# MA 1024 Calculus IV Lecture Notes

Binan Gu

May 6, 2025

# Contents

# Part I

# Differential Calculus in Multiple Dimensions

## 1 Lecture I: Introduction to Multivariable Functions

The first lecture is meant to introduce the notion of functions of several variables at a high level. There is more intuition involved than technical details. This course will use your spatial imagination which ultimately involves your experience with geometry.

### 1.1 Notations

To shorten terms in written format, we adhere to the following notation standard. They should follow from your previous calculus classes.

| Symbol | Meaning |
|---|---|
| $\{\ldots, \ldots, \ldots\}$ | set of numbers |
| $\{\cdots \mid \cdots\}$ | set of numbers such that |
| $\in$ | is an element of |
| $\notin$ | is not an element of |
| $\subset$ | is a proper subset of |
| $\subseteq$ | is a subset of |
| $\bigcup$ | union |
| $\bigcap$ | intersection |
| $\mathbb{R}$ | all real numbers |
| $\mathbb{R}^+$ | all positive real numbers |
| $\mathbb{R}^n = \mathbb{R} \times \cdots \times \mathbb{R}$ | $n$-tuples of elements of $\mathbb{R}$ |
| $\mathbb{Z}$ | all whole numbers |
| $\mathbb{N} = \mathbb{Z}^+$ | all natural numbers (positive whole numbers) |

For example, we say $(a, b) \in \mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ if $a \in \mathbb{R}$ and $b \in \mathbb{R}$. The "$\times$" here is not multiplication, but simply introducing an **ordered pair** (or $n$-tuple). We can also say, for example, that

$$\left(\sqrt{2}, 7\right) \in \mathbb{R} \times \mathbb{N}.$$

## 1.2 Domain, Range and Examples

You walked off Calculus III with an introduction to vectors. You might be able to answer the following question.

**Question 1.1.** What is a vector?

"Something" with a **direction** and a **magnitude**. Sure.

**Question 1.2.** What is direction? Does this require the notion of "points" and the space in which they lie in?

**Question 1.3.** How do you characterize a photo in mathematical terms?

Each pixel in a 2D photo can be viewed as points in the $xy$-plane plane. Let's pick an arbitrary point $(x, y)$, also known as an **ordered pair**. Each pixel is associated with a color made from various weights of RGB (the fundamental colors). For $(x, y)$, we can associate an RGB reading of $(R, G, B)$. Thus, to fully characterize one pixel of a photo, we have the 5-tuple

$$(x, y, R, G, B) .$$

This 5-tuple can be considered as a vector if we define an origin $O = (0, 0, 0, 0, 0)$ such that each pixel now can be represented as a vector

$$v = \langle x, y, R, G, B \rangle .$$

So, now, what is "direction" here? It is certainly more abstract than "north", "south", "east" and "west", or "up" and "down".

But, there is still no "function" here. Why do we even need "function" anyways?

**Definition 1.1.** *A function relates two sets $X$ and $Y$ (of mathematical objects preferably, but they can be anything), with a rule that each element in $Y$ corresponds to exactly one element in $X$. This rule is the well-known "vertical line test" (but now, in higher dimensions, be careful of what you mean by "vertical" and "line").*

| $X$ | $Y$ |
|---|---|
| *domain* | *codomain* |

*In this course, I will not use the word "range" for $Y$ because it leads to two possible interpretations. Your textbook still uses "ranges", which means the set of numbers in $Y$ that the function actually takes value in.*

*We write*

$$f : X \to Y.$$

*A quick example differentiating domain, codomain and range is the following.*

$$f(x) = 2x + 1, \quad x = 1, 2, 3, 4.$$

*The domain is the set of what can be plugged into $x$, here the set $\{1, 2, 3, 4\}$.*

*The codomain is pre-determined. We can say that the codomain is all whole numbers because it seems our function $2x + 1$ where $x = 1, 2, 3, 4$ can only give us whole numbers. Thus, the codomain may be $Y = \mathbb{Z}$, that is, the set of **possible** outputs.*

*The range is what $f$ **actually** outputs. Here*

$$R(f) = \{3, 5, 7, 9\}.$$

*(in fact, the codomain may be $\{1, 2, 3, \ldots, 8, 9, 10\}$, the set of whole numbers between $0$ and $10$, or all positive integers). **The codomain always contains the range**.*

**Definition 1.2.** *A function of $1$ variable relates $X$ and $Y$ where each element in $X$ is a single value ($1$-tuple).*

**Definition 1.3.** *A function of $n$ variable relates $X$ and $Y$ where each element in $X$ is an $n$-tuple of values.*

In fact, we are used to seeing functions of several variables.

**Example 1.1.** Consider the area of a rectangle with height $h$ and length $l$. Then,

$$A(h, l) = hl.$$

This means,

$$f : X \to Y$$

where $X = [0, \infty) \times [0, \infty)$ and $Y = [0, \infty)$. Sometimes, we also write $[0, \infty) = \mathbb{R}^+$ in short, that is, positive real numbers.

In this course, we work with $X = \mathbb{R}^n$ where $n = 2, 3$, most of the time. $Y = \mathbb{R}$ most of the time. Below, we define a function when $Y = \mathbb{R}^n$ for $n \geq 2$. All you really need to think about is a relationship between more than three variables, one dependent and two independent. Here are more examples.

**Example 1.2.** Given a point $(x, y) \in \mathbb{R}^2$, its distance to the origin is given by

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

Domain of the function $d$ is

$$\operatorname{dom}(d) = \mathbb{R} \times \mathbb{R} = \mathbb{R}^2$$

and the range is

$$\mathcal{R}(d) = [0, \infty)$$

because the output of a square root is always nonnegative.

**Example 1.3.** The Ideal Gas Law, with one variable isolated, e.g., consider a box with nonzero volume $V$, then the pressure satisfies

$$P(n, V, T) = \frac{nRT}{V}$$

where $n$ is the number of particles, $R$ the avogadro's number (constant), $T$ temperature, and $V$ volume. This function has domain

$$\mathrm{dom}(d) = \mathbb{Z}^{+} \times \mathbb{R} \times \mathbb{R}$$

where the range is

$$\mathcal{R}(d) = [0, \infty)$$

(though there is a notion called negative pressure).

Recall that Calculus III involves functions that map from one input to an output of multiple entries (not multiple outputs), i.e.

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

for example, a parametric curve

$$\vec{r}(t) = \vec{v}_0 + \vec{v}t$$

or

$$\vec{r}(t) = x(t)\,\hat{i} + y(t)\,\hat{j} + z(t)\,\hat{k} = (x(t), y(t), z(t)).$$

This course involves functions such as

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

which describes **surfaces**. We want to understand the **changes** in the surfaces, and of course, all calculus I concepts, such as graph of a function, limits, continuity, derivatives, extreme values (optimization) and integration. As you will see, the added dimension makes things a lot more interesting, and a lot more complex.

## 1.3 Back to High School

As we encounter more complicated functions of several variables, we ought to remember the domain and range of simpler functions

| Function | Typical Domain | Range |
|---|---|---|
| $\sin(x), \cos(x)$ | $\mathbb{R}, [0, 2\pi]$ | $[-1, 1]$ |
| $\tan(x)$ | $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ | $(-\infty, \infty)$ or $\mathbb{R}$ |
| $\ln(x)$ | $(0, \infty)$ | $(-\infty, \infty)$ or $\mathbb{R}$ |
| $\sqrt{x}$ | $[0, \infty)$ | $[0, \infty)$ |
| $x^{1/3}$ | $\mathbb{R}$ | $\mathbb{R}$ |
| $a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$ | $\mathbb{R}$ | $\mathbb{R}$ |

One must note that the domain in practice is usually chosen by the user. If a question asks for the domain of a function $f$, it is asking for the **largest** region whose values renders $f$ well-defined, often referred to as the **natural domain** of $f$. In this case, the domain of $\sin(x)$ would be $\mathbb{R}$, for example.

## 1.4 A Calculus IV Example

Since we now have at least two inputs, our domain is no longer an interval like in 1D functions. Instead, it is a subset of the $xy$-plane plane, i.e., $\mathbb{R}^2$.

**Example 1.4.** Find the domain and range of $f(x, y) = \sqrt{y - x^3}$.

Clearly, we must require $y - x^3 \geq 0$ since what goes under the square root must be nonnegative. Thus, the domain of $f$ is any ordered pair $(x, y)$ that satisfies $y \geq x^3$. In the $xy$-plane plane, these pairs form all the points $(x, y)$ that lies above the cubic function $y = x^3$ (including the curve itself due to "$\geq$").

## 1.5 The Cartesian Coordinate System

Consider the $xy$-plane plane. When we set $x = 0$, we are precisely talking about the vertical line $x = 0$, which we called the $y$-axis. When we set $y = 0$, we are then talking about the $x$-axis. Now, let's move up one dimension into the $xyz$-octant. Think of a cube where you slice three times, each at the midpoint of each dimension (which produces 8 little cubes, hence OCTANT). How do we characterize the coordinate axes now?

You may try setting $x = 0$ again. But now, $y$ and $z$ are still "freely" available, which means you are describing precisely the $yz$-plane. In fact, this procedure gives us the description of the **principal coordinate planes**, see Figure 1.

What about the principal axes then? Notice that you must determine **two** out of

| | $xyz$-octant | surface $z = f(x, y)$ |
|---|---|---|
| $x = 0$ | $yz$-plane | cross-section in the $yz$-plane |
| $y = 0$ | $xz$-plane | cross-section in the $xz$-plane |
| $z = 0$ | $xy$-plane (or birds' eye view) | cross-section in the $xy$-plane or the 0-level set |

Figure 1: Principal Coordinate Planes and surface cross-sections.

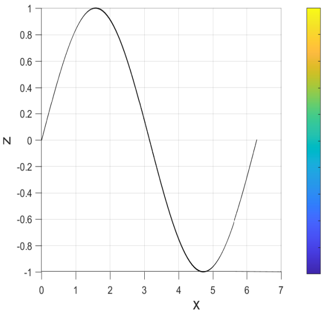
the three possible coordinates to narrow it down. Therefore, we must have

| | Principal Axis |
|---|---|
| $y = z = 0$ | $x$-axis |
| $x = z = 0$ | $y$-axis |
| $x = y = 0$ | $z$-axis |

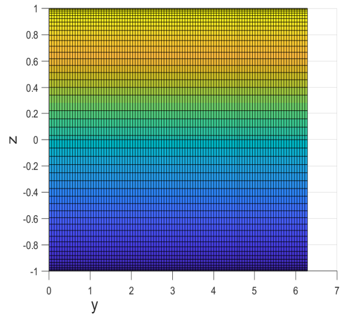
Now, if you were observant about how we determine the expression for the principal planes and axes, you might notice the following: any point characterized by the coordinate $(x, y, z)$ lives in $\mathbb{R}^3$, that is, in three dimensions. If we specify one of the three, then we reduce it to **two** dimensions, which is then a plane. If we specify two of the three, then we reduce it to **one** dimension, which is then a line.

In a 1D function $y = f(x)$, when we set $y = 0$, then we say the values satisfying $f(x) = 0$ are the **$x$-intercepts**; similarly, when we set $x = 0$, the corresponding $y$-value, i.e. $y = f(0)$ is called the **$y$-intercepts**. What are the equivalent notions when we consider a 2D function $z = f(x, y)$? See Figure 1 for the breakdown.

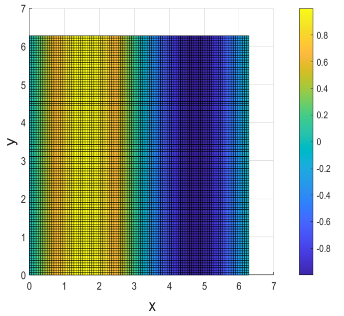
**Example 1.5.** Consider the function $z = f(x, y) = \sin(x)$, $0 \leq x \leq 2\pi$ and $0 \leq y \leq 2\pi$. Note that the function does NOT depend on $y$, so the only variation we should see in its 3D plot is in the $xz$-plane. See Figure 2 for a demonstration of the cross-sections.



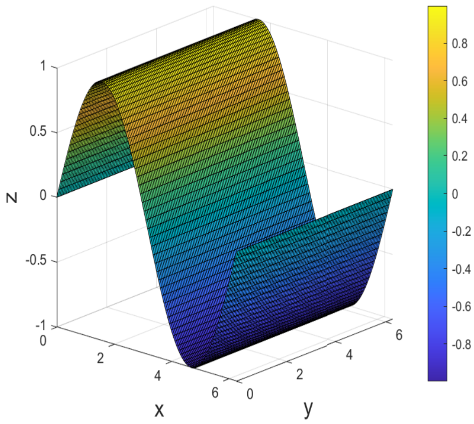

(a) $xz$ cross-section




(b) $yz$ cross-section




(c) $xy$ cross-section




(d) Full view

Figure 2: The three principal cross-sections and a full view of the surface $z = \sin(x)$. Note that in (a), we retrieve the basic relationship $z = \sin(x)$, just like in 1D; in (b), the color schemes tells us that for each $y$, $z$ is unaffected, as expected since there is no explicit $y$ dependence in the function; and in (c), we see the birds' eye view of the surface, which is precisely one period of $\sin(x)$, from its crest (yellow) to its trough (blue). In (d), we now can fully justify the surface we see.

# 2 Lecture II: Graph of Multivariable Functions and Level Sets

## 2.1 Graph of Multivariable Functions

See notes from last class, and play with the provided MATLAB codes for other functions.

## 2.2 Level Sets

One of the reasons we identify domain and range of a function of several variables is to plot its graph as a visual aid. So far, even though we are equipped with the domain and range of a function, we are still unable to graph it with much details.

In one dimension, we simply plug in values and trace out the shape of the graph. Can we use basic knowledge of the graph of functions of one variable to help us determine the looks of functions of several variables? Meanwhile, we certainly realize that plugging in values to a function

$$v = f(x, y, z, w)$$

is not a very smart move since we have too many points to consider.

The cross-sections we just saw are indeed helpful, since in each principal coordinate plane, we would be able to see the curve outlining the surface when one variable is held fixed, in particular, the $xz$-cross-section and $yz$-cross-section. The $xy$-cross-section, on the other hand, is more handy. Since $z = f(x, y)$ is a function, it must pass the "vertical plane" test. Therefore, specifying any two $z$-values must lead to two non-intersecting curves viewed in the $xy$-plane. This would be a great way to visualize what the surface looks like in the vertical direction, following the shape of these non-intersecting curves.

We introduce the **level curve**, that is, the set of points in the plane where a function $f(x, y)$ has a **constant** value (remember the contour maps in the brochures of a hiking trail?),

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

The set of all points $(x, y, f(x, y))$ in space, for $(x, y)$ in the domain of $f$, is called the **graph** of $f$. The graph of $f$ is also called the **surface**

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

In essence, for a function of two variables, the **level curves** are what the function looks like in birds-eye-view. These curves lie on the $xy$-plane.

**Example 2.1.** Describe the level curves of the function $f(x,y) = 100 - x^2 - y^2$.

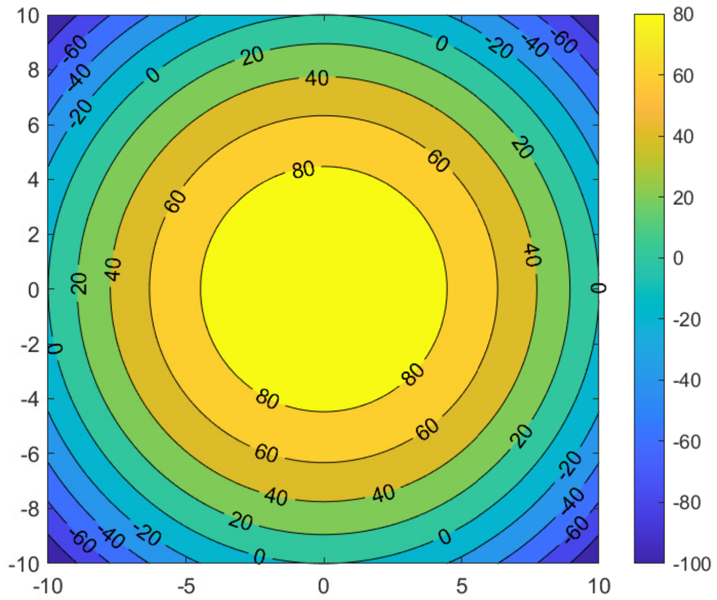


Figure 3: Level curves of $f$. Note that they are all circles of various radius, depending on the level.

Before one finds the level curve, one **must** identify the domain and range of $f$. Here, the domain of all of $\mathbb{R}^2$, while the range is $(-\infty, 100]$ since we note that the function value cannot exceed 100.

Thus, for a level set of $f$, we just need to set

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

where $c \in (-\infty, 100]$. More precisely,

$$100 - x^2 - y^2 = c \implies x^2 + y^2 = 100 - c,$$

that is, **an equation of a circle** of radius $\sqrt{100-c}$. Picking our favourite perfect squares, such as $c = 19, 36, 51$, etc., we can sketch the level curves of $f$, concentric circles with various radii (see figure).

If one selects $c > 100$, then we say the level set corresponding to this level is the *empty set*.

Note further that if we set $x$ or $y$ to be a fixed constant, then in the $yz$- or $xz$-planes, you will see an inverted parabola.

**Definition 2.1.** *We say that the curve in space in which the plane $z = c$ cuts a surface $z = f(x, y)$ is called the **contour curve**. Using the previous example, let $c = 19$. The level curve is the curve described by*

$$x^2 + y^2 = 81, \quad z = 0,$$

*that is, this curve lives in the $xy$-plane. The contour curve with $c = 19$ is described by*

$$x^2 + y^2 = 81, \quad z = 19,$$

*elevated from the level curve to the exact height of the function value. This distinction comes up in maps – the elevation lines you see in maps are actually called **contour lines** because they are meant to also represent specific heights. Calling them **level curves** would mean they are at sea level.*

**Definition 2.2.** *The set of points $(x, y, z)$ in space where a function of three independent variables has a constant value $f(x, y, z) = c$ is called a **level surface** of $f$.*

**Remark 2.1.** *In fact, though as it appears that this relies on solving*

$$f(x, y, z) = c$$

*which should result in*

$$z = g(x, y; c)$$

*namely, a surface described by $g$ with parameter $c$, it is not guaranteed that $f(x, y, z) = c$ is solvable for all $(x, y, z)$ and $c$. We come back to this subtlety in a later section in this course.*

## 2.3 Limits of Functions of Single Variable: A Review

Recall that we say a function $f(x)$ approaches the **limit** $L$ as $x$ approaches $x_0$, written as

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

In simpler terms,

$$x, x_0 \text{ not too far} \implies f(x), L \text{ not too far}.$$

Note that this **limiting value** $L$ is NOT necessarily the **function value** $f(x_0)$, even though $x$ is tending towards $x_0$. See the following example:

**Example 2.2.** Consider the function

$$f(x) = \begin{cases} 0, & x \neq 1; \\ 1, & x = 1. \end{cases}$$

The graph of this function is the entire $x$-axis but a hole at the point $(1, 0)$, because the function value at $x = 1$ is $f(1) = 1$. This function constitutes a jump discontinuity. However, we can still compute the left and right limit of this function,

$$\lim_{x \to 1^-} f(x) = 0$$
$$\lim_{x \to 1^+} f(x) = 0$$

because as $x$ goes close to 1 (but not exactly at 1), the function value is very close (in fact, constant) to 0 – the function value is NOT going anywhere near a height of 1, unless we hit $x = 1$ exactly. Therefore, the limit of this function as $x \to 1$ is 0, different from the **function value** $f(1) = 1$.

