where the the point is regular. That a parameterization renders a point irregular does not entail that the surface is actually irregular at that point.

Example 7.11: MT Section 7.3 Problem 22

The image of the parameterization

$$\Phi(u, v) = (a \sin u \cos v, b \sin u \sin v, c \cos u)$$

with $b < a$, $0 \leq u \leq \pi$, $0 \leq v \leq 2\pi$ parameterizes the ellipsoid.

(a) Show that all points in the image of Φ satisfy

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$$

(the Cartesian equation of an ellipsoid).

(b) Show that the image surface is regular at all points.

(a)

$$\begin{aligned}\frac{x(u,v)^2}{a^2} + \frac{y(u,v)^2}{b^2} + \frac{z(u,v)^2}{c^2} &= \frac{a^2 \sin^2 u \cos^2 v}{a^2} + \frac{b^2 \sin^2 u \sin^2 v}{b^2} + \frac{c^2 \cos^2 u}{c^2} \\ &= \sin^2 u \cos^2 v + \sin^2 u \sin^2 v + \cos^2 u \\ &= \sin^2 u + \cos^2 u \\ &= 1.\end{aligned}$$

(b) We first compute $\mathbf{n}$,

$$\begin{aligned}T_u &= (a \cos u \cos v, b \cos u, \sin v, -c \sin u) \\ T_v &= (-a \sin u \sin v, b \sin u \cos v, 0) \\ T_u \times T_v &= (bc \sin^2 u \cos v, ac \sin^2 u \sin v, ab \sin u \cos u \cos^2 v + ab \sin u \cos u \sin^2 v) \\ &= (bc \sin^2 u \cos v, ac \sin^2 u \sin v, ab \sin u \cos u) \\ &= \sin u(bc \sin u \cos v, ac \sin u \sin v, ab \cos u)\end{aligned}$$

This is 0 when $\sin u = 0$, where $u = 0$ and $u = \pi$. If $\sin u \neq 0$ but $\mathbf{n} = 0$, then $\sin v = \cos v = 0$, which is not possible. So for $\Phi(u, v)$ for any $u \neq 0$, $\Phi(u, v)$ is regular. Letting v be arbitrary,

$$\begin{aligned}\Phi(0, v) &= (0, 0, c) \\ \Phi(\pi, v) &= (0, 0, -c)\end{aligned}$$

Thus every point in the image of Φ is regular except potentially $(0, 0, \pm c)$.

If we consider the parameterization

$$\Phi(a \cos u, b \sin u \cos v, c \sin u \sin v),$$

we claim that this is also the parameterization of the same ellipsoid, and importantly, has surface normal vector

$$\sin u(bc \cos u, ac \sin u \cos v, ab \sin u \sin v)$$

which has an irregularity at $u = 0, \pi$. $\Phi(0, v)$ here is $(\pm a, 0, 0)$. Then every point on the ellipsoid except potentially $(\pm a, 0, 0)$ is regular, so the point at $(0, 0, \pm c)$ is regular. The first parameterization then confirms that $(\pm a, 0, 0)$ is regular, so every point on the ellipsoid is regular.

7.3.3 Tangent Planes to Parameterized Surfaces

Definition 7.6: Tangent Plane of to a Parameterized Surface

If a parameterized surface $\Phi : D \subset \mathbb{R}^2 \mapsto \mathbb{R}^3$ is regular at $\Phi(u_0, v_0)$, we define the tangent plane of the surface at $\Phi(u_0, v_0)$ to be the plane determined by T_u and T_v given as

$$(x - x(u,v), y - y(u,v), z - z(u,v)) \cdot (T_u(u_0, v_0) \times T_v(u_0, v_0)) = 0.$$

More compactly, let $\Phi(u_0, v_0) = (x_0, y_0, z_0)$ and let $\mathbf{n} = (T_u \times T_v)(u_0, v_0)$. Then the tangent plane can be given as

$$(x - x_0, y - y_0, z - z_0) \cdot \mathbf{n} = 0.$$

We can get more extreme if we define $r = (x, y, z)$,

$$(r - \Phi(u_0, v_0)) \cdot \mathbf{n} = 0.$$

Example 7.12: MT Section 7.3 Problem 3

Find an equation for the plane tangent to $(x, y, z) = (u^2, u \sin e^v, \frac{1}{3} u \cos e^v)$ at $(13, -2, 1)$.

As we need to compute $\mathbf{n}$ for the u and v that yield $(13, -2, 1)$, we might try to find those u and v. This is not necessary here, and in general, it is better to compute $T_u \times T_v$ first and see if we get lucky, since nothing of value is lost, as you would need to compute $T_u \times T_v$ anyways.

$$\begin{aligned}
T_u &= \frac{\partial}{\partial u}\left(u^2, u \sin e^v, \frac{1}{3} u \cos e^v\right) \\
&= \left(2u, \sin e^v, \frac{1}{3} \cos e^v\right) \\
T_v &= \frac{\partial}{\partial v}\left(u^2, u \sin e^v, \frac{1}{3} u \cos e^v\right) \\
&= \left(0, u e^v \cos e^v, -\frac{1}{3} u e^v \sin e^v\right) \\
T_u \times T_v &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 2u & \sin e^v & \frac{1}{3} \cos e^v \\ 0 & u e^v \cos e^v & -\frac{1}{3} u e^v \sin e^v \end{vmatrix} \\
&= \left(-\frac{1}{3} u e^v \sin^2 e^v - \frac{1}{3} u e^v \cos^2 e^v, \frac{2}{3} u^2 e^v \sin e^v, 2u^2 e^v \cos e^v\right)
\end{aligned}$$

$$= \left(-\frac{1}{3}ue^v, \frac{2}{3}u^2e^v \sin e^v, 2u^2e^v \cos e^v\right)$$
$$= -\frac{1}{3}ue^v(1, -2u \sin e^v, -6u \cos e^v)$$

Since $-\frac{1}{3}ue^v$ is a scalar quantity, $T_u \times T_v$ is parallel to $(1, -2u \sin e^v, -6u \cos e^v)$, that vector is also normal to the surface, and we can use this to characterize the plane instead. The key observation is that

$$-2u \sin e^v = -2y$$
$$-6u \cos e^v = -18z$$

So the vector

$$(1, -2y, -18z)$$

is normal to the plane.

Since $(x, y, z) = (13, -2, 1)$, this yields the normal vector

$$(1, 4, -18).$$

Then the plane is given as

$$1(x - 13) + 4(y + 2) + -18(z - 1) = 0.$$

Example 7.13: MT Section 7.3 Problem 14

Find the equation of the plane tangent to the surface

$$(x, y, z) = (u^2, v^2, u^2 + v^2)$$

at the point $(u, v) = (1, 1)$.

Computing the normal vector by the tangent vectors,

$$T_u = (2u, 0, 2u)$$
$$T_v = (0, 2v, 2v)$$
$$T_u \times T_v = (-4uv, -4uv, 4uv).$$

As $(u, v) = (1, 1)$, this gives $\mathbf{n} = (-4, -4, 4)$. It now remains to find the actual point in $\mathbb{R}^3$ that $(u, v) = (1, 1)$ maps to. This is $(x, y, z) = (1, 1, 2)$. Hence the tangent plane is

$$-4(x - 1) - 4(y - 1) + 4(z - 2) = 0.$$

Example 7.14: MT Section 7.3 Problem 15

Find a parameterization of the surface

$$z = 3x^2 + 8xy$$

and use it to find the tangent plane at $x = 1$, $y = 0$ and $z = 3$. Compare your answer with that using graphs.

We present the parameterization of letting x and y play the value of u and v, since z is easily expressible in terms of x and y:

$$\Phi(u,v) = (u, v, 3u^2 + 8uv).$$

Computing the normal vector by finding the tangent vectors,

$$\begin{aligned} T_u &= (1, 0, 6u + 8v) \\ T_v &= (0, 1, 8u) \\ T_u \times T_v &= (-6u - 8v, -8u, 1) \end{aligned}$$

Then $\mathbf{n} = (-6, -8, 1)$, and then the plane is given as

$$-6(x-1) - 8(y-0) + 1(z-3) = 0.$$

The tangent plane, when expressed as a graph, requires us to find the partial derivatives. Differentiating,

$$\begin{aligned} f_x &= 6x + 8y \\ f_y &= 8x. \end{aligned}$$

Then $(f_x, f_y)(1, 0) = (6, 8)$. This gives the tangent plane

$$z = 3 + 6(x-1) + 8(y-0)$$

which is the same after rearrangement.

Example 7.15: MT Section 7.3 Problem 18

Given a sphere of radius 2 centered at the origin, find the equation for the plane that is tangent to the point $(1, 1, \sqrt{2})$ by considering the sphere as:

(a) A surface parameterized as

$$\Phi(\theta, \phi) = (2\cos\theta\sin\phi, 2\sin\theta\sin\phi, 2\cos\phi)$$

(b) The level surface

$$f(x,y,z) = x^2 + y^2 + z^2$$

(c) The graph of

$$g(x,y) = \sqrt{4 - x^2 - y^2}.$$

(a) It's simpler to represent the parameterization as

$$\Phi(\theta, \phi) = 2(\cos\theta\sin\phi, \sin\theta\sin\phi, \cos\phi).$$

Then, finding $\mathbf{n}$,

$$T_\theta = 2(-\sin\theta\sin\phi, \cos\theta\sin\phi, 0)$$

$$\begin{aligned} T_\phi &= 2(\cos\theta\cos\phi, \sin\theta\cos\phi, -\sin\phi) \\ T_\theta \times T_\phi &= 4(-\cos\theta\sin^2\phi, -\sin\theta\sin^2\phi, -\sin^2\theta\sin\phi\cos\phi - \cos^2\theta\sin\phi\cos\phi) \\ &= 4(-\cos\theta\sin^2\phi, -\sin\theta\sin^2\phi, -\sin\phi\cos\phi) \\ &= -4\sin\phi(\cos\theta\sin\phi, \sin\theta\sin\phi, \cos\phi) \end{aligned}$$

We must now find the (θ, ϕ) such that $\Phi(\theta, \phi) = (1, 1, \sqrt{2})$. Since $z = 2\cos\phi$, $\phi = \pi/4$ as ϕ ranges from 0 to π. As $\sin\theta = \cos\theta = 1/\sqrt{2}$, $\theta = \pi/4$. Then $\mathbf{n}$ is parallel to

$$(1, 1, \sqrt{2})$$

This gives the tangent plane

$$1(x-1) + 1(y-1) + \sqrt{2}(z - \sqrt{2}).$$

(b) Taking the gradient of the function gives

$$\nabla f = (2x, 2y, 2z).$$

Then, evaluating gives

$$\nabla f(2, 2, 2\sqrt{2})$$

which gives the tangent plane

$$2(x-1) + 2(y-1) + 2\sqrt{2}(z - \sqrt{2}).$$

(c) The partial derivatives are

$$\begin{aligned} f_x &= \frac{-x}{\sqrt{4 - x^2 - y^2}} \\ f_y &= \frac{-y}{\sqrt{4 - x^2 - y^2}}. \end{aligned}$$

Evaluating gives

$$\begin{aligned} f_x(1,1) &= \frac{-1}{\sqrt{2}} \\ f_y(1,1) &= \frac{-1}{\sqrt{2}} \end{aligned}$$

which yields the tangent plane

$$z = \sqrt{2} - \frac{1}{\sqrt{2}}(x-1) - \frac{1}{\sqrt{2}}(x-1).$$

7.4 "Area of a Surface" (MT 7.4)

Just as how we used arc-length as an introduction to the idea of integrating over a path, we will do something similar, talking about the total "content" that a surface contains, here, the surface area.

Definition 7.7: Area of a Parameterized Surface

We define the surface area of a parameterized surface, denoted $A(S)$, as

$$A(s) = \iint_D \|T_u \times T_v\| \, du \, dv.$$

Example 7.16: MT Section 7.4 Problem 1

Find the surface area of the unit sphere S represented parametrically by $\Phi : D \mapsto S \subset \mathbb{R}^3$, where D is the rectangle $0 \leq \theta \leq 2\pi$, $0 \leq \phi \leq \pi$ and Φ is given by the equations

$$(x, y, z) = (\cos\theta \sin\phi, \sin\theta \sin\phi, \cos\phi).$$

$$\begin{aligned} T_\theta &= (-\sin\theta\sin\phi, \cos\theta\sin\phi, 0) \\ T_\phi &= (\cos\theta\cos\phi, \sin\theta\cos\phi, -\sin\phi) \\ T_\theta \times T_\phi &= (-\sin^2\phi\cos\theta, -\sin^2\phi\sin\theta, -\sin\phi\cos\phi) \\ &= -\sin\phi(\sin\phi\cos\theta, \sin\phi\sin\theta, \cos\phi) \\ \|T_\theta \times T_\phi\| &= \sin\phi. \end{aligned}$$

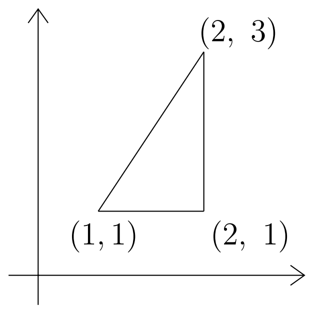
Integrating,

$$\begin{aligned} \int_0^{2\pi}\int_0^{\pi} \sin\phi \, d\phi \, d\theta &= 2\pi \int_0^{\pi} \sin\phi \, d\phi \\ &= 2\pi \left[-\cos\phi\right]_{\phi=0}^{\phi=\pi} \\ &= 4\pi. \end{aligned}$$

Example 7.17: MT Section 7.4 Problem 7

Use a surface integral to find the area of the triangle T in $\mathbb{R}^3$ with vertices at $(1, 1, 0)$, $(2, 1, 2)$ and $(2, 3, 3)$.
Hint: Write the triangle as the graph $z = g(x, y)$ over a triangle T^* in the xy-plane.

To identify g, we first find T^* by projecting our vertices. This gives the triangle here:



We can represent this with the bounds of $1 \leq x \leq 2$ and $1 \leq y \leq 2x - 1$. It now suffices to find g. Using a system of linear equations, we find that

$$g(x, y) = -2x - \frac{1}{2}y + \frac{5}{2}.$$

We now give the parameterization of

$$\Phi(u, v) = (u, v, g(u, v)).$$

We now compute the surface area differential,

$$\begin{aligned}
T_u &= (1, 0, g_u) \\
&= (1, 0, -2) \\
T_v &= (0, 1, g_v) \\
&= (0, 1, -1/2) \\
T_u \times T_v &= (-1/2, -2, 1) \\
\|T_u \times T_v\| &= \sqrt{\frac{1}{4} + 4 + 1} \\
&= \sqrt{\frac{21}{4}} \\
&= \frac{\sqrt{21}}{2}.
\end{aligned}$$

Integrating,

$$\begin{aligned}
\iint_D \frac{\sqrt{21}}{2}\, dA &= \frac{\sqrt{21}}{2} \iint_D dA \\
&= \frac{\sqrt{21}}{2}
\end{aligned}$$

as $\iint_D dA$ is the area of the triangle, which has area 1, since it has base 1 and height 2.