**Definition 2.3.** *A function $f(x)$ is said to be continuous at $x = x_0$ when*

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

*Note that a priori, we need the LHS to exist first, and then confirm that the* ***limiting value*** *and the* ***function value*** *coincide. They are not always the same the begin with (per example).*

**Remark 2.2.** *This is to say that we CANNOT just plug in the value $x_0$ to find the limit of $f(x)$ as $x \to x_0$. Plugging in automatically assumes that the function is continuous, which is NOT always the case.*

*Common continuous functions are: $\sin(x)$, $\cos(x)$, polynomials, rational functions $\frac{p(x)}{q(x)}$ where $p$ and $q$ are both polynomials, and $q(x) \neq 0$, $e^x$, $\ln(x)$, etc.*

**Question 2.1.** Do we still have just two ways of approaching the point of interest when the domain is now a region instead of an interval? If not, what are the new directions? Is it still **computationally feasible** to abide by the "left limit equals right limit" rule?

# 3 Lecture III: Limits and Continuity of Multivariable Functions

## 3.1 Limits of Functions of Several Variable: Similar Stuff

In higher dimensions, similar behaviours may happen as well. We just need to re-quantify what it means for $(x, y)$ to be approaching a point $(x_0, y_0)$ in the domain of $f(x, y)$.

**Definition 3.1.** *We say that a function $f(x, y)$ approaches the **limit** $L$ as $(x, y)$ approaches $(x_0, y_0)$ (in the xy-plane), and write*

$$\lim_{(x,y)\to(x_0,y_0)} f(x, y) = L,$$

*if, for every number $\epsilon > 0$, there exists a corresponding $\delta > 0$ such that for all $(x, y)$ in the domain of $f$,*

$$|f(x, y) - L| < \epsilon$$

*whenever*

$$0 < \sqrt{(x - x_0)^2 + (y - y_0)^2} < \delta.$$

*Note that $\sqrt{(x - x_0)^2 + (y - y_0)^2}$ is no other than the (Euclidean) distance between the points $(x, y)$ and $(x_0, y_0)$. It is **one** metric that measures the closeness of points in the plane.*

*In other words, if the pair of points $(x, y)$ and $(x_0, y_0)$ are pretty close to each other, then $f(x, y)$ and $L$ also aren't too far from each other. In fact, as the pair of points closes onto each other, the limiting value hits exactly $L$ (one must note that this limiting value may **NOT** be the function value $f(x_0, y_0)$).*

**Example 3.1.** (Optional) Show that $\lim_{(x,y)\to(x_0,y_0)} f(x, y) = x_0$ where $f(x, y) = x$.

We need to show that for any distance (of our choice) between $f(x, y)$ and $L$, we can always conjure up a number $\delta$ so that

$$0 < \sqrt{(x - x_0)^2 + (y - y_0)^2} < \delta \implies |f(x, y) - L| < \epsilon$$

where we identify that $L = x_0$ and $f(x, y) = x$. We choose an $\epsilon > 0$ (know its value and use everywhere we want). We ask, what should $\delta$ be so that the above statement is true? We **want** $|f(x, y) - L| = |x - x_0| < \epsilon$. Looking at what we want, we find that

$$|x - x_0| = \sqrt{(x - x_0)^2} \le \sqrt{(x - x_0)^2 + (y - y_0)^2} < \delta.$$

So, it would suffice to let $\delta = \epsilon$.

$$\sqrt{\left(x - x_0\right)^2 + \left(y - y_0\right)^2} < \epsilon \implies \left|f\left(x, y\right) - L\right| = \left|x - x_0\right| < \epsilon.$$

**Remark 3.1.** *Note that in $\epsilon$-$\delta$ proofs, we have already identified a candidate limiting value. But it is only a candidate. To qualify it as the true limit, one goes through the $\epsilon$-$\delta$ procedures by massaging $\left|f\left(x\right) - L\right|$ into something related to $\sqrt{\left(x - x_0\right)^2 + \left(y - y_0\right)^2} < \delta$, via algebraic methods.*

## 3.2 Rules of Limits

Given that

$$\lim_{(x,y)\to(x_0,y_0)} f\left(x, y\right) = L, \qquad \lim_{(x,y)\to(x_0,y_0)} g\left(x, y\right) = M$$

where $L, M \in \mathbb{R}$, we have

1. (Sum/Difference)

$$\lim_{(x,y)\to(x_0,y_0)} \left[f\left(x, y\right) \pm g\left(x, y\right)\right] = L \pm M.$$

2. (Constant multiple) For $k \in \mathbb{R}$,

$$\lim_{(x,y)\to(x_0,y_0)} kf\left(x, y\right) = kL.$$

3. (Product)

$$\lim_{(x,y)\to(x_0,y_0)} f\left(x, y\right) g\left(x, y\right) = LM.$$

4. (Quotient) If $M \neq 0$,

$$\lim_{(x,y)\to(x_0,y_0)} \frac{f\left(x, y\right)}{g\left(x, y\right)} = \frac{L}{M}.$$

5. (Power) Given $n \in \mathbb{N}$,

$$\lim_{(x,y)\to(x_0,y_0)} \left[f\left(x, y\right)\right]^n = L^n.$$

6. (Root) Given $n \in \mathbb{N}$, and if $n$ is even, we further require that $L > 0$,

$$\lim_{(x,y)\to(x_0,y_0)} \sqrt[n]{f\left(x, y\right)} = \sqrt[n]{L} = L^{\frac{1}{n}}.$$

One property of a limit not mentioned above is that the limit, if exists, is unique. This means that if a limit exists, then its value is determined independent of how $(x, y)$ approaches $(x_0, y_0)$. At the same time, this also gives rise to situations where before checking the existence of a limit, we should try approaching $(x_0, y_0)$ via different paths from $(x, y)$. Should we find two distinct limits, we say that the limit as $(x, y) \to (x_0, y_0)$ does not exist, since we expect it to be unique.

**Remark 3.2.** *In one dimension, there are only two ways to approach a point, that is, from the left and from the right. When left limit is equal to the right limit, we say the limit exists.* ***In two dimensions and above, there are infinitely many ways to approach a point, and there is no exhaustive methods to really consider them all.*** *We must resort to $\epsilon$-$\delta$ proofs where this pathing argument is avoided, once a target limit value is identified.*

**Example 3.2.** Find $\lim_{(x,y)\to(0,0)} \frac{x^2-xy}{\sqrt{x}-\sqrt{y}}$.

**Solution.** Direct "plug in" yields "0/0" which means we can't get our limits this way. Multiply top and bottom by the conjugate of the denominator.

$$\begin{aligned} \lim_{(x,y)\to(0,0)} \frac{x^2 - xy}{\sqrt{x} - \sqrt{y}} &= \lim_{(x,y)\to(0,0)} \frac{x^2 - xy}{\sqrt{x} - \sqrt{y}} \frac{\sqrt{x} + \sqrt{y}}{\sqrt{x} + \sqrt{y}} \\ &= \lim_{(x,y)\to(0,0)} \frac{x\,(x-y)\left(\sqrt{x} + \sqrt{y}\right)}{x - y} \end{aligned}$$

**The reason why we can now cancel $x - y$ from top and bottom is because the line $y = x$ is NOT in the domain of our function.** This is where **identifying the domain** becomes highly relevant in a calculation. If we don't check the domain, we actually won't justify this cancellation. The final limit is 0 after cancellation. See Figure 4 for the domain of the function of interest.

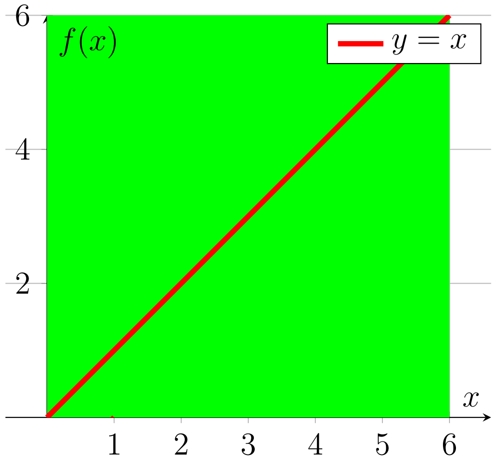



Figure 4: Domain of $f(x, y) = \frac{x^2 - xy}{\sqrt{x} - \sqrt{y}}$ in green. The red line $y = x$ is forbidden (not in the domain).

**Example 3.3.** Find $\lim_{(x,y)\to(1,1), x\neq y} \frac{x^2-y^2}{x-y}$.

**Solution.** Similar style as last one. One factors the numerator. We can cancel $x - y$ because we are given that $x \neq y$. Therefore, the final limit is 2. Fill in the details.

## 3.3 Two-Path Test for Non-Existence of a Limit (TPT)

We already know by now that for a limit to exist at a point $(x_0, y_0)$ for a multivariable function, unless the form of the function is really nice, we need to resort to the technical definition of a limit. However, because it is so difficult to obtain a limit, it is now easy to prove the **non-existence** of a limit. The idea is to select two distinct routes to approach the point of interest $(x_0, y_0)$. If one obtains two distinct limits along these routes, then the limit is said to not exist.

Let's begin with an example.

**Example 3.4.** Find the limit or prove that it does not exist.

$$\lim_{(x,y)\to(1,1)} \frac{xy^2 - 1}{y - 1}.$$

**Solution 3.1.** Always, the first step is to try plugging the target point in – here we get $\frac{0}{0}$ which means we can't get a limit this way. So, it seems we may have trouble near the target point. A good way to start is by doing the two-path test because if it goes through, there is no limit, which is a definitive answer.

Now, let's try approaching the target point $(1,1)$ via different paths. **One path** $y = x$ is a line that goes through this point. Thus,

$$\lim_{(x,y)\to(1,1),\quad y=x} \frac{xy^2-1}{y-1} = \lim_{x\to 1} \frac{x\left(x\right)^2-1}{x-1} = \lim_{x\to 1} \frac{x^3-1}{x-1} = \lim_{x\to 1} \frac{(x-1)\left(x^2+x+1\right)}{(x-1)} = 3.$$

**Another path** that goes through $(1,1)$ would be along the vertical line $x = 1$.

$$\lim_{(x,y)\to(1,1)} \frac{xy^2-1}{y-1} = \lim_{y\to 1,\quad x=1} \frac{xy^2-1}{y-1} = \lim_{y\to 1} \frac{y^2-1}{y-1} = \lim_{y\to 1} (y+1) = 2$$

which is not equal to 3, the limit when we approach $(1,1)$ along $y = x$. Obtaining different limits along **two distinct paths** implies that the limit DOES NOT EXIST.

**Remark 3.3.** *For the second path through* $(1,1)$*, you may also choose to approach it along the line* $y = -x+2$*. The limit computation then goes as*

$$\begin{aligned}
\lim_{(x,y)\to(1,1),\quad y=-x+2} \frac{xy^2-1}{y-1} &= \lim_{x\to 1} \frac{x\left(-x+2\right)^2-1}{-x+2-1} \\
&= \lim_{x\to 1} \frac{x\left(x^2-4x+4\right)-1}{1-x} \\
&= \lim_{x\to 1} \frac{x^3-4x^2+4x-1}{1-x} \\
&= -\lim_{x\to 1} \frac{x^3-1-4x\left(x-1\right)}{x-1} \\
&= -\lim_{x\to 1} \frac{(x-1)\left(x^2+x+1\right)-4x\left(x-1\right)}{x-1} \\
&= -\lim_{x\to 1} x^2+x+1-4x \\
&= 1 \\
&\neq 3
\end{aligned}$$

*Note that you cannot approach* $(1,1)$ *along the horizontal line* $y = 1$ *because the line* $(x,1)$ *(for any* $x$*) is NOT in the domain.*

The above procedure is known as the **Two-Path Test for Nonexistence of a Limit**. The idea is to approach $(x_0, y_0)$ (or other target point in $xy$-plane) with two different paths and check if we obtain different limits. If so, then there is NO limit due to uniqueness of limit.

More precisely, if a function $f(x,y)$ has different limits along two different paths in the domain of $f$ as $(x,y)$ approaches $(x_0,y_0)$, then $\lim_{(x,y)\to(x_0,y_0)} f(x,y)$ DOES NOT EXIST.

**Remark 3.4.** *Meanwhile, even though you check through the limits along all straight lines, it doesn't guarantee the existence of the limit at* $(x_0, y_0)$*. One may approach the point using other type of functions in the* $x$*-*$y$ *plane that go through* $(x_0, y_0)$*.* ***It is only a trick to prove NONEXISTENCE, but not a trick to prove existence.***

**Remark 3.5.** *Using all lines* $y = mx + b$ *to approach a point* $(x_0, y_0)$ *and find that all limits (exhausting all pairs of* $m$ *and* $b$*) are equal* ***DOES NOT*** *mean that the limit exists.*

**Remark 3.6.** *To prove existence of a limit in high dimensions, one almost always goes through the* $\epsilon$*-*$\delta$ *proof (as it is a rock solid definition), unless the function is obviously continuous, such as a polynomial, or rational function with non-vanishing denominator (and many other examples such as exponentials and sine and cosine without asymptotes).*

# 4 Lecture IV: Continuity and Partial Derivatives

In this lecture, we will explore the notion of continuity in dimension more than one by first drawing parallels to that from 1D calculus, and then noticing some subtle differences. Continuity hinges on your understanding of the existence of a limit, as the definition requires that the limit must exist a priori, and then the limiting value is equal to the function value. We will see that it is almost equally "easy" to NOT have continuity, by means of the two-path test for non-existence of a limit. You must make the logical connection that continuity requires existence of a limit.

## 4.1 Continuity

If a function is continuous, it means small fluctuations in inputs will result in small fluctuations in outputs. Now, with the help of limits, we can characterize these "fluctuations" already, with the help of $\epsilon$-$\delta$ definitions. However, we steer away from the technical definition and highlight the essence below.

A function is continuous at a point $(x_0, y_0)$ if

1. $f$ is defined at $(x_0, y_0)$, that is, $(x_0, y_0)$ lies in the domain of $f$ (another reason to know how to find domain of $f$).

2. $\lim_{(x,y)\to(x_0,y_0)} f(x, y)$ exists (this is where some functions may fail to achieve – check this to **disprove** continuity using TPT).

3. $\lim_{(x,y)\to(x_0,y_0)} f(x, y) = f(x_0, y_0)$.

**Example 4.1.** (**To disprove continuity**) Consider the function

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

Show that it is continuous at every point except the origin.

*Proof.* First, we note that the function obtains values of a rational function when $(x, y) \neq (0, 0)$. Rational functions are continuous in their domain. We now focus on the origin. Even though the function is defined at the origin, we may need to check the limit of the function as we approach $(0, 0)$ – it **may** not even exist to begin with.

**Key Logic**: to disprove continuity, we check if at least one of the three criteria fails. In this example, Criteria 1 and 3 are given. So, it will be Criteria 2 that we are trying to "break" using **the technique of Two-path Test for Non-existence of Limit** since Criteria 2 is a statement about existence of the limit.

We first need to identify some easy ways of approaching the origin. We employ a more general and robust method than the two-path test for non-existence (in fact, an upgrade). In the $xy$-plane, approaching the origin can be done via lines of various slopes, i.e.

$$y = mx, \quad m \in \mathbb{R}, \quad x \neq 0.$$

The reason why we kept this slope $m$ general rather than specified is that if the limiting value depends on $m$, that means we can select a slope $m$ to obtain a different limiting value, which automatically breaks the uniqueness of the limit, achieving the goal of proving non-existence.

Now, let us compute the limit if we travel on these lines. When we take a limit, we are away from the origin, so the procedure is simply "plug-in".

$$\begin{aligned} \lim_{(x,y)\to(0,0)} f(x,y) &= \lim_{(x,mx)\to(0,0)} \frac{2xy}{x^2+y^2} \\ &= \lim_{x\to 0} \frac{2mx^2}{x^2+m^2x^2} \\ &= \frac{2m}{1+m^2} \end{aligned}$$

where we can cancel $x^2$ since we are only approaching but not at $x = 0$. This means, along any straight lines, the function takes a constant limiting value $\frac{2m}{1+m^2}$, independent of $x, y$ but **dependent** on the slope $m$. This means, the function approaches **various limiting values** given different approach slopes, which breaks the uniqueness property of a limit. There is no single number we call the limit of $f$ as $(x, y)$ approaches the origin. Thus, the function does NOT have a limit at the origin, and thus certainly is NOT continuous at the origin. $\square$

**Remark 4.1.** *If limit does not exist at a point, we will NOT have continuity there since criterion **2** of continuity fails. Therefore, the Two-Path Test for Nonexistence of a Limit is also useful to show that a function is not continuous at certain points.*

**Remark 4.2.** *Now, sometimes, even after approaching the origin using the line $y = mx$, we end up with a limiting value that is also independent of $m$. Does this mean that the limit exists? NO. A Big NO. It just means the limit has the same for all lines approaching the origin, not all paths.*

## 4.2 Partial Derivatives

We will use geometric intuition from the procedures of graphing functions of several variables $z = f(x, y)$. In addition to the level set technique, we can also learn the graph of a function by looking at the cross-sections, i.e. $z = f(x, 0)$, that is, the

$xz$-plane, or $z = f(0, y)$, the $yz$-plane, in addition to $c = f(x, y)$ the level sets that live in the $xy$-plane.

Partial derivatives at a particular point of the function relate to the cross-sections quite a bit. We make the connection after seeing the definitions.

$$\begin{aligned} \frac{\partial f}{\partial x} \mid_{(x_0, y_0)} &= \lim_{h \to 0} \frac{f(x_0 + h, y_0) - f(x_0, y_0)}{h} \\ \frac{\partial f}{\partial y} \mid_{(x_0, y_0)} &= \lim_{h \to 0} \frac{f(x_0, y_0 + h) - f(x_0, y_0)}{h} \end{aligned}$$

Note that in the difference quotient, $y_0$ or $x_0$ is fixed respectively for $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$.