7.5 "Surface Integrals of Vector Fields" (MT 7.6)

Now that we have discussed parameterized surfaces, a means to represent arbitrary surfaces in $\mathbb{R}^3$ using a a map Φ, we are ready to define surface integrals, a means to extend the idea of line integrals and work done over a path to the sum total work done on a surface.

Definition 7.8: Surface Integral

Let F be a vector field defined on S, the image of a parameterized surface Φ defined over a domain D. The surface integral of F over Φ, denoted $\iint_\Phi F \cdot dS$ is defined as

$$\iint_\Phi F \cdot dS = \iint_D F \cdot (T_u \times T_v)\, du\, dv.$$

For a given path c, we have not discussed the notion of orientation. Observe that if we have some path c that describes a squiggly line in space, there are two ways to traverse said line, from beginning to end, and from end to beginning, if you pick one point as the beginning arbitrarily. For this reason, a path cannot only be defined as c, but also must have the end points of the path (generally a and b specified). Note that for a given path

$$c(t) : (c_i(t)), \quad a \le t \le b,$$

we can define the reverse of the path,

$$c_r(t) : (c_i(b-t)), \quad a \le t \le b$$

which has the same image, but just traverses it backwards. This becomes particularly important when we describe surfaces, since there's not a really a good way to identify the orientation from really looking at it, since it's a two dimensional object.

Definition 7.9: Oriented Surface

An oriented surface is a two-sided surface with one side specified as the outside or positive side, and the other side the inside or negative side. At each point (x, y, z) on the surface, there are two unit normal vectors $\mathbf{n}_1$ and $\mathbf{n}_2$ where $\mathbf{n}_1 = -\mathbf{n}_2$. Each normal can be associated with one side.
To choose an orientation, we pick a normal and label it as the positive side.

Möbius strips are examples of surfaces that do not have two sides. We will only work with surfaces of two sides.

Example 7.18: MT Section 7.6 Problem 2

Evaluate the surface integral

$$\iint_S F \cdot dS$$

where $F(x, y, z) = x\mathbf{i} + y\mathbf{j} + z^2\mathbf{k}$ and S is the surface parameterized by $\Phi(u, v) = (2\sin u, 3\cos u, v)$ where $0 \le u \le 2\pi$ and $0 \le v \le 1$.

First, composing F and Φ,

$$F(\Phi(u, v)) = (2\sin u, 3\cos u, v^2).$$

Next, computing the surface normal vector,

$$T_u = (2\cos u, -3\sin u, 0)$$

$$\begin{aligned} T_v &= (0,0,1) \\ T_u \times T_v &= (-3\sin u, -2\cos u, 0). \end{aligned}$$

Taking the dot product of the composition and the surface normal vector,

$$\begin{aligned} F(\Phi(u,v)) \cdot T_u \times T_v &= (2\sin u, 3\cos u, v^2) \cdot (-3\sin u, -2\cos u, 0) \\ &= -6\sin^2 u - 6\cos^2 u + 0 \\ &= -6. \end{aligned}$$

Integrating,

$$\begin{aligned} \iint_S F \cdot dS &= \int_0^{2\pi} \int_0^1 -6\, dv\, du \\ &= -12\pi. \end{aligned}$$

Example 7.19: MT Section 7.6 Problem 3

Let $F(x,y,z) = (x,y,z)$. Evaluate

$$\iint_S F \cdot dS$$

where S is

(a) The upper hemisphere of radius 3, centered at the origin.
(b) The entire sphere of radius 3, centered at the origin.

We can parameterize both of these by using spherical coordinates as our parameterization,

$$\Phi(u,v) = (3\sin u\cos v, 3\sin u\sin v, 3\cos u).$$

In both cases, v ranges from 0 to 2π. In (a), u ranges from 0 to $\pi/2$, and in (b), u ranges from 0 to π.

The composition is given as

$$F(\Phi(u,v)) = (3\sin u\cos v, 3\sin u\sin v, 3\cos u).$$

The surface normal vector is then

$$\begin{aligned} T_u &= (3\cos u\cos v, 3\cos u\sin v, -3\sin u) \\ T_v &= (-3\sin u\sin v, 3\sin u\cos v, 0) \\ T_u \times T_v &= (9\sin^2 u\cos v, 9\sin^2 u\sin v, 9\cos^2 v\sin u\cos u + 9\sin^2 v\sin u\cos u) \\ &= (9\sin^2 u\cos v, 9\sin^2 u\sin v, 9\sin u\cos u) \\ &= 9\sin u(\sin u\cos v, \sin u\sin v, \cos u). \end{aligned}$$

The dot product of the two is then

$$F(\Phi(u,v)) \cdot T_u \times T_v = (3\sin u\cos v, 3\sin u\sin v, 3\cos u) \cdot 9\sin u(\sin u\cos v, \sin u\sin v, \cos u)$$

$$\begin{aligned}
&= 9\sin u[3\sin^2 u\cos^2 v + 3\sin^2 u\sin^2 v + 3\cos^2 u] \\
&= 9\sin u[3\sin^2 u + 3\cos^2 u] \\
&= 9\sin u(3) \\
&= 27\sin u.
\end{aligned}$$

(a) Integrating,

$$\begin{aligned}
\iint_S F\cdot dS &= \int_0^{2\pi}\int_0^{\pi/2} 27\sin u\, du\, dv \\
&= 54\pi\int_0^{\pi/2} \sin u\, du \\
&= 54\pi\left[-\cos u\right]_{u=0}^{u=\pi/2} \\
&= 54\pi.
\end{aligned}$$

(b) Integrating,

$$\begin{aligned}
\iint_S F\cdot dS &= \int_0^{2\pi}\int_0^{\pi} 27\sin u\, du\, dv \\
&= 54\pi\int_0^{\pi} \sin u\, du \\
&= 54\pi\left[-\cos u\right]_{u=0}^{u=\pi} \\
&= 108\pi.
\end{aligned}$$

Example 7.20: MT Section 7.6 Problem 4

Let $F(x,y,z) = 2x\mathbf{i} - 2y\mathbf{j} + z^2\mathbf{k}$. Evaluate

$$\iint_S F\cdot dS,$$

where S is the cylinder $x^2+y^2=4$ where $z\in[0,1]$.

We use cylindrical coordinates as our parameterization,

$$\Phi(u,v) = (2\cos u, 2\sin u, v),$$

where $u\in[0,2\pi]$ and $v\in[0,1]$. Composing,

$$F(\Phi(u,v)) = (4\cos u, -4\sin u, v^2).$$

Computing the surface normal vector,

$$\begin{aligned}
T_u &= (-2\sin u, 2\cos u, 0) \\
T_v &= (0,0,1)
\end{aligned}$$

$$\begin{aligned} T_u \times T_v &= (2\cos u, 2\sin u, 0) \\ &= 2(\cos u, \sin u, 0). \end{aligned}$$

Taking the dot product,

$$\begin{aligned} F(\Phi(u,v)) \cdot T_u \times T_v &= (4\cos u, -4\sin u, v^2) \cdot 2(\cos u, \sin u, 0) \\ &= 8\cos^2 u - 8\sin^2 u \\ &= 8\cos 2u. \end{aligned}$$

Integrating,

$$\begin{aligned} \iint_S F \cdot dS &= \int_0^{2\pi} \int_0^1 8\cos 2u \, dv \, du \\ &= 8 \int_0^{2\pi} \cos 2u \, du \\ &= 4\left[\sin 2u\right]_{u=0}^{u=2\pi} \\ &= 0. \end{aligned}$$

Example 7.21: MT Section 7.6 Problem 14

Evaluate the surface integral $\iint_S F \cdot \mathbf{n} \, dA$, where $F(x,y,z) = \mathbf{i} + \mathbf{j} + z(x^2+y^2)^2\mathbf{k}$ and S is the surface of the cylinder $x^2 + y^2 \leq 1$, $0 \leq z \leq 1$.

This doesn't look the standard surface integral, but $n \, dA = T_u \times T_v$, as the latter is normal to the plane (the $\mathbf{n}$), and also the surface area differential, as have discussed before, so this really is just the normal surface integral. However, in this problem, identifying the surface area differential and the normal vector is easier than compute the surface normal vector.

Importantly, however, the surface of the sphere is composed of three parts, the two circular caps, and the actual sheet of paper that is taped into a roll. So we actually have to do three surface integrals.

$$\Phi_1(u,v) = (\cos u, \sin u, v) \quad (u,v) \in ([0,2\pi],[0,1]) \qquad \text{(side)}$$
$$\Phi_2(u,v) = (u\cos v, u\sin v, 0) \quad (u,v) \in ([0,1],[0,2\pi]) \qquad \text{(bottom)}$$
$$\Phi_3(u,v) = (u\cos v, u\sin v, 1) \quad (u,v) \in ([0,1],[0,2\pi]) \qquad \text{(top)}$$

We actually have to make sure these are all oriented the same way. Let's try to make all of their normals point outside. The side needs to point out–since each point is of the form (x,y,z), and we want it to point away from the origin and the surface is a cylinder's side, the normal is the same direction as the actual point, so $\mathbf{n} = (x,y,0)$. This is a unit normal vector since $x^2+y^2 = 1$. The top needs to point up, which we can accomplish as $(0,0,1)$. Similarly, the bottom will be $(0,0,-1)$.

Composing,

$$F(\Phi_1(u,v)) = (1,1,v) \qquad \text{(side)}$$

$$F(\Phi_2(u,v)) = (1,1,0) \tag{bottom}$$
$$F(\Phi_3(u,v)) = (1,1,u^4). \tag{top}$$

We can then find $\mathbf{n} \cdot dA$ by considering the area differentials. For the polar maps on the top and bottom, the area differential is r. Then, integrating,

$$\begin{aligned}
\iint_S F \cdot \mathbf{n}\, dA &= \int_0^{2\pi} \int_0^1 (1,1,v) \cdot (\cos(\theta), \sin(\theta), 0)\, dz\, d\theta && \text{(side)} \\
&= \int_0^{2\pi} \cos(\theta) + \sin(\theta)\, d\theta \\
&= [\cos(\theta) - \sin(\theta)]_{\theta=0}^{\theta=2\pi} \\
&= 0 \\
\iint_S F \cdot \mathbf{n}\, dA &= \int_0^1 \int_0^{2\pi} (r\cos(\theta), r\sin(\theta), 0) \cdot (0,0,-1) u\, d\theta\, dr && \text{(bottom)} \\
&= \int_0^{2\pi} 0\, d\theta \\
&= 0 \\
\iint_S F \cdot \mathbf{n}\, dA &= \int_0^1 \int_0^{2\pi} (r\cos(\theta), r\sin(\theta), r^4) \cdot (0,0,1) r\, d\theta\, dr && \text{(top)} \\
&= 2\pi \int_0^1 r^4\, dr \\
&= \frac{2\pi}{5} \left[r^5\right]_{r=0}^{r=1} \\
&= \frac{2\pi}{5}.
\end{aligned}$$

Summing all of them gives the overall surface integral of

$$\frac{2\pi}{5}.$$

8 "The Integral Theorems of Vector Analysis" (MT 8)

The unifying idea between differentiation and integration of single variable calculus is the fundamental theorem of calculus. If we have some function $F(x)$ where $\frac{d}{dx}F(x) = f(x)$, then

$$\int_a^b f(x)\,dx = F(b) - F(a),$$

which lets us represent a total action as a special related function only at the end points.

We will see extensions of this idea with vector analysis in many different forms, as we can define the idea of total action over some region in many different ways, be it line, surface, or volumetric integrals.

8.1 "Green's Theorem" (MT 8.1)

Green's theorem relates a line integral on a closed curve C in $\mathbb{R}^2$ to a double integral over the region enclosed by C. In this sense, the total action is the double integral, and the end points of the region is the boundary.

Following from our previous discussions on orientation, Green's theorem requires a counter-clockwise traversal over the boundary.

Theorem 8.1: Green's Theorem (Vector form)

Let D be a simple region and let C be its boundary. Let $F = (F_1, F_2)$ where F_1 and F_2 are of class C^1. Let $c(t)$ be a clockwise parameterization of C. Then

$$\int_{C^+} F \cdot\, ds = \int_a^b F(c(t))\|c'(t)\|\,dt = \iint_D \operatorname{curl} F \cdot \mathbf{k}\,dA$$

The middle term in the equality is just to explicitly indicate what's going on in the line integral.

Green's theorem is not generally presented in this form. Instead, Green's theorem is presented in differential form, but I find this to make its connection with its generalization, Stoke's Theorem (which we discuss later), less clear.

Let $F = (P, Q)$, and let $(dx, dy) = (x'(t), y'(t))$. Then we can alternatively write

$$\int_{C^+} P\,dx + Q\,dy = \iint_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)$$

The main idea of Green's theorem, its proof, and a way to remember how to handle regions that have punctures is the following diagram:

For any given region, we can split it into multiple regions and apply Green's theorem on each region and sum all of them together. For the sub-regions that contain boundaries that are not boundaries of the original region, as those are on the inside of the region, since they border another subregion which runs along said border in another direction, these cancel out. See this with any edge inside the cloud having two arrows in opposite directions. If we divide infinitely down, we still retain just the boundary.

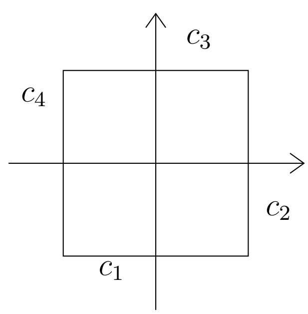
The final idea is that the infinitesimal line integral around an infinitesimal region is its curl, which I will loosely justify it as the line integral counterclockwise around a region is a rotational description, which is captured by the curl.

Example 8.1: MT Section 8.1 Problem 3

Verify Green's theorem for

$$D = [-1, 1] \times [-1, 1] \quad F = (P, Q) = (-y, x).$$

Let's first do the line integral. Unfortunately, there's not a great way to parameterize this as a single curve, so we'll have to use a union of curves, one for each edge. The region is given as



and we parameterize the curves as

$$\begin{aligned} c_1 &= (t, -1) \\ c_2 &= (1, t) \\ c_3 &= (t, 1) \\ c_4 &= (-1, t). \end{aligned}$$

Note that these curves must be oriented counterclockwise for Green's theorem, and c_2 and c_4 are not done so. This can be readily fixed – we just take the negative of their contributions to flip their orientations.