In fact, $\frac{\partial f}{\partial x}$ evaluated at $(x_0, y_0)$ is the ordinary derivative (against $x$) of the curve $f(x, y_0)$ with $y_0$ fixed, and then evaluated at $x = x_0$. More precisely, define $g(x) = f(x, y_0)$ and we find that

$$\frac{\partial f}{\partial x} \mid_{(x_0, y_0)} = \frac{d}{dx} g(x) \mid_{x = x_0} .$$

**Example 4.2.** Compute the partial derivatives of $f(x, y) = x^2 + y^2$ at $(x, y) = (1, 1)$, using the definition and the last observation to compare.

$$\begin{aligned} \frac{\partial f}{\partial x} \mid_{(x,y)=(1,1)} &= \lim_{h \to 0} \frac{f(1 + h, 1) - f(1, 1)}{h} \\ &= \lim_{h \to 0} \frac{(1 + h)^2 + 1^2 - (1^2 + 1^2)}{h} \\ &= \lim_{h \to 0} \frac{(h^2 + 2h + 1) + 1^2 - (1^2 + 1^2)}{h} \\ &= \lim_{h \to 0} h + 2 \\ &= 2. \end{aligned}$$

We check using the observation above: indeed, define $g(x) = f(x, 1) = x^2 + 1$, we have

$$\frac{\partial f}{\partial x} \mid_{(1,1)} = \frac{d}{dx} g(x) \mid_{x=1},$$

Carrying on, we have

$$\frac{d}{dx} g(x) \mid_{x=1} = 2x \mid_{x=1} = 2$$

Thus, $\frac{\partial f}{\partial x}(1, 1) = 2$ as it is from the limit definition.

In essence, whenever we take a partial derivative, we treat the other variable as a constant.

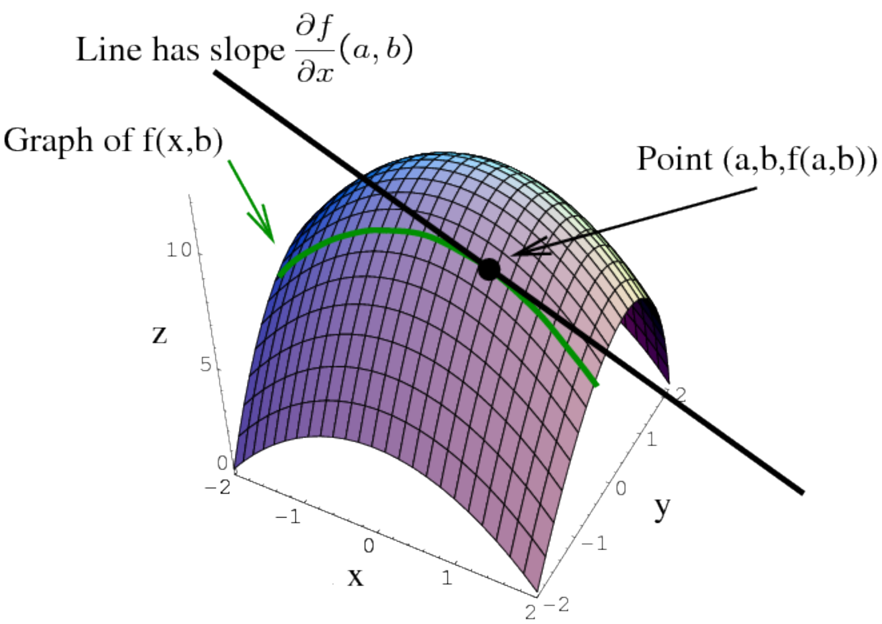
# 5 Lecture V: Examples and the Geometric Interpretation of Partial Derivatives

The partial derivative of $f(x, y)$ is the slope of the tangent in a very **specific** direction. For example, consider the partial derivative of $f$ against the variable $x$ at a particular point $(x_0, y_0)$. The intersection of the surface represented by $f$ and the vertical plane $y = y_0$ is the curve $g(x) = f(x, y_0)$. Now, are you able to take a derivative against $x$ for the function $g$? Of course. That's Calculus I stuff.

$$\frac{\partial f}{\partial x}(x_0, y_0) = g'(x_0).$$

In other words, you hold $y_0$ as if it is just another **constant**, and simply differentiate against $x$, as if there is no more other independent variables.

This description has a precise geometric interpretation. Consider the surface defined by $f(x, y) = 10 - x^2 - y^2$.




This surface is the function $f(x, y) = 10 - x^2 - y^2$. The tangent line (in black) to the curve (in green) defined by $g(x) = f(x, -1)$ at $x = 1$ has slope equal to the partial derivative

$$\frac{\partial f}{\partial x}(1, -1) = -2x \mid_{(x,y)=(1,-1)} = -2,$$

which is negative as you note where positive $x$ is – going towards the bottom corner. In other words, at $(x, y) = (1, -1)$, the instantaneous rate of change in the $x$-direction of the surface $f(x, y)$ is $-2$ unit/time.

Can you tell what the partial derivatives are at $(0, 0)$?

## 5.1 Examples: Partial Derivatives

Let's see a few more examples with other notations of partial derivatives.

The following typically applies when you are referring to the partial derivatives as **functions**.

$$\frac{\partial f}{\partial x}, \quad f_x$$

If we care about the **value** of the derivative at a specific point $(x_0, y_0)$, we write

$$\frac{\partial f}{\partial x}(x_0, y_0), \quad f_x(x_0, y_0)$$

**Takeaway 5.1.** When taking a partial derivative against one independent variable, treat **all other independent variables** as constants.

**Example 5.1.** Compute the partial derivative of $f(x, y) = y \sin(xy)$.

**Solution 5.1.**

$$\frac{\partial f}{\partial x} = y \cdot (y \cos(xy)) = y^2 \cos(xy).$$
$$\frac{\partial f}{\partial y} = \sin(xy) + y \cdot (x \cos(xy)) = \sin(xy) + xy \cos(xy).$$

**Example 5.2.** Find $f_x$ and $f_y$ as functions if

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

**Solution 5.2.**

$$\begin{aligned} f_x &= \frac{(y + \cos(x)) \frac{\partial}{\partial x}(2y) - 2y \frac{\partial}{\partial x}(y + \cos(x))}{(y + \cos(x))^2} \\ &= \frac{0 - 2y(-\sin(x))}{(y + \cos(x))^2} \\ &= \frac{2y \sin(x)}{(y + \cos(x))^2} \end{aligned}$$

$$
\begin{aligned}
f_y &= \frac{(y+\cos(x))\frac{\partial}{\partial y}(2y) - 2y\frac{\partial}{\partial y}(y+\cos(x))}{(y+\cos(x))^2} \\
&= \frac{2(y+\cos(x)) - 2y}{(y+\cos(x))^2} \\
&= \frac{2\cos(x)}{(y+\cos(x))^2}
\end{aligned}
$$

**Example 5.3.** (Implicit differentiation) Find $\frac{\partial z}{\partial x}$ when

$$yz - \ln z = x + y$$

where we know $z = z(x, y)$ is a function of two variables.

**Solution 5.3.** Taking $\frac{\partial}{\partial x}$ on both sides, by treating $z = z(x, y)$, we must have

$$y\frac{\partial z}{\partial x} - \frac{1}{z}\frac{\partial z}{\partial x} = 1 \implies \frac{\partial z}{\partial x} = \frac{1}{y - \frac{1}{z}} = \frac{z}{yz - 1}.$$

when $y \neq 1/z$.

## 5.2 Second-Order Partial Derivatives

$$
\begin{aligned}
\frac{\partial^2 f}{\partial x^2} &= \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right), \quad f_{xx} \\
\frac{\partial^2 f}{\partial y^2} &= \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right), \quad f_{yy}
\end{aligned}
$$

Mixed derivatives

$$
\begin{aligned}
\frac{\partial^2 f}{\partial x \partial y} &= \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right), \quad f_{yx} \\
\frac{\partial^2 f}{\partial y \partial x} &= \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right), \quad f_{xy}
\end{aligned}
$$

The above may seem confusing. Confuse no more, there is a nice theorem telling us under certain conditions, the mixed derivatives are equal to each other.

**Theorem 5.1.** *(Clairaut's Theorem) If $f(x, y)$ and its partial derivatives $f_x, f_y, f_{xy}$ and $f_{yx}$ are defined throughout an open region containing a point $(a, b)$ and are all continuous at $(a, b)$, then*

$$f_{xy}(a, b) = f_{yx}(a, b).$$

## 5.3 A Connection between Partial Derivatives and Level Curves

We have learned that the partial derivative computes the instantaneous rate of change of the elevation/function value in either the $x$-direction (via $f_x$) or the $y$-direction (via $f_y$). However, in practice, we are only provided elevation data at a spatial resolution not fine enough to make $h \to 0$ precise. Can we still **estimate** what $f_x$ and $f_y$ are based on the given data?

A typical survey of data would be **level curves**, curves at which the function maintains the same value. Let's consider the level curves for the function $f(x,y) = 100 - x^2 - y^2$.

What if we want $\frac{\partial f}{\partial y} \mid_{(2,0)}$? We then need to look vertically (in the direction of $y$). The next available data is still the level set at $c = 3$, but mind you that the distance $h$ from the point of interest to the level curve needs to be calculated. The 16-level curve has equation $x^2 + y^2 = 4^2$. So if $x = 2$, then $y = \pm\sqrt{12}$. Since $h$ is a distance (nonnegative number), $h = \sqrt{12}$. Thus, by a similar calculation, we have

$$\begin{aligned} \frac{\partial f}{\partial y} \mid_{(2,0)} &= \lim_{h \to 0} \frac{f(2, 0+h) - f(2,0)}{h} \\ &\overset{h=\sqrt{12}}{\approx} \frac{f(2, 0+h) - f(2,0)}{h} \\ &= \frac{f\left(2, \sqrt{12}\right) - f(2,0)}{\sqrt{12}} \\ &= \frac{16 - 21}{\sqrt{12}} \\ &= -\frac{5}{\sqrt{12}} \\ &\approx -1.44. \end{aligned}$$

Now, the true partial derivative is

$$\frac{\partial f}{\partial y} \mid_{(2,0)} = -2y \mid_{(2,0)} = -4$$

which we now miss by quite a margin. The reason is straightforward – our $h$ is too big. How shall we remedy this error? In practice, you may use a query point closer to the point of interest, or you may need to survey more data.

## 5.4 Optional Reading: Important Connection between Partial Derivatives and Continuity

If a function of one variable $f(x)$ is differentiable at $x = x_0$, then the function is automatically continuous at that point. However, when we are in higher dimensions,

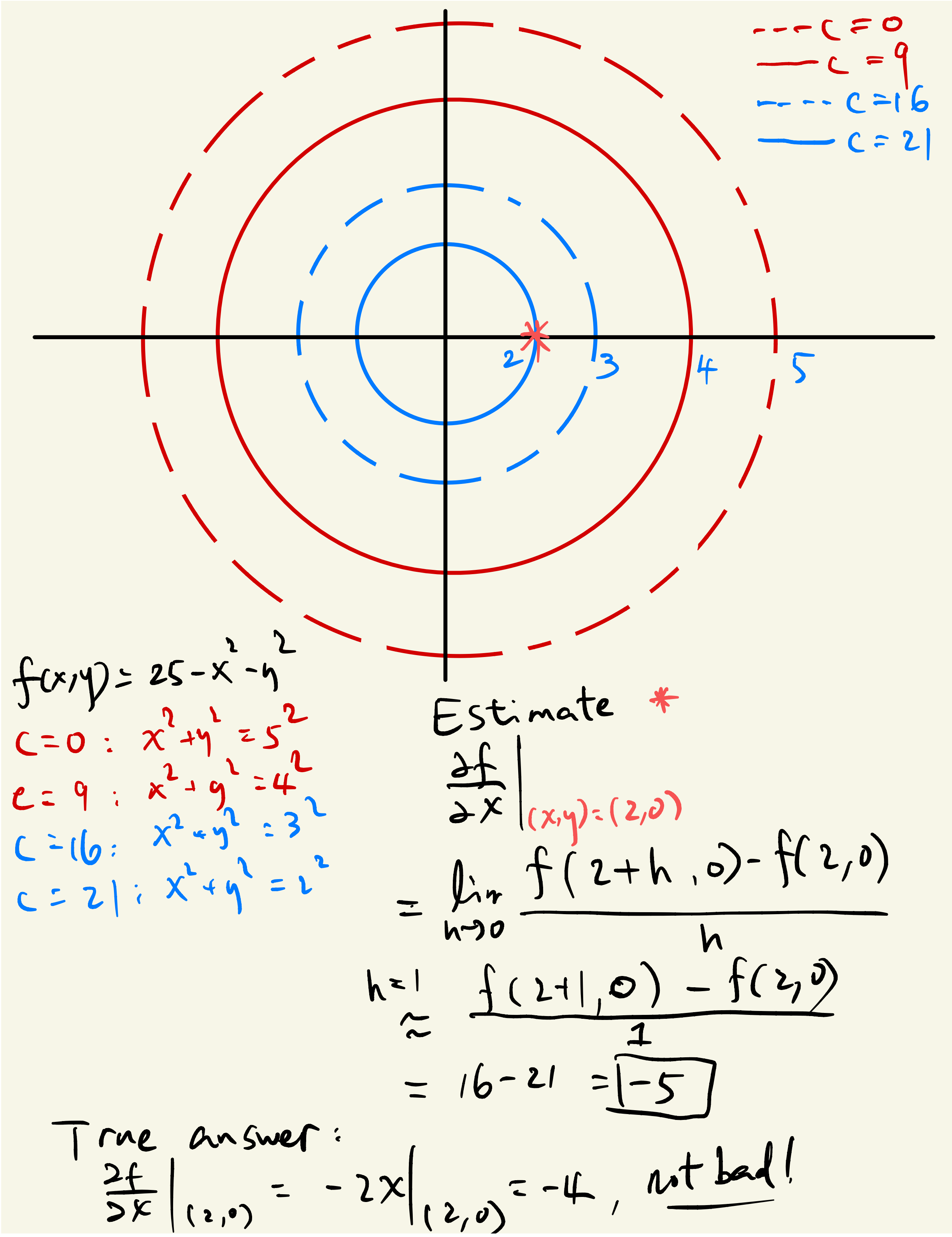



Figure 5: The figure shows the procedures to obtain an estimate for $\frac{\partial f}{\partial x}\mid_{(2,0)}$.

this is no longer true because the partial derivatives are only limits taken in a very specific direction. The function may behave very "badly" away from these directions.

Consider the following function

$$z = f(x, y) = \begin{cases} 0, & xy \neq 0; \\ 1, & xy = 0. \end{cases}$$

This function consists of two lines at height $z = 1$ parallel to the $x$ and $y$ axis respectively, and four disconnected flat planes elsewhere at height $z = 0$. This surface looks like

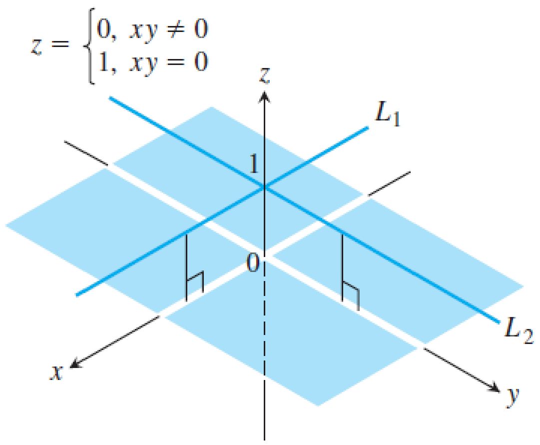



**FIGURE 14.21** The graph of

$$f(x, y) = \begin{cases} 0, & xy \neq 0 \\ 1, & xy = 0 \end{cases}$$

consists of the lines $L_1$ and $L_2$ (lying 1 unit above the $xy$-plane) and the four open quadrants of the $xy$-plane. The function has partial derivatives at the origin but is not continuous there (Example 8).

1. Find the limit of $f(x, y)$ as we approach the original along the line $y = x$.

**Solution 5.4.** We have

$$\lim_{(x,y)\to(0,0),\quad y=x} f(x,y) = \lim_{x\to 0} 0 = 0.$$

Note that $f(0,0) = 1 \neq 0$ which immediately implies that $f$ is NOT continuous at the origin.

2. Show that both partial derivatives exist at the origin.

**Solution 5.5.** Using the definition of partial derivatives,

$$\frac{\partial f}{\partial x}\mid_{(x,y)=(0,0)} = \frac{d}{dx} f(x,0)\mid_{x=0} = \frac{d}{dx}(1)\mid_{x=0} = 0.$$

and

$$\frac{\partial f}{\partial y}\mid_{(x,y)=(0,0)} = \frac{d}{dy} f(0,y)\mid_{y=0} = \frac{d}{dy}(1)\mid_{x=0} = 0.$$

This example informs us that we need stronger requirements in higher dimensions to obtain full differentiability. In fact, we need the partial derivatives to be continuous in an open region containing the point of interest – not just continuous at it. The precise definition of **differentiability in higher dimensions** is in the textbook, Section 14.3.

# 6 Lecture VI: Tangent Planes to Explicit Surfaces and Linearization

We have learned about the tangent line in Calculus I since its slope is precisely the derivative value at the point of tangency. Now, we are working with functions of several variables, and thus the notion of tangency must also upgrade. A tangent line now becomes a tangent plane. But how does this transformation take place geometrically?

The key word here is **explicit**, and we will revisit tangent planes to **implicit** surfaces in a later lecture.

## 6.1 Tangent Planes to Explicit Surfaces

**Definition 6.1** (Tangent Plane)**.** *Let $P_0 = (x_0, y_0, z_0)$ be a point on a surface $S$, and $C$ be any curve passing through $P_0$ and lying entirely in $S$. If the tangent lines to all such curves $C$ at $P_0$ lie in the same plane, then this plane is called the **tangent plane** to $S$ at $P_0$.*

This definition is legendary because it is building upon collecting all **tangent lines**, a notion we understand, into a **tangent plane**, a new object. In other words, a tangent plane is a plane spanned by vectors parallel to all tangent lines at $P_0$. This observation helps us find the equation of a tangent plane.

**Definition 6.2** (Equation of Tangent Plane given an **Explicit** Surface $z = f(x, y)$)**.** *Let $S$ be a surface defined by a differentiable function $z = f(x, y)$, and let $P_0 = (x_0, y_0)$ be a point in the domain of $f$. Then, the equation of the tangent plane to $S$ at $P_0$ is given by*

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

Why is this formula correct? Let's try to identify two possible tangent lines at the point of interest $P_0$.

Remember from last class that a partial derivative against $x$ at $x_0$ is the slope of the tangent line to the curve, formed by intersecting the surface and the vertical plane $y = y_0$. Thus, this tangent line has point-slope form

$$l_1(x) = f(x_0, y_0) + f_x(x_0, y_0)(x - x_0).$$

Similarly, another tangent line can be found via the partial derivative against $y$, namely,

$$l_2(x) = f(x_0, y_0) + f_y(x_0, y_0)(y - y_0).$$

The plane that contains these two tangent lines is the tangent plane, whose normal is a vector perpendicular to both of these two tangent lines. As of now, we need to identify the direction of the **vector** each line represents.