Computing the line integrals,

$$\begin{aligned}
\int_{-1}^{1} F(c_1(t)) \cdot c_1' t(t)\, dt &= \int_{-1}^{1} (1, t) \cdot (1, 0)\, dt = \int_{-1}^{1} 1\, dt = 2 \\
\int_{-1}^{1} F(c_2(t)) \cdot c_2'(t)\, dt &= \int_{-1}^{1} (-t, 1) \cdot (0, 1)\, dt = \int_{-1}^{1} 1\, dt = 2 \\
-\int_{-1}^{1} F(c_3(t)) \cdot c_3'(t)\, dt &= -\int_{-1}^{1} (-1, t) \cdot (1, 0)\, dt = -\int_{-1}^{1} -1\, dt = 2 \\
-\int_{-1}^{1} F(c_4(t)) \cdot c_4'(t)\, dt &= -\int_{-1}^{1} (-t, -1) \cdot (0, 1)\, dt = -\int_{-1}^{1} -1\, dt = 2.
\end{aligned}$$

Summing gives 8 for the total line integral.

Now, as a double integral, we note that $Q_x - P_y = 1 + 1 = 2$, so we integrate

$$\int_{-1}^{1} \int_{-1}^{1} 2\, dx\, dy = 8.$$

8.1.1 Green's Theorem for Area

Since the right hand side of Green's theorem is a double integral, if we pick a function F such that its scalar curl is 1, we get a regular area integral for a rather quirky way to compute area.

Theorem 8.2: Area of a Region

If C is a simple closed curve that bounds a region to which Green's theorem applies, then the area of the region D bounded by $C = \partial D$ is

$$A = \frac{1}{2} \int_{\partial D} (-y, x)\, ds.$$

Example 8.2: MT Section 8.1 Problem 10

Find the area of the disc D of radius R using Green's theorem.

We use polar coordinates and parameterize the boundary as

$$c(\theta) = (\cos\theta, \sin\theta).$$

The derivative is then

$$c'(\theta) = (-\sin\theta, \cos\theta).$$

Composing, $F \circ c = (-\sin\theta, \cos\theta)$. Taking the dot product,

$$\begin{aligned} F(c(\theta)) \cdot c'(\theta) &= (-\sin\theta, \cos\theta) \cdot (-\sin\theta, \cos\theta) \\ &= \sin^2\theta + \cos^2\theta \\ &= 1. \end{aligned}$$

Then, integrating over the boundary,

$$\frac{1}{2}\int_0^1 \int_0^{2\pi} 1\, d\theta\, dr = \pi.$$

8.2 "Stokes' Theorem" (MT 8.2)

We are now ready to discuss what is probably the most famous (or infamous) theorem in vector calculus, Stokes' theorem, a generalization of Green's theorem to arbitrary parameterized surfaces. Stokes' theorem relates the line integral of a vector field around a simple closed curve C in $\mathbb{R}^3$ to an integral over a surface S for which C is the boundary. Likewise, the total action is the surface, and the endpoints are the boundary.

We first discuss Stokes' theorem over the special case of graphs, and then generalize this to arbitrary parameterized surfaces.

8.2.1 On Graphs

On graphs, that $z = f(x, y)$ is greatly simplifying. Observe that for the surface defined as the graph of f, if the domain of f is a subset D which has some boundary that can be parameterized by the function $c(t) = (x(t), y(t))$ for $t \in [a, b]$, we can describe the boundary of the graph object by evaluating the given x and y values (for that t) by f as well. In a certain sense, if D is some shape on the xy-plane, then the graph of f is that shape, lifted into xyz-space, with deformations in the z-direction. To capture the deformations on the edges, we simply the boundary of the domain–given to us described by c, and append the proper z value. Since $z = f(x, y)$, the mapping $p(t) = (x(t), y(t), f(x(t), y(t)))$ is a mapping that describes this boundary in space. Since t going from a to b versus b to a are flipped, the image of p similarly has its orientation induced by p.

Theorem 8.3: Stokes' Theorem on Graphs

Let S be an oriented surface defined by the graph of a C^2 function $f : D \subset \mathbb{R}^2 \mapsto \mathbb{R}$ where Green's theorems applies on D, and let F be a C^1 vector field on S. Let ∂S be the oriented boundary curve of the graph.
Then

$$\iint_S \operatorname{curl}(F) \cdot dS = \int_{\partial S} F \cdot ds.$$

The proof is *long*, and you should totally skip it if you want.

Proof. The key observation is to write the right hand side as a line integral over a region in the xy-plane. This is possible as $z = f(x, y)$ and that S is the graph of f–hence specifying

$(x, y) \in D$, with f, describes S exactly, since for each x and y, we have that $z = f(x, y)$. More importantly, however, is that $z = f(x, y)$ together with the chain rule transforms the right hand side into a line integral over D, to which we can use Green's theorem.

We manipulate both sides independently, then show equality. Let $F = (F_1, F_2, F_3)$

First, we handle the left hand side. Skipping the curl calculation, $\text{curl}(F) = (D_yF_3 - D_zF_2, D_zF_1 - D_xF_3, D_xF_2 - D_yF_1)$. Using the surface integral of graphs formula, we then have that

$$\iint_S \text{curl}(F) \cdot dS = \iint_D -(D_xf)(D_yF_3 - D_zF_2) - (D_yf)(D_zF_1 - D_xF_3) + (D_xF_2 - D_yF_1)\,dA$$

We now handle the right hand side. Unpacking the integral gives

$$\int_{\partial S} F \cdot ds = \int_a^b F_1x'(t) + F_2y'(t) + F_3z'(t).$$

As $z = f(x, y)$ and x and y are functions of t, the chain rule gives that $z'(t) = D_xz \cdot x'(t) + D_yz \cdot y'(t)$. We now consider the parameterized line integral and factor:

$$\begin{aligned}\int_{\partial S} F \cdot ds &= \int_a^b F_1x'(t) + F_2y'(t) + F_3(D_xz \cdot x'(t) + D_yz \cdot y'(t))\,dt \\ &= \int_a^b (F_1 + F_3D_xz)x'(t) + (F_2 + F_3D_yz)y'(t)\,dt\end{aligned}$$

Since we have that Green's theorem applies to the region D, and skipping the curl calculation, this gives that

$$\iint_{\partial S} F \cdot ds = \iint_D D_x(F_2 + F_3D_yz) - D_y(F_1 + F_3D_xz)\,dA.$$

The chain rule and product rule unpacks this further–note that F are functions of x, y and z, but z is a function of both x and y. For clarify, note that x and y are completely independent – $D_xy = D_yx = 0$. We handle each term at a time.

Firstly, $D_x(F_2) = D_xF_2 + D_zF_2 \cdot D_xz$. Then, $D_x(F_3 \cdot D_yz)$, by the product rule, is $F_3 \cdot D_x(D_yz) + D_x(F_3) \cdot (D_yz)$. $D_x(D_yz)$ is $D_{xy}z$. By using the chain rule on $D_x(F_3)$, we obtain $D_xF_3 + D_zF_3 \cdot D_xz$. This sums together for $D_xF_2 + D_zF_2 \cdot D_xz + D_xF_3 \cdot D_yz + D_zF_3 \cdot D_xz \cdot D_yz + F_3 \cdot D_{xy}z$.

Repeating this for the second term of the integrand gives the negative of $D_yF_1 + D_zF_1 \cdot D_yz + D_yF_3 + D_xz + D_zF_3 \cdot D_yz \cdot D_xz + F_3 \cdot D_{yx}z$.