Given a line $y = mx + b$, can you identify a vector in the parallel to this line knowing that the slope is $m$? Yes,

$$v = \langle 1, m \rangle = \mathbf{i} + m\mathbf{j}.$$

But this is in the $xy$-plane. The two lines we found above live in $\mathbb{R}^3$, which corresponds to vectors of the form $v = \langle a, b, c \rangle = a\mathbf{i} + b\mathbf{j} + c\mathbf{k}$. So, a vector parallel to $l_1$ is

$$v_1 = \langle 1, 0, m \rangle = \mathbf{i} + m\mathbf{k} = \mathbf{i} + f_x(x_0, y_0)\,\mathbf{k}$$

where the slope is put in the $\mathbf{k}$ component. Similarly, a vector parallel to $l_2$ is

$$v_2 = \langle 0, 1, f_y(x_0, y_0) \rangle = \mathbf{j} + f_y(x_0, y_0)\,\mathbf{k}$$

Now, we are ready to find the equation of the tangent plane spanned by $v_1$ and $v_2$ by recalling from Calculus III that we need to compute the cross product of these two vectors, which produces the normal vector $\mathbf{n}$ to the plane.

$$\mathbf{n} = v_1 \times v_2 = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & 0 & f_x(x_0, y_0) \\ 0 & 1 & f_y(x_0, y_0) \end{vmatrix} = \left(-f_x(x_0, y_0)\right)\mathbf{i} + \left(-f_y(x_0, y_0)\right)\mathbf{j} + \mathbf{k}$$

Therefore, the equation of the tangent plane is given by the orthogonality condition

$$\mathbf{n} \cdot \mathbf{r} = \langle -f_x(x_0, y_0), -f_y(x_0, y_0), 1 \rangle \cdot \langle x - x_0, y - y_0, z - z_0 \rangle = 0$$

Carrying out the dot product, we have

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

Simplifying by isolating $z$, we obtain

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

as promised.

**Remark 6.1.** *Note that the tangent plane is flat, i.e., $z = z_0$, when both partial derivatives vanish. Kind of make sense?*

**Example 6.1.** Find the equation of the tangent plane to the surface defined by the function $f(x, y) = x^2 + y^2$ at point $(1, 2)$.

**Solution 6.1.** We need the following ingredients

$$\begin{aligned} z_0 = f(1,2) &= 1^2 + 2^2 = 5 \\ f_x(1,2) &= 2x \,|_{(x,y)=(1,2)} = 2 \\ f_y(1,2) &= 2y \,|_{(x,y)=(1,2)} = 4 \end{aligned}$$

Thus, the equation of the tangent plane, using the formula is

$$\begin{aligned} z &= z_0 + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0) \\ &= 5 + 2(x-1) + 4(y-2). \end{aligned}$$

## 6.2 Linearization

Recall that in Calculus I, we can approximate a differentiable function $f(x)$ near $x = x_0$ by its tangent line, i.e.,

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

This notion extends naturally and beautifully to higher dimensions.

**Definition 6.3** (Linearization of an Explicit Surface)**.** *Given a function $z = f(x,y)$ with continuous partial derivatives that exist at the point $(x_0, y_0)$, the **linear approximation** of $f$ at the point $(x_0, y_0)$ is given by the equation*

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

**Remark 6.2.** *Make a strong note that this expression is precisely the equation of the tangent plane. This is saying that near the point of interest $(x_0, y_0)$, the surface $z = f(x,y)$ can be approximated by the tangent plane at $(x_0, y_0)$.*

So, let's approximate some function values using linearization.

**Example 6.2.** Given the function $f(x,y) = e^{5-2x+3y}$, approximate $f(4.1, 0.9)$ using the point $(4,1)$ to four decimal places?

**Solution 6.2.** Near the point of interest $(4.1, 0.9)$, the function $f$ can be approximated by its linearization/tangent plane at $(4,1)$. The ingredients we need for the expression of the linearization are

$$\begin{aligned} f(x_0, y_0) = f(4,1) &= e^{5-2\cdot4+3\cdot1} = e^0 = 1 \\ f_x(4,1) &= e^{5-2x+3y}(-2) \,|_{(x,y)=(4,1)} = e^0(-2) = -2 \\ f_y(4,1) &= e^{5-2x+3y}(3) \,|_{(x,y)=(4,1)} = e^0(3) = 3 \end{aligned}$$

Therefore, the linearization is given by

$$\begin{aligned} L(x,y) &= f(x_0,y_0) + f_x(x_0,y_0)(x-x_0) + f_y(x_0,y_0)(y-y_0) \\ &= f(4,1) + f_x(4,1)(x-4) + f_y(4,1)(y-1) \\ &= 1 - 2(x-4) + 3(y-1) \end{aligned}$$

We evaluate $L$ at the point of interest,

$$\begin{aligned} L(4.1,0.9) &= 1 - 2(4.1-4) + 3(0.9-1) \\ &= 1 - 2(0.1) + 3(-0.1) \\ &= 1 - 0.2 - 0.3 \\ &= 0.5 \end{aligned}$$

The true value is $f(4.1,0.9) \approx 0.6065$. So, we missed about 20%. But not bad using the simple method.

# 7 Lecture VII: The Chain Rule and Dependency Diagrams

Recall from Calculus I that given a composite function $f(g(x))$, the derivative satisfies

$$\frac{d}{dx}(f(g(x))) = f'(g(x))\,g'(x).$$

This formula holds for single-variable functions $f$ and $g$. Now, what if $f$ has two independent variables $x$ and $y$ which both further depend on some new variable $t$? We are looking for an expression for

$$\frac{d}{dt}f(x(t), y(t)).$$

Note that this expression is well-defined because ultimately, $f$ depends only on $t$, since it is the only innermost variable.

A side note. The function $f(x(t), y(t))$ has a physical interpretation: the $t$ dependence can be thought of as time, which means $(x(t), y(t))$ is the position of a particle, traveling according to the law governed by $x(t)$ in the $x$-direction, and $y(t)$ in the $y$-direction. Think of it as a bug crawling on the surface at a height $f(x(t), y(t))$, or you hiking in the mountain ranges, both of which are **dynamics**.

## 7.1 The Chain Rule

For $z = f(x, y)$ with $x = x(t)$ and $y = y(t)$, we have

$$\begin{aligned}\frac{dz}{dt} &= f_x(x(t), y(t))\,x'(t) + f_y(x(t), y(t))\,y'(t)\\ &= \frac{\partial z}{\partial x}\frac{dx}{dt} + \frac{\partial z}{\partial y}\frac{dy}{dt}.\end{aligned}$$

The formula in fact requires a proof, nicely demonstrated by your textbook on page 406.

We will focus on its interpretation which helps us understand and thus remember the formula. Let's focus on the first term

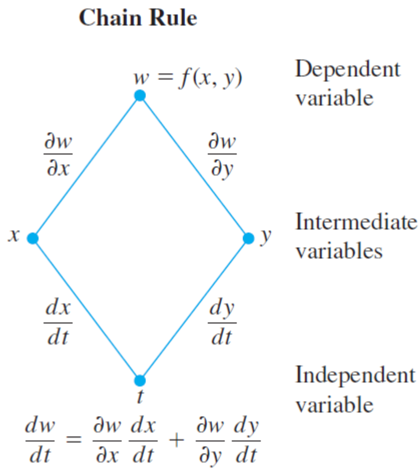
$$\frac{\partial z}{\partial x}\frac{dx}{dt}.$$

Given the physical intuition we spoke about earlier, we note that $\frac{dx}{dt}$ is no other than the velocity of the particle <u>in the $x$-direction</u>. If you are climbing an incline with slope $\frac{\partial z}{\partial x}$ at a velocity of $\frac{dx}{dt}$, then the rate of elevation gain should be precisely the product of the slope and the velocity.

Meanwhile, a particle doesn't just venture in one single direction. It may incur some velocity in the $y$-direction. The rate of elevation gain must then also add in the contribution from the $y$-direction climb, thus forming the second part of the sum.

## 7.2 Dependency Diagrams

This chain rule with one independent variable $t$ and two intermediate variables $x$ and $y$ has the following dependency graph:




**Example 7.1.** Find $\frac{dw}{dt}$ where $w = xy$ and $x = \cos(t)$ and $y = \sin(t)$, and draw the dependency diagram.

**Solution.** We do exactly as the Chain rule suggests. Write down the full form

$$
\begin{aligned}
\frac{dw}{dt} &= \frac{\partial w}{\partial x}\frac{dx}{dt} + \frac{\partial w}{\partial y}\frac{dy}{dt} \\
&= y(-\sin(t)) + x\cos(t) \\
&= \cos(t)(-\cos(t)) + \sin(t)\sin(t) \\
&= \cos^2(t) - \sin^2(t) \\
&= \cos(2t)\,. \quad \text{(optional step)}
\end{aligned}
$$

We can certainly extend a similar dependency diagram, to functions with two independent variables $(t, s)$ and two intermediate variables $(x, y)$. In fact, we can generalize to $n$ independent variables and $m$ intermediate variables, for arbitrary

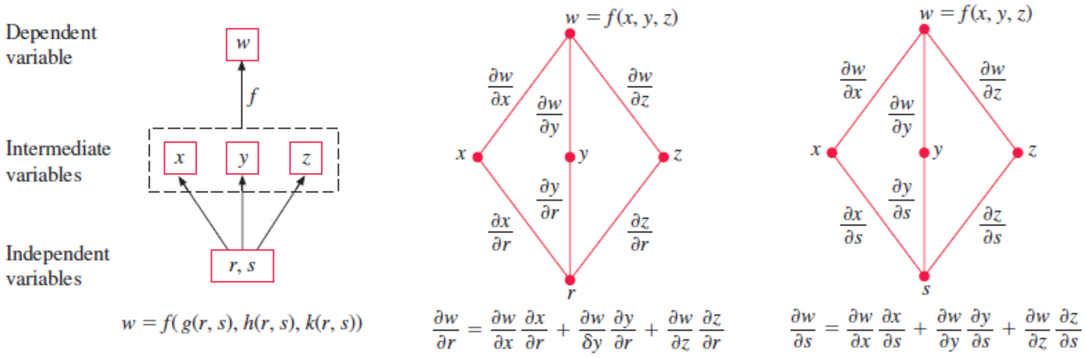



Figure 6: Dependency Diagram for Example 7.2.

integers $n$ and $m$. Of course, dependency diagram will look a bit more complicated, with more "leaves" and "branches".

**Example 7.2.** Let $w = f(x, y, z)$, $x = x(r, s)$, $y = y(r, s)$ and $z = z(r, s)$. Find all partial derivatives in $r$ and $s$.

$$\frac{\partial w}{\partial r} = \boxed{\frac{\partial w}{\partial x}}\frac{\partial x}{\partial r} + \boxed{\frac{\partial w}{\partial y}}\frac{\partial y}{\partial r} + \boxed{\frac{\partial w}{\partial z}}\frac{\partial z}{\partial r}$$

and

$$\frac{\partial w}{\partial s} = \boxed{\frac{\partial w}{\partial x}}\frac{\partial x}{\partial s} + \boxed{\frac{\partial w}{\partial y}}\frac{\partial y}{\partial s} + \boxed{\frac{\partial w}{\partial z}}\frac{\partial z}{\partial s}.$$

**Remark 7.1.** *Notice the boxed terms: they are the repeated part in each partial derivative. This means you just need to compute each of one time, and use them when necessary.* ***These repeated partial derivatives are precisely the partial derivatives against the intermediate variables.***

Let's do a concrete example with explicit functions.

**Example 7.3.** Calculate $\frac{\partial z}{\partial u}$ and $\frac{\partial z}{\partial v}$ using the following functions

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

**Solution.** Here, notice that $z$ depends on two intermediate variables, which depend on two independent variables individually. We have a total of 6 partial derivatives in total to calculate. Putting things together, we have

$$\begin{aligned}\frac{\partial z}{\partial u} &= \boxed{\frac{\partial z}{\partial x}}\frac{\partial x}{\partial u} + \boxed{\frac{\partial z}{\partial y}}\frac{\partial y}{\partial u} \\ &= (6x - 2y)(3) + (-2x + 2y)(4) \\ &= 10x + 2y \\ &= 10(3u + 2v) + 2(4u - v) \\ &= 38u + 18v\end{aligned}$$

and

$$\begin{aligned}\frac{\partial z}{\partial v} &= \boxed{\frac{\partial z}{\partial x}}\frac{\partial x}{\partial v} + \boxed{\frac{\partial z}{\partial y}}\frac{\partial y}{\partial v} \\ &= (6x - 2y)(2) + (-2x + 2y)(-1) \\ &= 14x - 6y \\ &= 14(3u + 2v) - 6(4u - v) \\ &= 18u + 34v\end{aligned}$$

Notice again that we have a few terms here that only require computation one time.

The fact that we need to compute the partial derivatives against the intermediate variables ONLY ONE TIME means that we can save some memory when giving this job to a computer. The following optional section checks the computational savings when computing a full partial derivative in a more general setting, when the number of intermediate variables and independent variables is arbitrary.

## 7.3 Optional: The General Chain Rule and Computational Savings

Let's consider the chain rule in a more general setting. Suppose we have $z = f(x, y)$ still, but now $x = x(t, s)$ and $y = y(t, s)$, that is, the original "independent variables" now depend on two more independent variables. This is saying that $z$ actually takes $t$ and $s$ as independent variables, which then may have two partial derivatives, $\frac{\partial z}{\partial t}$ and $\frac{\partial z}{\partial s}$.

$$\frac{\partial z}{\partial t} = \boxed{\frac{\partial z}{\partial x}}\frac{\partial x}{\partial t} + \boxed{\frac{\partial z}{\partial y}}\frac{\partial y}{\partial t}$$

and

$$\frac{\partial z}{\partial s} = \boxed{\frac{\partial z}{\partial x}}\frac{\partial x}{\partial s} + \boxed{\frac{\partial z}{\partial y}}\frac{\partial y}{\partial s}.$$

Here comes the observation on memory saving. Our goal is to calculate $\frac{\partial z}{\partial t}$ and $\frac{\partial z}{\partial s}$ with given function forms of $f, x$ and $y$. Note that the boxed terms are computed twice respectively, if we were to each partial derivative separately. In practice, you only need $\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial y}$ once. So here, the computational saving is 2 terms out of a total of 8 terms, so we save 25%.

Even more generally, consider

$$z = f\left(x_1, \ldots, x_n\right)$$

a function of $n$ variables, while each $x_i = x_i\left(w_1, \ldots, w_m\right)$ is a function of $m$ variables. Then, the chain rule looks like a sum

$$\frac{\partial z}{\partial w_i} = \boxed{\frac{\partial z}{\partial x_1}}\frac{\partial x_1}{\partial w_i} + \boxed{\frac{\partial z}{\partial x_2}}\frac{\partial x_2}{\partial w_i} + \cdots + \boxed{\frac{\partial z}{\partial x_n}}\frac{\partial x_n}{\partial w_i} = \sum_{j=1}^{n} \boxed{\frac{\partial z}{\partial x_j}}\frac{\partial x_j}{\partial w_i}.$$

This formula holds for $i = 1, 2, 3, \ldots, m$. The boxed terms only need to be computed once since they show up for every $i$ considered.

Let's now do some complexity analysis. For each $i$, $\frac{\partial z}{\partial w_i}$ requires computing $2n$ derivatives, and thus totals $2nm$ number of derivatives if we were to brute force the computation on $\frac{\partial z}{\partial w_i}$ for every $i = 1, 2, \ldots, m$. But from each computation, we save $n$ calculations of the boxed terms, so in total, we only need to compute $2nm - n\left(m-1\right) = nm + n$. The percentage of computation saved is

$$\frac{nm + n}{2nm} = \frac{1}{2} + \frac{1}{2m}$$

which is more than half. What's more, the computational saving is independent of $n$, indicating that no matter how many variables $z$ has, we can always save by more than half the computations.

# 8 Lecture VIII: Directional Derivative as a Rate of Change

## 8.1 Preface to Lectures VIII and IX

From the expression of the tangent plane, to the construction of the Chain rule from a physical point of view, we saw expressions like

$$L(x,y) = f(x_0,y_0) + \boxed{f_x(x_0,y_0)(x-x_0) + f_y(x_0,y_0)(y-y_0)}.$$

or

$$\frac{\partial z}{\partial t} = \boxed{\frac{\partial z}{\partial x}}\frac{\partial x}{\partial t} + \boxed{\frac{\partial z}{\partial y}}\frac{\partial y}{\partial t}.$$

And to go in a higher dimension, we saw expressions like

$$\frac{\partial w}{\partial r} = \boxed{\frac{\partial w}{\partial x}}\frac{\partial x}{\partial r} + \boxed{\frac{\partial w}{\partial y}}\frac{\partial y}{\partial r} + \boxed{\frac{\partial w}{\partial z}}\frac{\partial z}{\partial r}$$

and

$$\frac{\partial w}{\partial s} = \boxed{\frac{\partial w}{\partial x}}\frac{\partial x}{\partial s} + \boxed{\frac{\partial w}{\partial y}}\frac{\partial y}{\partial s} + \boxed{\frac{\partial w}{\partial z}}\frac{\partial z}{\partial s}.$$

What do they have in common? They are all sums of products.

In calculus III, you have learned the dot product between two vectors

$$\boldsymbol{v} \cdot \boldsymbol{u} = v_1 u_1 + v_2 u_2,$$

which is also incidentally a sum of products. Any coincidence here? The dot product has deep geometric meanings. Do these meanings somehow lend themselves to partial derivatives?

## 8.2 Rate of Change Along a Line

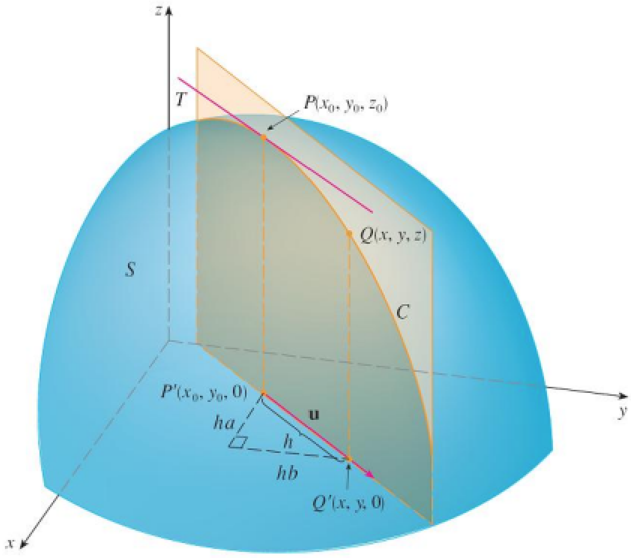



Figure 7: A vertical plane represents the direction **u** in the $xy$-plane, intersecting the surface at the curve $C$. We are interested in learning the slope of this curve along the surface, captured by the directional derivative precisely.

Suppose a particle travels on a straight line is in the direction of $\boldsymbol{u} = (u_1, u_2) = u_1\boldsymbol{i} + u_2\boldsymbol{j}$ and $\boldsymbol{u}$ is a **unit vector** (see figure, the vertical plane can be characterized by the point $P$ and the direction $\boldsymbol{u}$). Then, any arbitrary point in space passing through the point $(x_0, y_0)$ satisfy

$$\begin{aligned} x &= x_0 + u_1 s; \\ y &= y_0 + u_2 s, \end{aligned}$$

where $s \in \mathbb{R}$ ($s$ here simply takes care of the magnitudes of any vector in the direction of $\boldsymbol{u}$). Now, the question is, how do we compute the rate of change in elevation as the particle traverses the surface along this straight line? The story is philosophically the same as before in 1D calculus, namely, rise over run.