As the differential operator is commutative with itself, $F_3 \cdot D_{xy}z = F_3D_{yx}z$ and $D_zF_3 \cdot D_xz \cdot D_yz = D_zF_3 \cdot D_yz \cdot D_xz$, so the last two terms of each expression drop, which gives us the integral

$$\iint_D D_xF_2 + D_zF_2 \cdot D_xz + D_xF_3 \cdot D_yz - (D_yF_1 + D_zF_1 \cdot D_yz + D_yF_3)\,dA$$

Since $z = f(x, y)$, we can replace all some instances of z with f (namely the ones that line up to $z = f(x, y)$, rather than F). We now rearrange and group, and essentially factor by using the sum derivative rule:

$$\begin{aligned}
\text{RHS} &= \int_{\partial S} F \cdot ds \\
&= \iint_D D_x F_2 + D_z F_2 \cdot D_x f + D_x F_3 \cdot D_y f - (D_y F_1 + D_z F_1 \cdot D_y f + D_y F_3 \cdot D_x f)\, dA \\
&= \iint_D D_z F_2 \cdot D_x f - D_y F_3 \cdot D_x f + D_x F_3 \cdot D_y f - D_z F_1 \cdot D_y f + D_x F_2 - D_y F_1 \, dA \\
&= \iint_D D_x f \cdot (D_z F_2 - D_y F_3) + D_y f \cdot (D_x F_3 - D_z F_1) + (D_x F_2 - D_y F_1)\, dA \\
&= \iint_D -(D_x f)(D_y F_3 - D_z F_2) + -(D_y f)(D_z F_1 - D_x F_3) + (D_x F_2 - D_y F_1)\, dA \\
&= \text{LHS}
\end{aligned}$$

This completes the proof. □

An important consequence of this *version* of Stokes' theorem is that for graphs, once we identify D, we can more or less immediately transform the line integral to a double integral over D. Alternatively,

$$\textstyle\int_{\partial S} ds \longrightarrow \iint_S dS \longrightarrow \iint_D dA$$

Example 8.3: MT Section 8.2 Problem 3

Verify Stokes' theorem for the given surface

$$S = \{(x, y, z) \mid x^2 + y^2 + z^2 = 1, \quad z \geq 0\}$$

oriented as a graph, the boundary

$$\partial S = \{(x, y) \mid x^2 + y^2 = 1\}$$

with the vector field

$$F = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}.$$

Before anything, let's just compute the curl of F.

$$\text{curl}\, F = 0$$

so the surface integral of the curl is just 0,

$$\iint_S \text{curl}\, F \cdot dS = \iint_S 0 \cdot dS = 0.$$

Now we consider the line integral. We present the familiar polar parameterization of the circle,

$$c(t) = (\cos(t), \sin(t), 0).$$

Then $F(c(t)) = (\cos(t), \sin(t), 0)$ and $c'(t) = (-\sin(t), \cos(t), 0)$ which has dot product $F(c(t)) \cdot c'(t) = -\sin(t)\cos(t) + \sin(t)\cos(t) = 0$, so the line integral is also 0, so they're the same.

8.2.2 On Parameterized Surfaces

In it's full power, Stokes' theorem applies for any surface.

Theorem 8.4: Stokes' Theorem on Parameterized Surfaces

Let S be an oriented surface defined by the image of the injective parameterization $\Phi : D \subset \mathbb{R}^2 \mapsto S \subset \mathbb{R}^3$ where D is a region where Green's theorem applies. Let ∂S denote the oriented boundary and let F be a C^1 vector field on S.
Then

$$\iint_S \operatorname{curl}(F) \cdot dS = \int_{\partial S} F \cdot ds.$$

The proof is more or less the same as the graph case, but with more terms; we replace the x and y chain rules with u and v.

If S has no boundary that can be described by a path, the surface integral of the curl F is 0, since the right is a work integral over a boundary, of which there is none.

Green's theorem uses the scalar curl/$\mathbf{k} \cdot \nabla \operatorname{curl}(F)$ by taking S to be the xy-plane in $\mathbb{R}^3$, parameterized by $\Phi(x, y) = (x, y, 0)$. Then the normal vector is given by $T_x \times T_y = (1, 0, 0) \times (0, 1, 0) = (0, 0, 1) = \mathbf{k}$. This is also more or less how the scalar curl is defined.

Example 8.4: MT Section 8.2 Problem 13

Let S be a a hemisphere sitting on top of a cylinder, represented as the union of two surfaces S_1 and S_2, where

$$S_1 = \{(x, y, z) \mid x^2 + y^2 = 1, \quad 0 \leq z \leq 1\}$$

and

$$S_2 = \{(x, y, z) \mid x^2 + y^2 + (z-1)^2 = 1, \quad z \geq 1\}.$$

Let

$$F(x, y, z) = (zx + z^2 y + x)\mathbf{i} + (z^3 yx + y)\mathbf{j} + z^4 x^2 \mathbf{k}.$$

Compute

$$\iint_S \nabla \times F \cdot dS.$$

Hint: Stokes' theorem holds for this surface.

A key idea here is that the boundary of the surface between S_1 and S_2 cancel out, since their boundaries completely intersect, and as we discuss with cases like Green's theorem, they cancel out since when we "encounter" them for the surface integral of S_1 and when we encounter them for S_2, they are oppositely oriented, so that boundary is more or less cancelled out.

Instead, the boundary we work with is where S_1 touches the xy-plane, the unit circle on the xy-plane, given by the parameterization

$$c(t) = (\cos t, \sin t, 0).$$

This gives the tangent vector

$$c'(t)(-\sin t, \cos t, 0)$$

and the composition

$$f(c(t)) = (\cos t, \sin t, 0).$$

The dot product is then

$$\begin{aligned} f(c(t)) \cdot c'(t) &= (\cos t, \sin t, 0) \cdot (-\sin t, \cos t, 0) \\ &= -\sin t \cos t + \sin t \cos t \\ &= 0 \end{aligned}$$

so the line integral is then 0, so by Stokes' theorem, so is the surface integral of the curl.

Example 8.5: MT Section 8.2 Problem 14

Let c consist of straight lines joining $(1, 0, 0)$, $(0, 1, 0)$ and $(0, 0, 1)$ and let S be the triangle with these vertices. Verify Stokes' theorem directly with

$$F = yz\mathbf{i} + xz\mathbf{j} + xy\mathbf{k}.$$

Computing the curl, we have

$$\operatorname{curl} F = (x - x, -y + y, z - z) = (0, 0, 0)$$

so the surface integral over the triangle will be 0.

We now consider the line integral, presenting the following paths:

$$\begin{aligned} c_1(t) &= (1 - t, t, 0) \\ c_2(t) &= (0, 1 - t, t) \\ c_2(t) &= (t, 0, 1 - t) \end{aligned}$$

all from $0 \leq t \leq 1$. These have tangent vectors

$$\begin{aligned} c_1'(t) &= (-1, 1, 0) \\ c_2'(t) &= (0, -1, 1) \\ c_3'(t) &+ (1, 0, -1) \end{aligned}$$

Composing, we have

$$\begin{aligned} F(c_1(t)) &= (0, 0, t - t^2) \\ F(c_2(t)) &= (t - t^2, 0, 0) \\ F(c_3(t)) &= (0, t - t^2, 0) \end{aligned}$$

which have dot products

$$\begin{aligned} F(c_1(t)) \cdot c_1'(t) &= 0 \\ F(c_2(t)) \cdot c_2'(t) &= 0 \\ F(c_3(t)) \cdot c_3'(t) &= 0 \end{aligned}$$

so all the line integrals are 0, which agrees with the surface integral.