**Definition 8.1.** *(Directional derivative) The derivative of $f$ at $(x_0, y_0)$ in the direction of the* ***unit vector*** $\boldsymbol{u} = u_1\boldsymbol{i} + u_2\boldsymbol{j}$ *is the number*

$$\left(\frac{df}{ds}\right)_{\boldsymbol{u},(x_0,y_0)} = \lim_{s\to 0} \frac{f\left(x_0 + su_1, y_0 + su_2\right) - f\left(x_0, y_0\right)}{s}$$

*provided the limit exists. We also write*

$$D_{\boldsymbol{u}} f\left(x_0, y_0\right), \quad D_{\boldsymbol{u}} f\mid_{\left(x_0, y_0\right)}, \quad \left(\frac{df}{ds}\right)_{\boldsymbol{u},\left(x_0, y_0\right)}.$$

**Remark 8.1.** *Note that the numerator is the "rise", represented by how much the function has changed after moving in the direction of* $\boldsymbol{u}$. *The denominator is the size of "run". This is the classical "rise over run" definition of a derivative, now just with a sense of direction since the domain is 2D.*

From this definition, we see that

$$\frac{\partial f}{\partial x} = D_{\hat{\boldsymbol{\imath}}} f, \quad \frac{\partial f}{\partial x} = D_{\hat{\boldsymbol{\jmath}}} f,$$

**that is, the partial derivatives are directional derivatives in the direction of the two principal axes respectively.** This fact makes directional derivative a more general notion of derivative than the partial derivative since the latter is a special case (with $\boldsymbol{u} = \hat{\boldsymbol{\jmath}} = (1, 0)$ for $f_x$ and $\boldsymbol{u} = \hat{\boldsymbol{\imath}} = (0, 1)$ for $f_y$).

So, can we now compute the directional derivative of a function given a direction? Certainly. Let's see in the following example how the formula applies.

**Example 8.1.** Consider the surface $f(x, y) = xy$. Compute the directional derivative at $(1, 2)$ in the direction of $\boldsymbol{u} = \left\langle \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right\rangle$.

**Solution 8.1.** Note that $\boldsymbol{u}$ has unit length, so we can proceed. If it is not, we need to normalize it by performing $\boldsymbol{u}/\left|\boldsymbol{u}\right|$, that is, divide each component by the length of the vector.

With $(x_0, y_0) = (1, 2)$, we have

$$\begin{aligned}
\left(\frac{df}{ds}\right)_{\boldsymbol{u},(1,2)} &= \lim_{s\to 0} \frac{f\left(x_0 + su_1, y_0 + su_2\right) - f\left(x_0, y_0\right)}{s} \\
&= \lim_{s\to 0} \frac{f\left(1 + s\frac{1}{\sqrt{2}}, 2 + s\frac{1}{\sqrt{2}}\right) - f(1, 2)}{s} \\
&= \lim_{s\to 0} \frac{\left(1 + \frac{s}{\sqrt{2}}\right)\left(2 + \frac{s}{\sqrt{2}}\right) - 1 \cdot 2}{s} \\
&= \lim_{s\to 0} \frac{2 + \frac{3}{\sqrt{2}} s + \frac{s^2}{2} - 1 \cdot 2}{s} \\
&= \lim_{s\to 0} \frac{\frac{3}{\sqrt{2}} s + \frac{s^2}{2}}{s}
\end{aligned}$$

$$= \lim_{s \to 0} \frac{3}{\sqrt{2}} + \frac{s}{2}$$
$$= \frac{3}{\sqrt{2}}.$$

This result says that the rate of change of the surface $f(x, y) = xy$ at $(1, 2)$ in the direction $\boldsymbol{u}$ is $\frac{3}{\sqrt{2}}$. Since it is a positive number, we then know that the function value is increasing in the direction of $\boldsymbol{u}$.

# 9 Lecture IX: Directional Derivative as a Dot Product and the Gradient Vector

We continue from last class. Looking at the last example, we saw that we went through an explicit limit calculation. But we didn't quite go through this procedure in Calculus I after we learned some differentiation rules. Is there going to be something similar now, i.e., a nicer formula that does not involve limits?

## 9.1 A Formula for the Directional Derivative

Given a function $f(x,y)$ is differentiable, if a particle is traveling on a curve given by $(x(s), y(s))$ on the surface, we can find its rate of change. More precisely, the **Chain rule** identifies the rate

$$\frac{df}{dt} = \frac{\partial f}{\partial x}\frac{dx}{ds} + \frac{\partial f}{\partial y}\frac{dy}{ds}.$$

**Fact 9.1.** *Note that the vector $\left(\frac{dx}{ds}, \frac{dy}{ds}\right)$ goes in the* ***direction*** *of the tangent line to the parametric curve given by $x = x(s)$ and $y = y(s)$. We will come back to use this later.*

The directional derivative concerns the **special** case where $(x(s), y(s))$ represent a line in the direction of $\boldsymbol{u} = (u_1, u_2)$. The parameterization of the vector $\boldsymbol{u} = (u_1, u_2)$ is

$$x = x_0 + u_1 s;$$
$$y = y_0 + u_2 s.$$

Using the Chain rule, we find, for $f(x,y) = f(x(s), y(s))$,

$$\left(\frac{df}{ds}\right)_{\boldsymbol{u},(x_0,y_0)} = \frac{\partial f}{\partial x}\big|_{(x_0,y_0)} \frac{dx}{ds} + \frac{\partial f}{\partial y}\big|_{(x_0,y_0)} \frac{dy}{ds} \tag{9.1a}$$

$$= \frac{\partial f}{\partial x}\big|_{(x_0,y_0)} u_1 + \frac{\partial f}{\partial y}\big|_{(x_0,y_0)} u_2 \tag{9.1b}$$

$$= \boxed{\left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right)\big|_{(x_0,y_0)}} \cdot (u_1, u_2). \tag{9.1c}$$

The boxed item here also warrants a definition.

**Definition 9.1.** *The* ***gradient vector*** *(or* ***gradient****) of $f(x,y)$ is the vector*

$$\nabla f := \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right) = \frac{\partial f}{\partial x}\boldsymbol{i} + \frac{\partial f}{\partial y}\boldsymbol{j}.$$

*We may evaluate the gradient vector at a point $(x_0, y_0)$ using notations*

$$\nabla f\,|_{(x_0,y_0)}, \quad \nabla f\,(x_0, y_0)\,.$$

$\nabla f$ *is read "grad $f$" as well as "gradient of $f$" and "del $f$" (also "nabla $f$").*

With this setup, the directional derivative can then be written as a dot product between the gradient and the direction.

$$\left(\frac{df}{ds}\right)_{\boldsymbol{u},(x_0,y_0)} = \nabla f\,(x_0, y_0) \cdot \boldsymbol{u} = \nabla f\,|_{(x_0,y_0)} \cdot\boldsymbol{u}.$$

## 9.2 Examples

**Example 9.1.** Now, let's verify the example from last lecture using the freshly learned formula for Directional Derivative. We were interested in finding the rate of change of the function $f\,(x, y) = xy$ at $(1, 2)$ in the direction of $\boldsymbol{u} = \left\langle \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right\rangle$.

**Solution 9.1.** We compute the gradient vector, namely, the vector whose components consist of the partial derivatives of $f$.

$$\nabla f\,(1, 2) = \left\langle \frac{\partial f}{\partial x}\,(x, y)\,, \frac{\partial f}{\partial y}\,(x, y) \right\rangle |_{(1,2)} = \langle y, x\rangle\,|_{(1,2)} = \langle 2, 1\rangle$$

Then, by Equation (9.1c), we have

$$\left(\frac{df}{ds}\right)|_{\boldsymbol{u},(1,2)} = \nabla f\,(1, 2) \cdot \boldsymbol{u} = \langle 2, 1\rangle \cdot \left\langle \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right\rangle = \frac{3}{\sqrt{2}},$$

as expected.

**Example 9.2.** Find the derivative of $f\,(x, y) = xe^y + \cos(xy)$ at the point $(2, 0)$ in the direction of $\boldsymbol{v} = 3\boldsymbol{i} - 4\boldsymbol{j}$.

One follows the dot product formulation,

$$\left(\frac{df}{dt}\right)_{\boldsymbol{v},(2,0)} = \nabla f\,|_{(2,0)} \cdot\hat{\boldsymbol{v}}$$

We now need grad $f$ and $\hat{\boldsymbol{v}}$.

$$\nabla f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right) = \left(e^y - y\sin(x)\,, xe^y - x\sin(y)\right)$$

and the normalized $\boldsymbol{v}$

$$\hat{\boldsymbol{v}} = \frac{\boldsymbol{v}}{\sqrt{3^2 + (-4)^2}} = \frac{3}{5}\boldsymbol{i} - \frac{4}{5}\boldsymbol{j}.$$

Evaluated at $(2,0)$, we have

$$\nabla f \mid_{(2,0)} = (1,2).$$

Therefore,

$$\left(\frac{df}{dt}\right)_{\boldsymbol{v},(2,0)} = \nabla f \mid_{(2,0)} \cdot \hat{\boldsymbol{v}} = (1,2) \cdot \left(\frac{3}{5}, -\frac{4}{5}\right) = -1.$$

**Remark 9.1.** *Normalizing is extremely important, and must be carried out when the given vector does not have unit length. See "Alternative Definition of Directional Derivative and Why Unit Vector" on Canvas main page.*

## 9.3 Operations with the Gradient

**Definition 9.2.** *The **gradient vector** (or **gradient**) of $f(x,y)$ is the vector*

$$\nabla f := \left\langle \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right\rangle = \frac{\partial f}{\partial x}\boldsymbol{i} + \frac{\partial f}{\partial y}\boldsymbol{j}.$$

*We may evaluate the gradient vector at a point $(x_0, y_0)$ using notations*

$$\nabla f \mid_{(x_0,y_0)}, \quad \nabla f(x_0, y_0).$$

*$\nabla f$ is read "grad $f$" as well as "gradient of $f$" and "del $f$" (also "nabla $f$").*

The gradient of $f(x,y)$ is defined as the **VECTOR**

$$\nabla f = \left\langle \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right\rangle.$$

As the gradient involves the operation of differentation, algebraic rules from regular 1D differentiation carry through.

1. (Sum/Difference) $\nabla(f \pm g) = \nabla f \pm \nabla g$. This is also known as **Term-by-Term differentiation**, which only applies in a sum/difference setting.
2. (Constant multiple) $\nabla(kf) = k\nabla f$, for $k \in \mathbb{R}$.
3. (Product) $\nabla(fg) = f\nabla g + g\nabla f$.
4. (Quotient) $\nabla\left(\frac{f}{g}\right) = \frac{g\nabla f - f\nabla g}{g^2}$, given $g \neq 0$ in the region where you want to find the derivatives for.

**Example 9.3.** Find the gradient of $\frac{x^2+y^2}{x+y}$, away from the line $y \neq -x$.

**Solution 9.2.** We make use of the Quotient rule. We identify

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

So, the given function is in the form of $\frac{f(x,y)}{g(x,y)}$, with a restriction that $y \neq -x$. All we need are the gradients of $f$ and $g$, respectively, and then we invoke the Quotient Rule (point 4 above).

$$\nabla f = \langle 2x, 2y \rangle, \quad \nabla g = \langle 1, 1 \rangle.$$

Thus,

$$\begin{aligned} \nabla\left(\frac{x^2+y^2}{x+y}\right) &= \nabla\left(\frac{f}{g}\right) \\ &= \frac{g\nabla f - f\nabla g}{g^2} \\ &= \frac{(x+y)\langle 2x, 2y\rangle - (x^2+y^2)\langle 1, 1\rangle}{(x+y)^2} \\ &= \frac{\langle 2x^2+2xy, 2xy+2y^2\rangle - \langle x^2+y^2, x^2+y^2\rangle}{(x+y)^2} \\ &= \frac{\langle 2x^2+2xy-x^2-y^2, 2xy+2y^2-x^2-y^2\rangle}{(x+y)^2} \\ &= \frac{\langle x^2+2xy-y^2, y^2+2xy-x^2\rangle}{(x+y)^2}. \end{aligned}$$

**Remark 9.2.** *Once we identify $f$ and $g$ from a given function, then we can compute its gradient directly without having to consider $\frac{\partial}{\partial x}$ and $\frac{\partial}{\partial y}$ separately.*

# 10 Lecture X: The Gradient Vector and the Level Curves

## 10.1 Interpretation of the Gradient Vector

Since the directional derivative is a dot product, it also has a strong geometric interpretation. Recall that the dot product between two vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ is

$$\boldsymbol{x} \cdot \boldsymbol{y} = \|\boldsymbol{x}\| \, \|\boldsymbol{y}\| \cos(\theta)$$

where $\|\cdot\|$ is the length of the vector, computed via Pythagorus (see step 2 in the last example).

$$D_{\boldsymbol{u}} f = \nabla f \cdot \boldsymbol{u} = \|\nabla f\| \, \|\boldsymbol{u}\| \cos(\theta) = \|\nabla f\| \cos(\theta)$$

since $\|\boldsymbol{u}\|$ is the length of a unit vector (length 1).

1. The directional derivative is at its <u>largest</u> when $\boldsymbol{u}$ is in the direction of $\nabla f$, i.e. the function $f$ increases the most rapidly in the direction of the gradient vector. The derivative in this direction is

   $$D_{\boldsymbol{u}} f = \|\nabla f\| \cos(0) = \|\nabla f\| .$$

   In other words, the function $f$ is **increasing** the fastest in the direction of $\nabla f$.

2. The directional derivative is at its <u>most negative</u> when $\boldsymbol{u}$ is in the direction of $-\nabla f$. The derivative in this direction is

   $$D_{\boldsymbol{u}} f = \|\nabla f\| \cos(\pi) = -\|\nabla f\| .$$

   In other words, the function $f$ is **decreasing** the fastest in the direction of $-\nabla f$. (A side comment: this very property led to a famous method called "Steepest Descent" when one seeks the local/global minimum of a multivariable function because along the negative gradient, the function decreases the fastest – hence the name Steepest Descent)

3. Any direction $\boldsymbol{u}$ *orthogonal* to a gradient is a direction of zero change in $f$ because $\theta = \frac{\pi}{2}$ and

   $$D_{\boldsymbol{u}} f = \|\nabla f\| \cos\left(\frac{\pi}{2}\right) = 0.$$

   In other words, the function $f$ is **not changing in value** in the direction perpendicular to $\nabla f$.

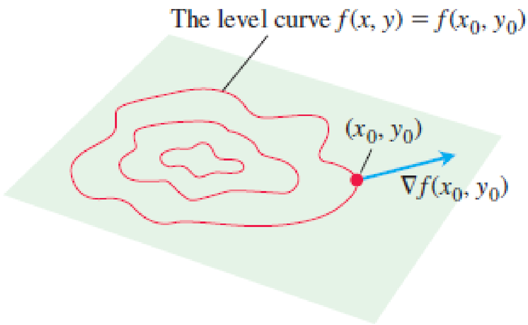
**Remark 10.1.** *The last property has deep connections to later topics in tangent planes and optimization. If the directional derivative is* 0*, it means in the direction* $\boldsymbol{u}$ *you are traveling, your height is NOT changing. This implies that if you travel in these directions, your trajectories are on the* ***level curve*** *of* $f$*. We will see how we can utilize this fact to construct tangent planes.*

## 10.2 Directional Derivative: a Recall and a Motivating Question

We know from criterion no.3 that

$$D_{\boldsymbol{u}} f = 0 \iff \nabla f \cdot \boldsymbol{u} = 0.$$

The LHS of the implication means that if you travel in the direction of $\boldsymbol{u}$, then the function value is not changing, i.e., along $\boldsymbol{u}$, you are travelling on the **level curve** of $f$. The RHS of the implication suggests that the gradient of $f$ (a vector) must be perpendicular/normal to the level curves.




**Example 10.1.** Consider the surface $z = f(x, y) = x^2 + y^2$ and its level set at $c = 4$. The level set is a circle of radius 2. Let the point of interest be $(0, 1)$. It is clear that that the tangent line at $(0, 1)$ is $y = 1$ – you can also say that the tangent line points in the direction of $\boldsymbol{T} = \langle 1, 0 \rangle$ (since it is parallel to the $x$-axis). We now check the gradient of $f$ evaluated at $(0, 1)$.

$$\nabla f(0, 1) = \langle 2x, 2y \rangle \,|_{(0,1)} = \langle 0, 2 \rangle \,.$$

Indeed, we find that $\nabla f(0, 1) = \langle 0, 2 \rangle$ dotted with the tangent vector $\boldsymbol{T}$ yields

$$\nabla f(0, 1) \cdot \boldsymbol{T} = \langle 0, 2 \rangle \cdot \langle 1, 0 \rangle = 0,$$

which implies that the gradient of $f$ at this point is perpendicular/normal to the tangent vector, and hence the level curve.

Now, the question is – can this beautiful relationship help us identify the expression of the tangent line at a given point on the level curve of a given function? It seems that $\nabla f$ is always perpendicular to the tangent line that contains the given point. Is this enough information? We first need to see the **normal form of a line**.

## 10.3 A Review: Normal Form of a Line

We are used to seeing the point-slope form of a line – given a point $(x_0, y_0)$ in the plane and the slope, we have

$$y - y_0 = m(x - x_0).$$

However, this expression does not generalize well to high dimensions because the notion of a slope (i.e. derivative) is hard to quantify, due to the very fact that you can speak of infinitely many derivatives at an arbitrary point (infinite ways of approaching a point in 3D).

We now formulate the **normal form** of a line. First of all, we know that a line points in certain direction, which can be formed by a vector in the plane. Suppose we want to characterize all the points in the line that contains $(x, y)$ and $(x_0, y_0)$. This line points in the direction of $\langle x - x_0, y - y_0 \rangle$. Now, if we also know the vector perpendicular to this line, say, $\boldsymbol{n} = \langle A, B \rangle$, then by the interpretation of the dot product, we have

$$0 = \boldsymbol{n} \cdot \boldsymbol{r} = \langle A, B \rangle \cdot \langle x - x_0, y - y_0 \rangle = \boxed{A(x - x_0) + B(y - y_0) = 0}.$$

The boxed expression is the celebrated **normal form** of a line. We call $\boldsymbol{n} = \langle A, B \rangle$ the normal vector of the line. All lines in the plane can be expressed in **normal form – and one should always consider lines as merely an expression of an orthogonal relationship**. One can always double check that a normal form can always be turned into the point-slope form, i.e.

$$A(x - x_0) + B(y - y_0) = 0 \implies y - y_0 = -\frac{A}{B}(x - x_0)$$

where the slope now is understood as $-\frac{A}{B}$.

Long story short, we need two ingredients in this case:

1. the **normal vector $\boldsymbol{n}$**;
2. a point in the line $(x_0, y_0)$.

To find the expression of the tangent line to a level curve utilizes the **normal form** of a line in a geometric fashion, and that's the beauty of Calculus IV (using materials from Calculus III). Below, we will see that the role of $\boldsymbol{n}$ will be filled by the gradient of a function $\nabla f$ when we are tasked to find the tangent line to a level curve.

## 10.4 Tangent Line to Level Curves

Suppose we have a surface $f(x, y)$. Consider its level curve $f(x, y) = c$. Pick a point $(x_0, y_0)$ on this curve. Let's find the expression of the tangent line at this point. We note that the points in this tangent line form the vector $(x - x_0, y - y_0)$ using the known point.

We know from the property of the directional derivative that the gradient of $f$ at $(x_0, y_0)$ (if exists) is always perpendicular/normal to its level curve, and hence perpendicular/normal to the tangent line at $(x_0, y_0)$. Great, we found both ingredients to express the tangent line. The tangent line (a set of points $(x, y)$) to the level curve of $f$ at $(x_0, y_0)$ satisfies the following equation

$$\nabla f \mid_{(x_0,y_0)} \cdot \langle x - x_0, y - y_0 \rangle = f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0) = 0.$$

**Example 10.2.** Consider the ellipse $\frac{x^2}{4} + y^2 = 2$. Find an equation for the tangent to this ellipse at the point $(-2, 1)$.

Define a surface

$$f(x, y) = \frac{x^2}{4} + y^2$$

so the ellipse is the 2-level curve $f(x, y) = 2$. By the normal form of a line

$$\boldsymbol{n} \cdot \langle x - x_0, y - y_0 \rangle = 0.$$

we need to identify the normal vector and a point on this line. We already know that $(x_0, y_0) = (-2, 1)$.

Since the gradient vector at this point is perpendicular to the level curve, it is also perpendicular to the tangent line, which suggests that it is a great candidate for the normal vector. Indeed,

$$\nabla f \mid_{(-2,1)} = \left\langle \frac{x}{2}, 2y \right\rangle \mid_{(-2,1)} = \langle -1, 2 \rangle.$$

Therefore, the equation of the tangent line is

$$\nabla f \cdot \langle x - x_0, y - y_0 \rangle = 0 \implies \langle -1, 2 \rangle \cdot \langle x + 2, y - 1 \rangle = 0$$

or more precisely,

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

Rearranging, we see that

$$y = \frac{x}{2} + 2.$$

**Remark 10.2.** *We can also solve this problem in a conventional way, that is, we first isolate $y$ and find a function that satisfies $y = f(x)$ and compute the standard tangent line. As a sanity check, we can in fact find the equation of the tangent using Calculus 1 method. We rewrite the equation of the ellipse as*

$$y = \sqrt{2 - \frac{x^2}{4}}.$$

*We then compute the slope,*

$$\frac{dy}{dx} = \frac{d}{dx}\left(2 - \frac{x^2}{4}\right)^{\frac{1}{2}} = \frac{-\frac{x}{2}}{2\sqrt{2 - \frac{x^2}{4}}} = \frac{-x}{4\sqrt{2 - \frac{x^2}{4}}}.$$

*At $x = -2$, we have*

$$\frac{dy}{dx}\,|_{x=-2} = \frac{2}{4\sqrt{2 - \frac{(-2)^2}{4}}} = \frac{1}{2}.$$

*So, we have a line of slope $\frac{1}{2}$ and it passes through $(-2, 1)$ – point-slope form yields*

$$y - 1 = \frac{1}{2}(x + 2) \implies y = \frac{1}{2}x + 2,$$

*as expected.*

## 10.5 A Review: Planes

A plane in space is identified by a point on the plane and its normal vector, i.e. a vector perpendicular to the plane. We use a property of the dot product to illustrate this orthogonality.

Let $\boldsymbol{n} = A\boldsymbol{i} + B\boldsymbol{j} + C\boldsymbol{k}$ be the normal vector, and $P_0 = (x_0, y_0, z_0)$ be a point on the plane. To find the equation of the plane, first, consider any other points on the plane $P = (x, y, z)$. The line between $P$ and $P_0$ can be represented by the vector

$$P - P_0 = \langle x - x_0, y - y_0, z - z_0 \rangle.$$

By orthogonality, we must have

$$\boldsymbol{n} \cdot (P - P_0) = 0.$$

This condition can be written out explicitly as

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

where $(x, y, z)$ represents all points in the plane, and $(x_0, y_0, z_0)$ is one sample point. Compare this to the **normal form of a line**

$$A(x - x_0) + B(y - y_0) = 0$$

in one dimension, that is, you have to identify a point and the normal vector of the line to fully characterize the line. Here, we are merely upgrading with one more dimension. Altogether, what we need for the equation of a plane is

1. A point on the plane.
2. The normal of the plane.

One may argue that in the plane (2 dimension), you need two points (0 dimension) to determine the equation of a line (1 dimension). Indeed, points are the building blocks of one dimension. However, in a 3D space, lines are the building blocks as they are 1 dimensional beings. In fact, knowing two lines automatically give us a plane, via the cross product. If you need a review of this, please consult your textbook on Section 2.5 in Volume 3.

## 10.6 Tangent Planes

Tangent lines to level curves $f(x, y) = c$ at $(x_0, y_0)$ is given by

$$\nabla f(x_0, y_0) \cdot \langle x - x_0, y - y_0 \rangle = f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0) = 0$$

since we have proved that the gradient vector is always perpendicular to the level curves (and hence the tangent line).

In a similar fashion, tangent planes to level surfaces $f(x, y, z) = c$ at $(x_0, y_0, z_0)$ is given by

$$\nabla f(x_0, y_0, z_0) \cdot \langle x - x_0, y - y_0, z - z_0 \rangle = f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0) + f_z(x_0, y_0)(z - z_0) = 0.$$

This result still came from the very fact that the directional derivative at $(x_0, y_0, z_0)$ on a level surface is zero (though it is a little harder to visualize than tangent lines to level curves). This means the gradient $\nabla f(x_0, y_0, z_0)$ is perpendicular to the level surface, and hence the tangent plane at the same point. The points that live in this plane form the vector $\langle x - x_0, y - y_0, z - z_0 \rangle$. If $\nabla f(x_0, y_0, z_0)$ is tangent to this vector, i.e.

$$\nabla f(x_0, y_0, z_0) \cdot \langle x - x_0, y - y_0, z - z_0 \rangle = 0,$$

then it is parallel to, $\boldsymbol{n}$, the **normal vector** of the plane, which means

$$\boldsymbol{n} = k \nabla f(x_0, y_0, z_0)$$

for some nonzero constant $k$. Thus, the equation of the plane satisfies

$$\boldsymbol{n} \cdot \langle x - x_0, y - y_0, z - z_0 \rangle = k \nabla f\left(x_0, y_0, z_0\right) \cdot \langle x - x_0, y - y_0, z - z_0 \rangle = 0$$

where we see that it doesn't really matter what $k$ is. This implies that the tangent plane to a **level surface** only requires

1. A point $(x_0, y_0, z_0)$.
2. Gradient of the function at $(x_0, y_0, z_0)$, i.e. $\nabla f\left(x_0, y_0, z_0\right)$.

**Example 10.3.** Let's confirm that the tangent plane to the northern hemisphere of radius 1, given by $f\left(x, y\right) = \sqrt{1 - x^2 - y^2}$, at the north pole is given by $z = 1$, using the level surface approach.

**Solution 10.1.** Consider $h\left(x, y, z\right) = x^2 + y^2 + z^2$. Then $f\left(x, y\right) = \sqrt{1 - x^2 - y^2}$ can be viewed as the 1-level surface of the function $h$, or more mathematically,

$$f\left(x\right) = \sqrt{1 - x^2 - y^2} \iff h\left(x, y, z\right) = x^2 + y^2 + z^2 = 1$$

So, in order to find the expression of the tangent plane at $(x, y, z) = (0, 0, 1)$, we need its normal vector. By previous results, we know that the gradient of $h$ is always perpendicular to its level surface, and thus is a constant multiple of sought normal vector (parallel to it).

To proceed with the gradient of $h$, we find that

$$\nabla h\left(0, 0, 1\right) = \langle 2x, 2y, 2z \rangle \mid_{(0,0,1)} = \langle 0, 0, 2 \rangle.$$

Therefore, the equation of the tangent plane is

$$\boldsymbol{n} \cdot \langle x - x_0, y - y_0, z - z_0 \rangle = 0 \implies \langle 0, 0, 2 \rangle \cdot \langle x - 0, y - 0, z - 1 \rangle = 0$$

Carrying out the dot product, we arrive at

$$2\left(z - 1\right) = 0 \implies z = 1,$$

as expected.

**Takeaway 10.1.** When tasked to find the expression of the tangent plane to a surface $z = f\left(x, y\right)$ at $(x_0, y_0)$, do

1. Express $z = f\left(x, y\right)$ as a level surface of some function $h\left(x, y, z\right)$.

**Remark 10.3.** *A typical one is*

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

*but more convenient expressions may be found to facilitate the next step: gradient calculation.*

2. Compute the gradient of $h$ at $(x_0, y_0, z_0)$, i.e.

$$\nabla h(x_0, y_0, z_0) = \langle h_x, h_y, h_z \rangle \mid_{(x_0, y_0, z_0)} .$$

3. Assemble the normal form of a (tangent) plane, by choosing the gradient vector $\nabla h(x_0, y_0, z_0)$ as the normal vector $\boldsymbol{n}$,

$$\boldsymbol{n} \cdot \langle x - x_0, y - y_0, z - z_0 \rangle = 0 \implies \nabla h(x_0, y_0, z_0) \cdot \langle x - x_0, y - y_0, z - z_0 \rangle = 0$$

**Theorem 10.1.** *Consider $h(x, y, z) = z - f(x, y)$, then $z = f(x, y)$ is the 0-level surface of $h$, and the tangent plane at $(x_0, y_0, z_0)$ where $z_0 = f(x_0, y_0)$ has normal vector*

$$\nabla h(x_0, y_0, z_0) = \langle -f_x, -f_y, 1 \rangle \mid_{(x_0, y_0, z_0)}$$

*and expression*

$$\langle -f_x, -f_y, 1 \rangle \mid_{(x_0, y_0, z_0)} \cdot \langle x - x_0, y - y_0, z - z_0 \rangle = 0$$

*which is exactly what we know before in Definition 6.2*

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

**Question 10.1.** Find three possible $h(x, y, z)$ such that the surface $f(x, y) = \sqrt{1 - x^2 - y^2}$ is a level surface of $h$. The level doesn't have to be 0 – it can be any level.

# 11 Lecture XI: An Introduction to Optimization in High Dimensions

Optimization problems are ubiquitous in our daily life. Vendors need to maximize sales based on variable and fixed costs. Civil engineers need to know how to minimize load of a beam. Statisticians estimate the parameters of a data set that yields the maximum likelihood. Computer scientists investigate algorithms and try to improve by minimizing the number of operations. All of these problems can be mathematically interpreted and cast as a mathematical optimization problem.

Problems in these scenarios often depend on more than one independent variable. The optimal result sometimes requires "all stars to align", meaning that multiple variables need to take certain value simultaneously to achieve optimum.

## 11.1 Key Definitions in Optimization in One Dimension

**Definition 11.1.** *(Local Extrema) Consider a closed interval* $[a-\delta, a+\delta]$ *in the domain of* $f(x)$.

1. *If* $f(a) \geq f(x)$*, for* $x \in [a-\delta, a+\delta]$*, then* $x = a$ *is a local maximum.*
2. *If* $f(a) \leq f(x)$*, for* $x \in [a-\delta, a+\delta]$*, then* $x = a$ *is a local minimum.*

The keyword here is the "a". As long as you find **one** neighborhood around the point of interest that has the above property, you then can qualify the point as a local extrema.

**Remark 11.1.** *So far,* $f$ *needs not be differentiable.*

If we happen to obtain a differentiable function, we know that if $x = a$ is a relative extrema, then we have as a necessary condition

$$f'(a) = 0.$$

**Theorem 11.1.** *(Fermat's Theorem, 1D Local Extrema) If a differentiable function* $f(x)$ *attains a local extrema at* $x = a$*, then* $f'(a) = 0$.

*Proof.* Suppose $x = a$ is a local maximum. This implies there exists a neighborhood $[a-\delta, a+\delta]$ such that

$$f(x) \leq f(a)$$

for all $x \in [a-\delta, a+\delta]$. Given that $f'(a)$ is well-defined, the derivative value is equal to the left and right limit of the difference quotient.

$$f'(a) = \lim_{x \to a^+} \frac{f(x) - f(a)}{x - a} \leq 0.$$

Meanwhile,

$$f'(a) = \lim_{x \to a^-} \frac{f(x) - f(a)}{x - a} \geq 0.$$

The only value $f'(a)$ can take is thus 0. $\square$

However, the **converse** is not always true, that is, having a zero derivative at a point does not imply that it is a relative extrema. This can be shown via the counterexample $f(x) = x^3$ at $x = 0$.

But, the points with zero derivatives are still worth studying, because they **may** be relative extremas, which is better than no information at all.

**Definition 11.2.** *(Critical points) An interior point of the domain of a function $f(x)$ where $f'(x) = 0$ or $f'(x)$ does not exist is a **critical point** of $f$.*

## 11.2 Optimization in Two Dimensions

**Definition 11.3.** *(Local Extrema) Consider **a** neighborhood $\mathcal{N}$ of a point $(x, y)$ in the domain of $f(x, y)$.*

1. *If $f(a, b) \geq f(x, y)$, for $(x, y) \in \mathcal{N}$, then $(x, y) = (a, b)$ is a local maximum.*
2. *If $f(a, b) \leq f(x, y)$, for $(x, y) \in \mathcal{N}$, then $(x, y) = (a, b)$ is a local minimum.*

**Theorem 11.2.** *(2D Local Extrema) If a differentiable function $f(x, y)$ attains a local extrema at $(x, y) = (a, b)$, then $\nabla f(a, b) = 0$.*

The proof considers the cross-sectional curves $f(x, b)$ and $f(a, y)$ and uses the 1D result directly.

**Question 11.1.** Local extrema gives us zero gradient at the point. How about directional derivatives at this point?

**Solution.** Directional derivative, in an arbitrary direction $\boldsymbol{u}$, at this point is

$$D_{\boldsymbol{u}} f(a, b) = \nabla f(a, b) \cdot \widehat{\boldsymbol{u}} = 0.$$

**Definition 11.4.** *A point $P$ is critical if $\nabla f(x, y) \mid_P = 0$ or undefined.*

## 11.3 Saddle Points and Examples

Now, the same questions (or logical pitfalls) still arise from the implication between local extrema and zero gradient. Will there be cases such that a zero gradient at a point yields neither a local max nor a local min, just like $f(x) = x^3$ at $x = 0$?

**Definition 11.5.** *A differentiable function $f(x, y)$ has a saddle point at a critical point $(a, b)$ if in every open disk centered at $(a, b)$, there are domain points $(x, y)$ where $f(x, y) > f(a, b)$ and domain points $(x, y)$ where $f(x, y) < f(a, b)$. The corresponding point $(a, b, f(a, b))$ on the surface $z = f(x, y)$ is called a saddle point of the surface.*

You may check this by simply plugging in nearby points to find functions value above and below $f(a, b)$. Another way would be picking a few directions and computing the directional derivative – you look for **different** signs of the directional derivative to qualify $(a, b)$ as a saddle.

## 11.4 Second Derivative Tests

The series of operations presented above seems quite cumbersome if the function is more complicated. In one dimension, we may further opt for second derivative tests, namely, **local maximum** is achieved when $f''(x) < 0$ (and local min otherwise). However, the second derivatives of a multivariable function are much more involved.

We here devise a set of rules that qualify the critical points.

**Theorem 11.3.** *Suppose that $f(x, y)$ and its first and second partial derivatives are continuous throughout a disk centered at $(a, b)$ and that $f_x(a, b) = f_y(a, b) = 0$. Then*

1. *$f$ has a **local maximum** at $(a, b)$ if $f_{xx} < 0$ and $f_{xx}f_{yy} - f_{xy}^2 > 0$ at $(a, b)$.*
2. *$f$ has a **local minimum** at $(a, b)$ if $f_{xx} > 0$ and $f_{xx}f_{yy} - f_{xy}^2 > 0$ at $(a, b)$.*
3. *$f$ has a **saddle point** at $(a, b)$ if $f_{xx}f_{yy} - f_{xy}^2 < 0$ at $(a, b)$.*
4. ***The test is inconclusive** if $f_{xx}f_{yy} - f_{xy}^2 = 0$ at $(a, b)$.*

The proof is based on Taylor expansion of multivariable functions, discussed in Section 14.9. You are welcome to consume it at your own leisure.

**Remark 11.2.** *Think about the execution of this theorem **algorithmically**. It seems that the **discriminant** $f_{xx}f_{yy} - f_{xy}^2$ is very important because all four criteria involve its sign. Suppose you now have a list of critical points (ones that give zero gradient), how do you write a program that utilizes the test above?*

**Algorithm 1** An Algorithmic View of the Second Derivative Test.

```
if f_{xx}*f_{yy} - f^2_{xy} = 0
return inconclusive
elseif f_{xx}*f_{yy} - f^2_{xy} > 0
if f_{xx} > 0
return local maximum
else
% f_{xx} is necessarily not equal to 0 here
% as the discriminant would otherwise be negative
return local minimum
end
else
return saddle point
end
```

**Example 11.1.** Confirm that $(0,2)$ is a local minimum of the function $f(x,y) = x^2 + y^2 - 4y + 9$ using the second derivative test.

**Solution 11.1.** We first compute the gradient of $f(x,y)$ and aim to find the critical points.

$$\nabla f(x,y) = \langle 2x, 2y-4 \rangle .$$

The critical point is $(0,2)$, indeed as promised. We continue with the second derivative test.

$$\Delta = f_{xx} f_{yy} - f_{xy}^2 = (2)(2) - 0^2 = 4 > 0.$$

The discriminant is a constant positive number, putting us in the first two criteria. Furthermore, $f_{xx} = 2 > 0$, putting us in the second criteria. Indeed, $(0,2)$ is a local minimum.

**Example 11.2.** Find the local extreme values of $f(x,y) = 3y^2 - 2y^3 - 3x^2 + 6xy$.

**Solution 11.2.**

$$\nabla f(x,y) = \left\langle -6x + 6y, 6y - 6y^2 + 6x \right\rangle .$$

We need to solve a nonlinear system of equations to find the critical points,

$$\begin{aligned} -6x + 6y &= 0 \\ 6y - 6y^2 + 6x &= 0 \end{aligned}$$

The first equation implies $x = y$, while using it in the second yields

$$6y - 6y^2 + 6y = 0 \implies 12y - 6y^2 = 0 \implies 6y(2-y) = 0$$

which gives $y = 0$ and $y = 2$ – these correspond to $x = 0$ and $x = 2$ respectively. Thus, we have two pairs of critical points

$$(0, 0), \quad (2, 2).$$

The discriminant is

$$\begin{aligned} \Delta(x, y) &= f_{xx} f_{yy} - f_{xy}^2 \\ &= (-6)(6 - 12y) - 6^2 \\ &= 72y - 72 \end{aligned}$$

- At $(0, 0)$, $\Delta(0, 0) = -72 < 0$, making $(0, 0)$ a saddle point (criterion 3).
- At $(2, 2)$, $\Delta(2, 2) = 144 - 72 = 72 > 0$, and $f_{xx}(2, 2) = -6 < 0$, making $(2, 2)$ a local maximum (criterion 1).