8.3 "Conservative Vector Fields" (MT 8.3)

We've previously previewed the fundamental theorem of line integrals, which states that for a given vector field F, if there exists a scalar function f such that $F = \nabla f$, then for any path $c(t)$ with $a \leq t \leq b$, that only the endpoints matter, and

$$\int_C F \cdot ds = f(c(b)) - f(c(a)).$$

This was detailed as a gradient field. These are also known as *conservative* vector fields. The most critical condition (as in, the easiest one to verify) is that

$$\text{curl}\, F = 0.$$

Theorem 8.5: Equal Characterizations of Conservative Vector Fields

Let F be a C^1 vector field defined on $\mathbb{R}^3$, except possibly for a finite number of points. The following conditions on F are all equivalent:

- For any simple oriented curve C, $\int_C F \cdot ds = 0$.
- For any two oriented simple curves C_1 and C_2 that have the same endpoints,

$$\int_{C_1} F \cdot ds = \int_{C_2} F \cdot ds.$$

- F is the gradient of some function f; that is, $F = \nabla f$ (and if F has one or more exceptional points where it fails to be defined, f is also undefined there).
- $\nabla F = 0$.

A vector field satisfying one (and, hence, all) of the conditions is called a conservative vector field.

8.3.1 Undoing the Gradient

Since we are interested in studying vector fields F where $F = \nabla f$, perhaps to use the fundamental theorem of line integrals to not do a regular line integral, then we often need to identify such an f.

Theorem 8.6: Undoing the Gradient

If we have that $F = (F_1, F_2, F_3)$ is irrotational, a scalar potential f can be determined as

$$f(x, y, z) = \int_0^x F_1(t, 0, 0)\, dt + \int_0^y F_2(x, t, 0)\, dt + \int_0^z F_3(x, y, t)\, dt + C.$$

Example 8.6: MT Section 8.3 Problem 1

Determine which of the following vector fields F in the plane is the gradient of a scalar function f. If such an f exists, find it.

(a) $F(x, y) = x\mathbf{i} + y\mathbf{j}$
(b) $F(x, y) = xy\mathbf{i} + xy\mathbf{j}$
(c) $F(x, y) = (x^2 + y^2)\mathbf{i} + 2xy\mathbf{j}$

I'll compute all the curls first.

(a) $\operatorname{curl} F = 0$

(b) $\operatorname{curl} F = y - x$

(c) $\operatorname{curl} F = 2x - 2x = 0$

So the vector fields (a), and (c) are the gradient of a scalar function f.

By inspection, (a) has scalar potential

$$f(x, y) = \frac{1}{2}x^2 + \frac{1}{2}y^2 + C$$

and (c) has scalar potential

$$f(x, y) = \frac{1}{3}x^3 + xy^2 + C$$

Example 8.7: MT Section 8.3 Problem 7

Let

$$F(x, y, z) = (2xyz + \sin x)\mathbf{i} + x^2z\mathbf{j} + x^2y\mathbf{k}.$$

Find a function f such that $F = \nabla f$.

$$f(x, y, z) = \int_0^x 2(t)(0)(0) + \sin t\, dt + \int_0^y x^2(0)\, dt + \int_0^z x^2y\, dt$$

$$\begin{aligned}
&= \int_0^x \sin t\,dt + \int_0^z x^2 y\,dt \\
&= [-\cos t]_{t=0}^{t=x} + \left[x^2 y\right]_{t=0}^{t=z} \\
&= -\cos(x) - 1 + x^2 yz + C.
\end{aligned}$$

Example 8.8: MT Section 8.3 Problem 13

Let

$$F(x, y, z) = (e^x \sin y)\mathbf{i} + (e^x \cos y)\mathbf{j} + z^2\mathbf{k}.$$

Evaluate the integral

$$\int_c F \cdot ds$$

where

$$c(t) = (\sqrt{t}, t^3, \exp\sqrt{t}), \quad 0 \le t \le 1.$$

I have a feeling this might be conservative, so let's check the curl.

$$\begin{aligned}
\operatorname{curl} F &= (0 - 0, 0 - 0, -e^x \sin y + e^x \sin y) \\
&= (0, 0, 0)
\end{aligned}$$

so the vector field is irrotational and hence conservative. The scalar potential is given as

$$\begin{aligned}
f(x, y, z) &= \int_0^x e^t \sin 0\,dt + \int_0^y e^x \cos t\,dt + \int_0^z t^2\,dt + C \\
&= \int_0^y e^x \cos t\,dt + \int_0^z t^2\,dt + C \\
&= e^x \left[\sin t\right]_{t=0}^{t=y} + \frac{1}{3}\left[t^3\right]_{t=0}^{z} + C \\
&= e^x \sin y + \frac{1}{3} z^3 + C
\end{aligned}$$

The end points are $c(0) = (0, 0, 1)$ and $c(1) = (1, 1, e)$. Then $f(c(1)) = e \sin 1 + (1/3)e^3 + C$ and $F(c(0)) = 1/3 + C$ so

$$\int_c F \cdot ds = f(c(1)) - f(c(0)) = e \sin 1 + \frac{1}{3}e^3 + C - \frac{1}{3} - C = e \sin 1 + \frac{1}{3}e^3 - \frac{1}{3}.$$

8.3.2 Divergence Free Fields May be Curl Fields

Theorem 8.7: Converse of Curl Fields are Divergence Free

If F is a C^1 vector field on *all* of $\mathbb{R}^3$ with $\operatorname{div} F = 0$, then there exists a C^1 vector field G with

$$F = \operatorname{curl} G.$$

Example 8.9: MT Section 8.3 Problem 3

For each of the following vector fields F, determine (i) if there exists a function g such that $\nabla g = F$, and (ii) if there exists a vector field G such that $\operatorname{curl} G = F$. (It is not necessary to find g or G.)

(a) $F(x, y, z) = (4xz - x, -4yz, z - 2y)$
(b) $F(x, y, z) = (e^x \sin y, e^x \cos y, z^2)$
(c) $F(x, y, z) = (\log(z^2 + 1) + y^2, 2xy, (2xz)/(z^2 + 1))$
(d) $F(x, y, z) = (x^2 + x \sin z, y \cos z - 2xy, \cos z + \sin z)$.

(a) $\operatorname{curl} F = (-2 + 4y, \ldots)$ already the curl is nonzero, so it cannot be a gradient field.

$\operatorname{div} F = 4z - 1 - 4z + 1 = 0$ so it can be a curl field since F does not have an exception point.

(b) $\operatorname{curl} F = (0 - 0, 0 - 0 - e^x \sin y + e^x \sin y) = (0, 0, 0), \ldots$, so F is a gradient field.

$\operatorname{div} F = e^x \sin y - e^x \sin y + 2z = 2z$ which is not zero so F can't be a curl field.

(c) $\operatorname{curl} F = (0 - 0, \frac{2z}{z^2+1}) - \frac{2z}{z^2+1}, 2y - 2y) = (0, 0, 0)$, so F is a gradient field.

$\operatorname{div} F = 0 + 2x + \frac{(z^+1)2x - 2xz(2z+1)}{(z^2+1)^2}$ which is not zero, so F can't be a curl field.

(d) $\operatorname{curl} F = (y \sin(z), \ldots)$ already the curl is nonzero, so it cannot be a gradient field.

$\operatorname{div} F = 2x + \sin z + \cos z - 2x - \sin z + \cos z = 2 \cos z$ which is nonzero, so it cannot be a curl field.

8.4 "Gauss' Theorem" or the Divergence Theorem (MT 8.4)

Gauss' theorem relates the total action of a volumetric integral over a region to the the endpoints that are the enclosing boundary surface of the divergence.