# 12 Lecture XII: Optimization Problems in Bounded and Unbounded Domains

## 12.1 Optimization over $\mathbb{R}^2$

Sometimes, the optimization problem may be posed without a constraint on the region. This is commonly called "free space optimization". The difficulty lies in a lack of boundary points to compare with so as to determine global extrema.

**Example 12.1.** Determine the local and global extreme values of $f(x,y) = x^2 + y^2 - 4y + 9$.

We find $\nabla f = (2x, 2y-4)$. Setting each component to zero, we have only one critical point

$$(x^*, y^*) = (0, 2).$$

Plugging this value in, we find

$$f(0,2) = 5.$$

However, upon further inspection of the function itself,

$$f(x,y) = x^2 + y^2 - 4y + 4 - 4 + 9 = x^2 + (y-2)^2 + 5 \geq 5$$

for all $(x,y) \in Dom(f)$. Thus, $(0,2)$ is a local minimum (in fact a global minimum).

**Remark 12.1.** *After obtaining the critical points, we must go through a series of mathematical operations to* ***qualify*** *them as local extremas or saddles.*

## 12.2 General Optimization

This section includes the algorithmic approach when we are posed with the optimization problem on a general closed and bounded domain $R$.

Relative extrema may occur on the interior, but they are not necessarily the absolute extrema. One must compare values achieved on the boundary of the domain with the relative extreme values to ultimately determine the general optima of a function.

1. List the **interior** points of $R$ where $f$ may have local maxima and minima and evaluate $f$ at these points. These are **critical points** of $f$.

2. List the **boundary** points of $R$ where $f$ has local maxima and minima and evaluate $f$ at these points.

**Remark 12.2.** *The boundary of $R$ is not necessarily a set of two points (in the case of 1D functions); it could be a curve or a line, represented by a function $f(x, y)$ or a parametrized curve $x = g(t)$ and $y = h(t)$.*

3. Look through the lists for the maximum and minimum values of $f$. These will be the absolute maximum and minimum values of $f$ on $R$.

**Example 12.2.** Find the absolute maximum and minimum values of

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

on the triangular region in the first quadrant bounded by the lines $x = 0$, $y = 0$ and $y = 9 - x$.

**Solution.** We proceed exactly as the above algorithm.

1. Interior

   Critical points lie at where the gradient vanishes or is undefined.

   $$\nabla f(x, y) = \langle 2 - 2x, 4 - 2y \rangle$$

   which implies

   $$(x^*, y^*) = (1, 2).$$

   The function value at this point is

   $$f(1, 2) = 7.$$

2. Boundary: note we have three pieces of the boundary.

   (a) Along $y = 0$, we have

   $$f(x, 0) = 2 + 2x - x^2, \quad 0 \leq x \leq 9.$$

   This function has a critical point at $x = 1$ that attains a value $f(1, 0) = 3$. At the endpoints, we have

   $$f(0, 0) = 2, \quad f(9, 0) = -61.$$

   (b) Along $x = 0$, we have

   $$f(0, y) = 2 + 4y - y^2, \quad 0 \leq y \leq 9.$$

   This function has a critical point at $y = 2$ that attains a value $f(0, 2) = 6$. At the endpoints, we have

   $$f(0, 0) = 2, \quad f(0, 9) = -43.$$

(c) Along the line $y = 9 - x$, we have

$$f(x, 9-x) = 2 + 2x + 4(9-x) - x^2 - (9-x)^2 = -43 + 16x - 2x^2.$$

This function has a critical point at $(x, y) = (4, 5)$ which then attains a value

$$f(4,5) = -11.$$

Altogether, we have

$$\begin{aligned} \text{(critical point) } f(1,2) &= 7 \\ f(0,0) &= 2 \\ f(1,0) &= 3 \\ f(9,0) &= -61 \\ f(0,9) &= -43 \\ f(0,2) &= 6 \\ f(4,5) &= -11 \end{aligned}$$

which we conclude

$$f(9,0) = -61$$

is an absolute minimum, and

$$f(1,2) = 7$$

is an absolute maximum.

# 13 Lecture XIII: Lagrange Multipliers

## 13.1 Motivation

Given a level curve $f(x, y) = c$, or a level surface $f(x, y, z) = c$, we learn before that the gradient of $f$ is always perpendicular to the level curve at any point $(x, y)$ (or $(x, y, z)$) in the domain of $f$. The property comes in very handy when an optimization problem has a constraint, typically described by another curve $g(x, y) = d$. However, before we start solving this type of problems, let's review what we know already.

1. Free space optimization: this is the type where we search for optimizers in the entire domain of the function $f$. No constraints.

2. Optimization on a bounded region: we search for optimizers in a bounded domain (typically a subset of $Dom(f)$). We are constrained to solutions that only lie in the bounded region.

3. Constrained optimization: this is even more restrictive than the previous case. We search for optimizers that must live on a certain curve rather than a region. In some sense, this problem should be easier to solve because we have shrunk down the search space from the entire domain to just a curve. Indeed, the technique to achieve this is simply brilliant – Lagrange Multipliers.

**Example 13.1.** Consider the circle $x^2+y^2 = 1$. Find the area of the largest rectangle inscribed in this circle.

**Solution.** We know that the four vertices must live on the circle. If we know one of them, we know all of them. But this fact restricts where we seek optimum. They cannot be **any** points in the space; whatever optimizer $(x, y)$ you find, the pair **has** to live on the circle, i.e., $x^2 + y^2 = 1$ is a constraint that must be satisfied.

Let $x$ and $y$ be the half width and half height of the rectangle. Then, the area of the rectangle is

$$A(x, y) = (2x)(2y) = 4xy.$$

This $A$ is called an **objective function**. It is the very function for which we seek optima. This problem can be written in the following form

$$\begin{aligned} \max_{Dom(A)} \quad & A(x, y) = 4xy \\ \text{subject to} \quad & x^2 + y^2 = 1. \end{aligned}$$

Now, how do we solve it? We know $x$ is related to $y$ via the circle. Shall we solve $y$ in terms of $x$ and substitute $y$ in the objective function? In fact, in the first quadrant, $y = \sqrt{1-x^2}$ without the need to worry about $\pm$ signs. Great. So, we have just "unconstrained" the problem, or, more mathematically, **reduced the dimension of the problem**. Just optimize the single variable function

$$A(x) = 4x\sqrt{1-x^2}, \quad 0 \le x \le 1.$$

It's a calculus 1 exercise. You carry it out.

The maximizing rectangle should be intuitively obvious – a square of side length $\sqrt{2}$ and thus maximizing area is 2. Don't believe it? Work out the calculus 1 exercise and show that $x = \frac{1}{\sqrt{2}}$ is a maximizer of $A(x)$ for $x \ge 0$.

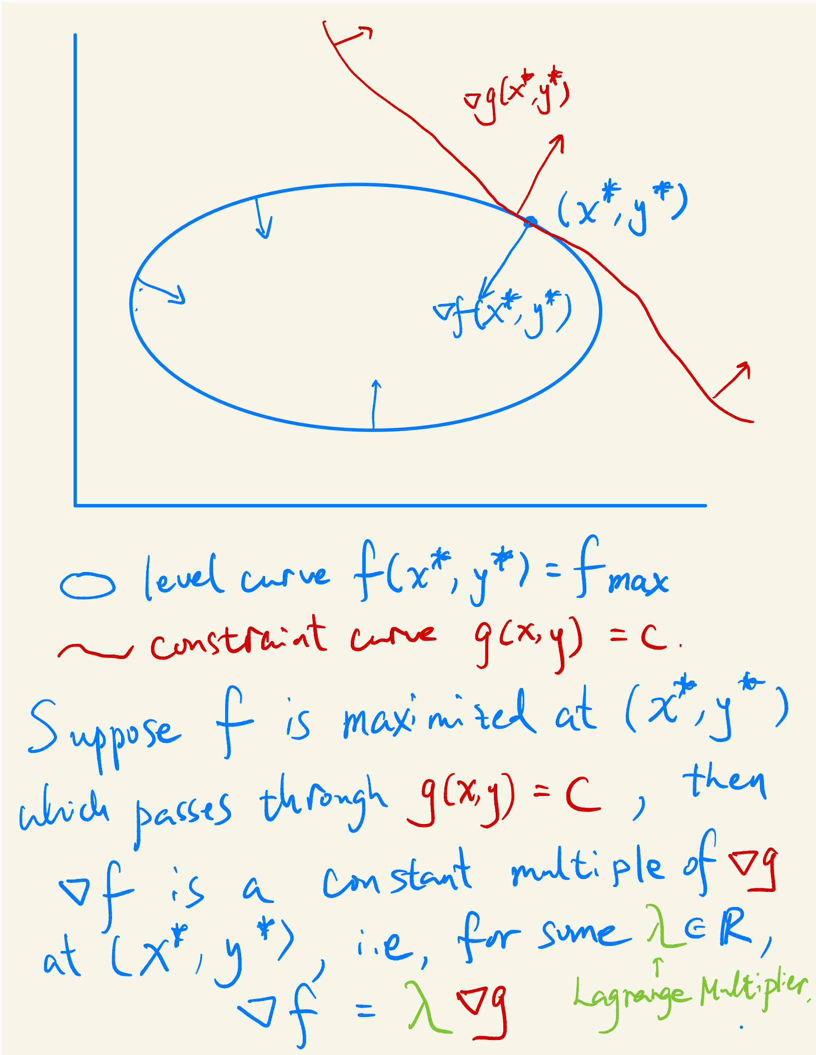
**Example 13.2.** Consider the ugly potato $x^4 + y^4 - x - y = 1$. Find the area of the largest rectangle inscribed in this circle.

**Solution.** The problem didn't seem to change much, except the constraint has changed from a circle to a weirdo. We have

$$\begin{aligned} \max_{Dom(A)} \quad & A(x,y) = 4xy \\ \text{subject to} \quad & x^4 + y^4 - x - y = 1. \end{aligned}$$

Now, can we still solve $y$ in terms of $x$ and do the **dimension reduction** trick? This equation is just awful to separate. We need something better.

### 13.2 The Lagrange Multiplier




One picture says it all. Suppose a function $f(x, y)$ achieves its maximum $f_{\max}$ at a point $(x^*, y^*)$. This point also passes through the constraint $g(x, y) = c$. Then, this very point sits on the **level curve** of $f$ and $g$ respectively. This implies that the gradient of $f$ is parallel to that of $g$, or in other words,

$$\nabla f = \lambda \nabla g$$

for some $\lambda$, to be determined. This number $\lambda$ is called **the Lagrange Multiplier**.

This equality between two gradients yields $n$ equations where $n$ is the number of independent variables in $f$ (or $g$). If we have $f(x, y)$ and $g(x, y)$, we have two equations

$$\begin{aligned}\frac{\partial f}{\partial x}\mid_{(x^*,y^*)} &= \lambda\frac{\partial g}{\partial x}\mid_{(x^*,y^*)};\\ \frac{\partial f}{\partial y}\mid_{(x^*,y^*)} &= \lambda\frac{\partial g}{\partial y}\mid_{(x^*,y^*)},\end{aligned}$$

which we use to find the optimizer $(x^*, y^*)$. But, $\lambda$ is the third unknown. No worries, the third equation is no other than the constrained equation

$$g(x^*, y^*) = c.$$

Three unknowns, three equations, we are ready to roll.

**Example 13.3.** (Inscribed Rectangle Revisited)

$$\begin{aligned}\max_{Dom(A)} A(x,y) &= 4xy\\ \text{subject to} x^2 + y^2 &= 1.\end{aligned}$$

So, we perform our solution technique using Lagrange Multipliers.

$$\begin{aligned}\frac{\partial f}{\partial x}\mid_{(x^*,y^*)} &= \lambda\frac{\partial g}{\partial x}\mid_{(x^*,y^*)},\\ \frac{\partial f}{\partial y}\mid_{(x^*,y^*)} &= \lambda\frac{\partial g}{\partial y}\mid_{(x^*,y^*)},\\ x^2 + y^2 &= 1\end{aligned}$$

Now, instead of writing $(x^*, y^*)$, we can drop the asterisks for notational simplicity.

Filling out the details, we have

$$\begin{aligned}4y &= \lambda(2x),\\ 4x &= \lambda(2y),\\ x^2 + y^2 &= 1.\end{aligned}$$

Looking at the first two equations, we notice several cases.

- $\lambda = 0$. Then, $x, y$ can be anything that live on the unit circle, which is NOT helpful.

  **Remark 13.1.** *A good note here also is that setting $\lambda = 0$ always gives you back the constraint equation.*

- $\lambda \neq 0$.

  We multiply the two equations to get

  $$16xy = \lambda^2 4xy.$$

  Now suppose $x \neq 0$ and $y \neq 0$ (they cannot be both zero, why?). This implies $\lambda^2 = 4 \implies \lambda = \pm 2$.

  - $\lambda = 2$

    $4y = 2(2x) \implies x = y$. Using this in the constrained equation, we have

    $$x = \pm\frac{1}{\sqrt{2}} \implies y = \pm\frac{1}{\sqrt{2}}.$$

  - $\lambda = -2$

    $4y = (-2)(2x) \implies x = -y$. Using this in the constrained equation, we have

    $$x = \pm\frac{1}{\sqrt{2}} \implies y = \mp\frac{1}{\sqrt{2}}$$

    (opposite sign).

  We evaluate these four points in the objective function and find

  $$f\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) = 2$$
  $$f\left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right) = 2$$
  $$f\left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) = -2$$
  $$f\left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right) = -2$$

  which confirms that $(x^*, y^*) = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$ is a maximizer.

**Example 13.4.** (Weirdo revisited)

$$\max_{x\geq 0, y\geq 0} f(x,y) = 4xy$$
$$\text{subject to} g(x,y) = x^4 + y^4 + 4x - 4y = 1$$

So, we perform our solution technique using Lagrange Multipliers.

$$
\begin{aligned}
\frac{\partial f}{\partial x} \mid_{(x^*,y^*)} &= \lambda \frac{\partial g}{\partial x} \mid_{(x^*,y^*)}, \\
\frac{\partial f}{\partial y} \mid_{(x^*,y^*)} &= \lambda \frac{\partial g}{\partial y} \mid_{(x^*,y^*)}, \\
x^4 + y^4 + x - y &= 1
\end{aligned}
$$

Now, instead of writing $(x^*, y^*)$, we can drop the asterisks for notational simplicity.

Filling out the details, we have

$$
\begin{aligned}
4y &= \lambda \left(4x^3 + 4\right), \\
4x &= \lambda \left(4y^3 - 4\right), \\
x^4 + y^4 + 4x - 4y &= 1
\end{aligned}
$$

Multiplying the first equation by $x$ and the second by $y$, we have

$$
\begin{aligned}
4xy &= \lambda \left(4x^4 + 4x\right), \\
4xy &= \lambda \left(4y^4 - 4y\right).
\end{aligned}
$$

Adding them yields

$$
\frac{8}{\lambda} xy = 4\left(x^4 + y^4\right) + 4x - 4y = 4 \implies \lambda = 2xy.
$$

Using this information in the first two equations, assuming $x \neq 0$ and $y \neq 0$, we have

$$
\begin{aligned}
4y = 2xy\left(4x^3 + 4\right) &\implies 2x^4 + 2x - 1 = 0 \\
4x = 2xy\left(4y^3 - 4\right) &\implies 2y^4 - 2y - 1 = 0
\end{aligned}
$$

Now what? We are ready to find $(x, y)$ but don't know how. This type of nonlinear equation calls for the Newton-Raphson Method, typically taught in the first course of **Numerical Analysis** (MATH 3257, C-term). We can only solve it numerically.

# Part II

# Integral Calculus in Multiple Dimensions

Integration is really just a fancy way of saying "adding up small parts in an organized way." In single-variable calculus, we figured out things like the "area under a curve" by adding up skinny vertical rectangles—this approximation is called a Riemann sum.

Now we're stepping things up. Our functions describe surfaces: $z = f(x, y)$. So instead of area under a curve, we're now looking at the volume under a surface. But before diving into the formulas, let's take a detour to your kitchen. Imagine you're chopping a potato.

Start by cutting the potato in half so that it can sit flat. That flat side represents the $xy$-plane. Now take your knife and slice the potato along its main axis, producing strips.

Each strip has a small volume—kind of like the thin rectangles from before, but now in 3D. If you slice carefully and don't lose any bits, then the sum of the volumes of all those little strips should give you the volume of the whole (half) potato. This is the idea behind a Riemann sum in two variables. Instead of just one direction of accumulation, we're adding up pieces that vary in both length and width—each contributing a little bit of volume.

We'll start by exploring how these two-dimensional Riemann sums work, what ideas carry over from single-variable calculus, and what new twists show up. Then we'll expand our mathematical toolbox by learning new coordinate systems—like polar and cylindrical—that help us tackle shapes and regions that are awkward to describe with plain old $x$ and $y$ (Cartesian coordinates).

## 14 Lecture XIV: Double Integrals over Rectangular Regions

### 14.1 Riemann Sum Revisited in Two Dimensions

The integral of a single-variable function $f(x)$ over a given region (interval) is the area under the curve over this region. The Riemann sum of a continuous function approximate this integral using the rectangles, namely,

$$\int_a^b f(x)\, dx = \lim_{|P| \to 0} \sum_{k=0}^{n} f(x_k)\, \Delta x_k$$

where the limit is taken as the number of subintervals goes to infinity, which is equivalent to shrinking the subinterval length to zero ($|P| \to 0$).

The integral of a two-variable function $f(x, y)$ over a given region is the volume under the surface over this region. We write it as

$$\iint_R f(x, y)\, dA = \iint_R f(x, y)\, dx dy.$$

The Riemann sum also undergoes a dimensional upgrade,

$$\iint_R f(x, y)\, dA = \lim_{|P| \to 0} \sum_{k=0}^{n} f(x_k, y_k)\, \Delta A_k$$

where $\Delta A_k$ is the area of a rectangle formed by changes in $x$ AND $y$.

However, the main complexity in double integral (compared to single integral) is that the region of integration may become more complicated. In a single definite integral, we are always dealing with the area under a curve over a fixed interval, i.e., $\int_a^b f(x)\, dx$. As an additional variable is added, we are able to describe various shapes in the plane using both $x$ and $y$. The region of integration thus has become far more complicated than mere intervals.