Theorem 8.8: Gauss' Divergence Theorem

Let W be a region in space that can be broken up into symmetric elementary regions. Denote by ∂W the oriented closed surface that bounds W. Let F be a smooth vector field defined on W. Then

$$\iiint_W \operatorname{div} F \, dV = \iint_{\partial W} F \cdot dS.$$

Example 8.10: MT Section 8.4 Problem 1

Verify the divergence theorem for

$$W = [0,1] \times [0,1] \times [0,1]$$

boundary ∂W oriented outward, and vector field

$$F = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}.$$

Let's do the volumetric space integral first. $\operatorname{div} F = 1 + 1 + 1 = 3$, and the total volume is 1, so the surface integral is 3.

The surface integral here is not the most pleasant, as must consider the 6 faces of the cube. I'll table this, too:

Face	Parameterization	Normal vector	$F(c(t))$	$F(c(t)) \cdot \mathbf{n}$
(1)	$(u, v, 0)$	$(0, 0, -1)$	$(u, v, 0)$	0
(2)	$(u, v, 1)$	$(0, 0, 1)$	$(u, v, 1)$	1
(3)	$(u, 0, v)$	$(0, -1, 0)$	$(u, 0, v)$	0
(4)	$(u, 1, v)$	$(0, 1, 0)$	$(u, 1, v)$	1
(5)	$(0, u, v)$	$(-1, 0, 0)$	$(0, u, v)$	0
(6)	$(1, u, v)$	$(1, 0, 0)$	$(1, u, v)$	1

Summing, the total surface integral is 3, which agrees.

Example 8.11: MT Section 8.4 Problem 2

Verify the divergence theorem for

$$W = [0,1] \times [0,1] \times [0,1]$$

boundary ∂W oriented outward, and vector field

$$F = zy\mathbf{i} + xz\mathbf{j} + xy\mathbf{k}.$$

The divergence of F is $\operatorname{div} F = 0$, so the volumetric integral of the region of the divergence is 0.

Table again!

Face	Parameterization	Normal vector	$F(c(t))$	$F(c(t)) \cdot \mathbf{n}$
(1)	$(u, v, 0)$	$(0, 0, -1)$	$(0, 0, uv)$	$-uv$
(2)	$(u, v, 1)$	$(0, 0, 1)$	(v, u, uv)	uv
(3)	$(u, 0, v)$	$(0, -1, 0)$	$(0, uv, 0)$	$-uv$
(4)	$(u, 1, v)$	$(0, 1, 0)$	(v, uv, u)	uv
(5)	$(0, u, v)$	$(-1, 0, 0)$	$(uv, 0, 0)$	$-uv$
(6)	$(1, u, v)$	$(1, 0, 0)$	(uv, v, u)	uv

All of these have bounds $(u, v) \in [0,1] \times [0,1]$, so all of their bounds are the same, so we can integrate all of them at once as a singular integrand of six terms. As we have three uvs and three $-uv$s, they all cancel and the integral is 0.

Example 8.12: MT Section 8.4 Problem 3

Verify the divergence theorem for

$$W = \{(x, y, z) \mid x^2 + y^2 + z^2 \leq 1\}$$

boundary ∂W oriented outward, and vector field

$$F = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}.$$

$\operatorname{div} F = 1 + 1 + 1 = 3$, and the region, the unit sphere, has area $(4/3)\pi$, so the volumetric integral of the region is 4π.

For the surface integral, we use the parameterization

$$c(t) = (\sin u \cos v, \sin u \sin v, \cos u)$$

which has surface normal vector

$$T_u \times T_v = \sin u(\sin u \cos v, \sin u \sin v, \cos u).$$

and $F(c(t)) = c(t)$. Integrating,

$$\begin{aligned} \int_0^{\pi} \int_0^{2\pi} \sin u \, dv \, du &= 2\pi \int_0^{\pi} \sin u \, du \\ &= 2\pi[-\cos u]_{u=0}^{u=\pi} \\ &= 4\pi \end{aligned}$$

which agrees.

Example 8.13: MT Section 8.4 Problem 5

Use the divergence theorem to calculate the flux of

$$F = (x - y)\mathbf{i} + (y - z)\mathbf{j} + (z - x)\mathbf{k}$$

out of the unit sphere.

The divergence is $\operatorname{div} F = 1 + 1 + 1 = 3$. Integrating 3 over the unit sphere is then 4π.

Example 8.14: MT Section 8.4 Problem 9

Let
$$F = y\mathbf{i} + z\mathbf{j} + xz\mathbf{k}.$$
Evaluate $\iint_{\partial W} F \cdot dS$ for each of the following regions W:
(a) $x^2 + y^2 \le z \le 1$
(b) $x^2 + y^2 \le z \le 1$ and $x \ge 0$
(c) $x^2 + y^2 \le z \le 1$ and $x \le 0$.

We'll instead use the divergence theorem and do the volumetric integral instead.

$$\operatorname{div} F = 0 + 0 + x = x.$$

(a) This is symmetric across x, so the total integral is 0.

For (b) and (c), we actually have to do the integral, so we integrate using cylindrical coordinates. The $x^2 + y^2 \le z$ in cylindrical coordinates is $r^2 \le z$, or $0 \le r \le \sqrt{z}$. Finally, $x = r\cos\theta$.

(b) Since $x \ge 0$, $-\pi/2 \le \theta \le \pi/2$. Integrating,

$$\begin{aligned}
\int_{-\pi/2}^{\pi/2}\int_0^1\int_0^{\sqrt{z}} r\cos\theta \cdot r\,dr\,dz\,d\theta &= \int_{-\pi/2}^{\pi/2}\int_0^1\int_0^{\sqrt{z}} r^2\cos\theta\,dr\,dz\,d\theta \\
&= \frac{1}{3}\int_{-\pi/2}^{\pi/2}\int_0^1 \cos\theta\left[r^3\right]_{r=0}^{r=\sqrt{z}} dz\,d\theta \\
&= \frac{1}{3}\int_{-\pi/2}^{\pi/2}\int_0^1 z^{3/2}\cos\theta\,dz\,d\theta \\
&= \frac{2}{15}\int_{-\pi/2}^{\pi/2}\cos\theta\,d\theta \\
&= \frac{2}{15}\left[\sin\theta\right]_{\theta=-\pi/2}^{\theta=\pi/2} \\
&= \frac{4}{15}.
\end{aligned}$$

(c) The bounds are the same, but in reverse, so it's

$$\begin{aligned}
\frac{2}{15}\int_{\pi/2}^{-\pi/2}\cos\theta\,d\theta &= -\frac{2}{15}\int_{-\pi/2}^{\pi/2}\cos\theta\,d\theta \\
&= -\frac{4}{15}.
\end{aligned}$$

Example 8.15: MT Section 8.4 Problem 16

Evaluate the surface integral $\iint_{\partial S} F \cdot \mathbf{n}\, dA$ where

$$F = \mathbf{i} + \mathbf{j} + z(x^2 + y^2)^2\mathbf{k}$$

and ∂S is the surface of the cylinder $x^2 + y^2 \leq 1$, $0 \leq z \leq 1$.

We'll use the divergence theorem. $\operatorname{div} F = (x^2 + y^2)^2$. Using cylindrical coordinates, $x^2 + y^2 = r^2$, so we are integrating r^4. Integrating,

$$\begin{aligned}
\int_0^{2\pi} \int_0^1 \int_0^1 r^4 \cdot r\, dz\, dr\, d\theta &= 2\pi \int_0^1 \int_0^1 r^5\, dz\, dr \\
&= 2\pi r^5\, dr \\
&= \frac{1}{3}\pi \left[r^6\right]_{r=0}^{r=1} \\
&= \frac{\pi}{3}.
\end{aligned}$$