## 14.2 Iterated Integrals

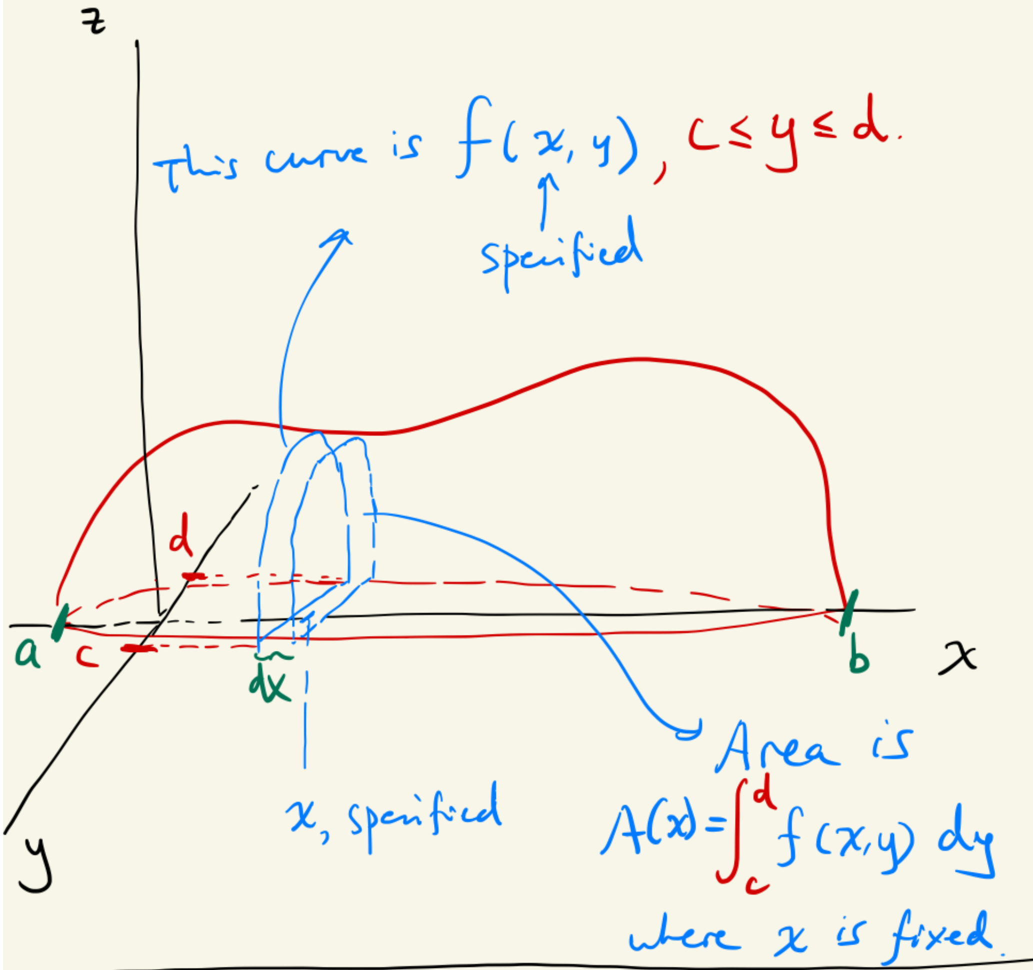
Though the Riemann sum formulation is all nice and theoretical, it does NOT provide us a concrete way of evaluating the integral. However, we now revisit the potato slicing "algorithm".

Suppose you are taking slices parallel to the $yz$-plane. Each slice is at a specific $x$ value, with thickness $dx$. The slice has a cross-sectional area, which is precisely the area under the curve $f(x, y)$ for a fixed $x$, across a range of $y$-values. In other words, at a particular $x$ value, the cross-sectional area is given by an integral over $y$ (just like in single-variable integral)

$$A(x) = \int_c^d f(x, y)\, dy$$

This slice of potato has thickness $dx$ and cross-sectional area $A(x)$, which implies that it has volume $A(x)\, dx$. How do we get the total volume of the potato? Add up all the slices,

$$V = \int_a^b A(x)\, dx = \int_a^b \left( \int_c^d f(x, y)\, dy \right) dx = \iint f(x, y)\, dA.$$




"Add up" all slices, i.e, integrate.

$$V = \int_a^b A(x)\,dx.$$

$$= \int_a^b \left( \int_c^d f(x,y)\,dy \right) dx$$

This type of integral is called an iterated integral, namely, integrate against one variable first and then the other. It involves precisely the idea of adding up cross-sectional area to find volume.

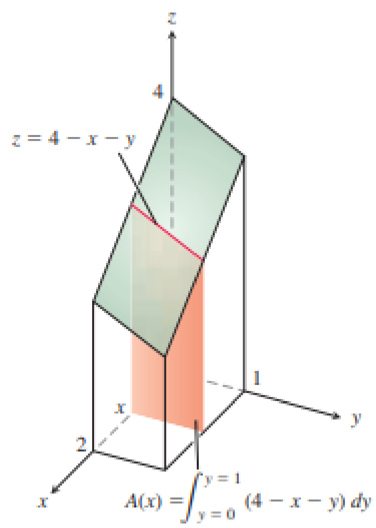
$$\int_c^d \int_a^b f(x,y)\,dxdy = \int_c^d \left(\int_a^b f(x,y)\,dx\right) dy = \int_c^d A(y)\,dy,$$

or

$$\int_a^b \int_c^d f(x,y)\,dydx = \int_a^b \left(\int_c^d f(x,y)\,dy\right) dx = \int_a^b A(x)\,dx.$$

Consider a triangular prism, with a right-triangle cross-section. The volume of this shape is straightforward – base triangle area times height/length. However, we can add up the volume in a different way.

**Example 14.1.** Compute the volume under the plane $z = 4 - x - y$ over the rectangular region $0 \le x \le 2$ and $0 \le y \le 1$.




**Solution 14.1.** First, we consider the cross-sectional area in the $yz$-plane, that is, consider an arbitrary $x$ value.

$$A(x) = \int_{y=0}^{y=1} (4 - x - y)\,dy.$$

One may wonder how this formula determines the cross-section area. Here, $x$ is a fixed, so consider it constant. Then, $z = 4 - x - y$ is merely the height of a line. Given the equation of the line, can we find the area under it? Of course, for each vertical slab spanning $[y, y+dy]$, its area is height times width, which yields $(4-x-y)\,dy$, i.e., the differential (of infinitesimal area). Adding across all $y$-values (here from 0 to 1), we retrieve the formula for $A(x)$.

As we've found the area formula for each $x$, we add up this area for all $x$'s (from 0 to 2),

$$\begin{aligned}
\int_{x=0}^{x=2} A(x)\,dx &= \int_{x=0}^{x=2} \left[\int_{y=0}^{y=1} (4-x-y)\,dy\right] dx \\
&= \int_{x=0}^{x=2} \left[4y - xy - \frac{y^2}{2}\right]_{y=0}^{y=1} dx \\
&= \int_{x=0}^{x=2} \left(\frac{7}{2} - x\right) dx \\
&= \left[\frac{7}{2}x - \frac{x^2}{2}\right]_{x=0}^{x=2} \\
&= 7 - 2 \\
&= 5
\end{aligned}$$

**Remark 14.1.** *This is called the* ***iterated integral*** *(or repeated integral).*

Now, let's try computing the crossectional area in the $xz$-plane by holding $y$ arbitrary (constant). Confirm that we have the same volume by swapping the order of integration.

$$A(y) = \int_{x=0}^{x=2} (4-x-y)\,dx.$$

We add up all the $y$'s and have

$$\begin{aligned}
\int_{y=0}^{y=1} A(y)\,dy &= \int_{y=0}^{y=1} \left(\int_{x=0}^{x=2} (4-x-y)\,dx\right) dy \\
&= \int_{y=0}^{y=1} \left[4x - \frac{x^2}{2} - xy\right]_{x=0}^{x=2} dy \\
&= \int_{y=0}^{y=1} [6-2y]\,dy \\
&= \left[6y - y^2\right]_{y=0}^{y=1} \\
&= 6 - 1 \\
&= 5
\end{aligned}$$

as expected.

# 15 Lecture XV: Baby Fubini and More Examples of Iterated Integrals

Last class, we introduced the double integral as a Riemann sum that approximates the **volume** under a surface $z = f(x, y)$. We looked at a specific domain of integration, *rectangles*, often described by the inequalities

$$D = \{(x, y) : a \leq x \leq b, \quad c \leq y \leq d\}.$$

We found a systematic way to compute the integral

$$V = \iint_D f(x, y)\, dA$$

using iterated integrals,

$$V = \int_c^d \int_a^b f(x, y)\, dxdy = \int_a^b \int_c^d f(x, y)\, dydx.$$

This statement suggests that we simply evaluate one integral against one variable after another, and it does NOT matter which variable we go with first. This result can be rigorously proved and is described in what I coined "baby Fubini Theorem".

## 15.1 Baby Fubini Theorem

**Theorem 15.1.** *(Baby Fubini) If $f(x, y)$ is continuous throughout the rectangular region $R:\ a \leq x \leq b$, $c \leq y \leq d$, then*

$$\iint_R f(x, y)\, dA = \int_c^d \int_a^b f(x, y)\, dxdy = \int_a^b \int_c^d f(x, y)\, dydx.$$

Here is an example to help you remember what was done last class.

**Example 15.1.** Find the volume of the paraboloid $z = 5 - x^2 - y^2$ over the region $0 \leq x \leq 1$ and $0 \leq y \leq 2$.

$$\begin{aligned}
\iint_R \left(5 - x^2 - y^2\right) dA &= \int_0^1 \int_0^2 \left(5 - x^2 - y^2\right) dydx \\
&= \int_0^1 \left[5y - yx^2 - \frac{y^3}{3}\right]_0^2 dx \\
&= \int_0^1 \left[10 - 2x^2 - \frac{8}{3}\right] dx
\end{aligned}$$

$$= \int_0^1 \left[\frac{22}{3} - 2x^2\right] dx$$
$$= \left[\frac{22}{3}x - \frac{2}{3}x^3\right]_0^1$$
$$= \frac{20}{3}.$$

We confirm that via Fubini's theorem,

$$\begin{aligned}
\int_0^2 \int_0^1 \left(5 - x^2 - y^2\right) dxdy &= \int_0^2 \left[5x - \frac{x^3}{3} - xy^2\right]_0^1 dy \\
&= \int_0^2 \left[\frac{14}{3} - y^2\right] dy \\
&= \left[\frac{14}{3}y - \frac{y^3}{3}\right]_0^2 \\
&= \frac{28}{3} - \frac{8}{3} \\
&= \frac{20}{3}.
\end{aligned}$$

We note further that if $f(x, y) = g(x)\,h(y)$ is a separable function, then we can further simplify the Baby Fubini formula.

**Corollary 15.1.** *Note that if $f(x, y) = g(x)\,h(y)$, meaning that $f$ can be decomposed into two functions that depend only on $x$ and $y$ respectively, then Fubini's theorem further reduces to*

$$\int_c^d \int_a^b g(x)\,h(y)\,dxdy = \int_c^d h(y) \left(\int_a^b g(x)\,dx\right) dy = \left(\int_a^b g(x)\,dx\right)\left(\int_c^d h(y)\,dy\right),$$

*i.e., a product of two single-variable integrals.*

## 15.2 Applications of Baby Fubini and Examples

Below, we study more examples of double integrals over rectangles.

**Example 15.2.** $\int_1^2 \int_0^4 2xydydx$.

**Solution 15.1.** First, we note that the integrand is separable, namely $2xy = g(x)\,h(y)$ where $g(x) = 2x$ and $h(y) = y$. The region of integration is a rectangle since the bounds of integration are constants. Therefore, by Fubini's theorem,

$$\int_1^2 \int_0^4 2xydydx = \left(\int_1^2 2xdx\right)\left(\int_0^4 ydy\right)$$

$$
\begin{aligned}
&= \left[x^2\right]_1^2 \left[\frac{y^2}{2}\right]_1^4 \\
&= \left[2^2 - 1^1\right] \times \frac{1}{2}\left[4^2 - 1^1\right] \\
&= 3 \times \frac{1}{2} \times 15 \\
&= \frac{45}{2}.
\end{aligned}
$$

**Example 15.3.** $\int_1^4 \int_0^4 \left(\frac{x}{2} + \sqrt{y}\right) dxdy.$

**Solution 15.2.** Even though the region of integration is still a rectangle, we note that the integrand is no longer separable into a product of two functions of single variable. Therefore, we treat it as an iterated integral, by first integrating against $x$ (and holding $y$ constant).

$$
\begin{aligned}
\int_1^4 \int_0^4 \left(\frac{x}{2} + \sqrt{y}\right) dxdy &= \int_1^4 \left(\int_0^4 \left(\frac{x}{2} + \sqrt{y}\right) dx\right) dy \\
&\overset{y \text{ constant}}{=} \int_1^4 \left[\frac{x^2}{4} + x\sqrt{y}\right]_{x=0}^{x=4} dy \\
&= \int_1^4 \left[\frac{4^2}{4} + 4\sqrt{y} - 0 - 0\right] dy \\
&= 4\int_1^4 \left[1 + \sqrt{y}\right] dy \\
&= 4\left[y + \frac{2}{3}y^{\frac{3}{2}}\right]_{y=1}^{y=4} \\
&= 4\left[4 + \frac{2}{3}4^{\frac{3}{2}} - 1 - \frac{2}{3}\right] \\
&= 4\left[4 + \frac{2}{3} \times 8 - 1 - \frac{2}{3}\right] \\
&= 4\left[3 + \frac{2}{3} \times 7\right] \\
&= 4\left[3 + \frac{14}{3}\right] \\
&= 4\left[\frac{23}{3}\right] \\
&= \frac{92}{3}
\end{aligned}
$$

The reason why all the algebra is written out is to encourage you to calculate first more slowly, and then speed up.

Fubini's theorem may not appear powerful when our regions are rectangles, since the order of integration did not really simplify the problem. However, it becomes a powerful tool when integrating over a general non-rectangular region. More work is to be done to convert the region of integration from a given perspective into a new perspective.

It is also possible to do iterated integrals when the region of integration is NOT a rectangle. The following serves as a nice introduction to what's to follow next class.

**Example 15.4** (Important)**.** $\int_0^1 \int_x^{2x} 2xydydx$.

The region of integration here is the triangular region between the two lines $y = x$ and $y = 2x$, and the vertical line $x = 1$. First, note that the inner integral is against $y$, so $x$ is held constant, which can be moved outside.

$$\begin{aligned}
\int_0^1 \int_x^{2x} xydydx &= \int_0^1 x \left( \int_x^{2x} 2ydy \right) dx \\
&= \int_0^1 x \left[ y^2 \right]_{y=x}^{y=2x} dx \\
&= \int_0^1 x \left[ (2x)^2 - x^2 \right] dx \\
&= 3 \int_0^1 x^3 dx \\
&= 3 \left[ \frac{x^4}{4} \right]_0^1 \\
&= \frac{3}{4}.
\end{aligned}$$

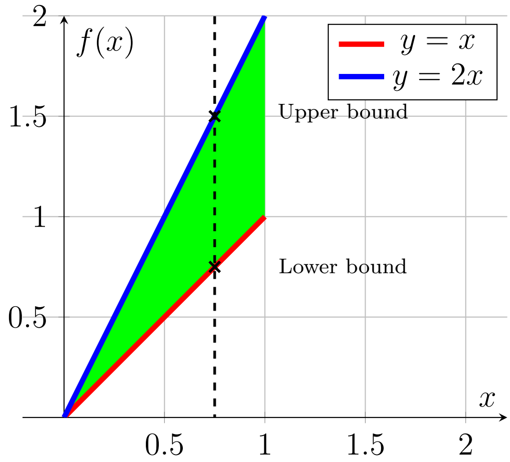



Figure 8: Region of integration for Example 15.4, $\left\{(x,y) \in \mathbb{R}^2 : {}^{x \le y \le 2x}_{0 \le x \le 1}\right\}$. The vertical line here is to the $x$-needle that first enters the region at the curve $y = x$ (lower bound) and exits the region at the curve $y = 2x$ (upper bound).

## 15.3 An Introduction to The Needle Method

This is the method we will use continuously for the rest of the term. Please pay great attention to the principles behind this method as it will simplify numerous integrals over general regions, in various coordinate systems.

From the last example, we saw that the integration over a triangular region is just another application of the iterated integral. In practice, we sometimes are only provided the sketch of a region, for which we must **identify** the bounds of integration for each variable. Notice that the bounds for $y$ in the last problem **depends** on $x$ because it is bounded between two lines (and eventually in harder problems, curves).

We identify the lower and upper curve by inserting a needle parallel to the $y$-axis, namely, a particular $x$-value, coined as the "$x$-needle", and poking it through the region from the bottom. The curve at which the needle enters the region is the **lower bound**, and the curve at which it exits is the **upper bound**.

**Takeaway 15.1** (Important)**.** The variable that has a function upper and lower bounds is to be integrated **first**.

**Example 15.5.** Consider the sketch in Figure 9. We identify that the lower bound is the curve $y = \sqrt{x}$ and the upper bound is the flat line $y = 1$. Therefore, we write, in this order, the description of the region $D$:

$$D = \left\{(x,y) \in \mathbb{R}^2 : {}^{\sqrt{x} \le y \le 1}_{0 \le x \le 1}\right\}$$

How did we find the range of values for $x$? We set the two curves equal and find that $x = 1$ and $x = 0$ are the two points of intersection.

Therefore, if one is to integrate a function $z = f(x, y)$ over this region, we will write

$$\iint_D f(x,y)\, dA = \int_0^1 \int_{\sqrt{x}}^1 f(x,y)\, dydx$$

noting that the order of integration is against $y$ first since it has **function** bounds.

**Example 15.6.** Consider the sketch in Figure 10. We identify that the lower bound is the curve $y = x^2$ and the upper bound is the line $y = x$. Therefore, we write, in this order, the description of the region $D$:

$$D = \left\{(x,y) \in \mathbb{R}^2 : {}^{x^2 \le y \le x}_{0 \le x \le 1}\right\}$$

How did we find the range of values for $x$? We set the two curves equal and find that $x = 1$ and $x = 0$ are the two points of intersection.

Therefore, if one is to integrate a function $z = f(x, y)$ over this region, we will write

$$\iint_D f(x,y)\, dA = \int_0^1 \int_{x^2}^x f(x,y)\, dydx$$

noting that the order of integration is against $y$ first since it has **function** bounds.

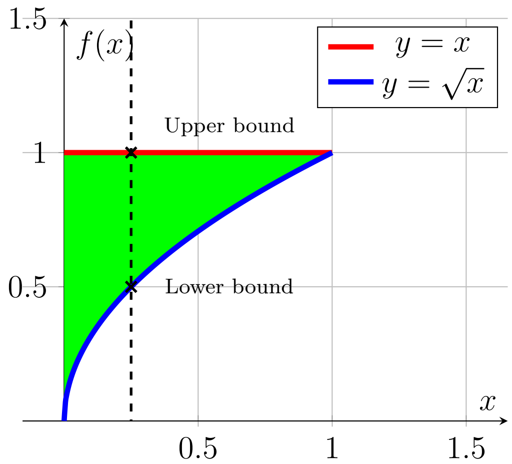



Figure 9: Region of integration for Example 15.5, $\left\{(x,y) \in \mathbb{R}^2 : {}^{\sqrt{x} \le y \le 1}_{0 \le x \le 1}\right\}$. The vertical line here is to the $x$-needle that first enters the region at the curve $y = \sqrt{x}$ (lower bound) and exits the region at the curve $y = 1$ (upper bound).

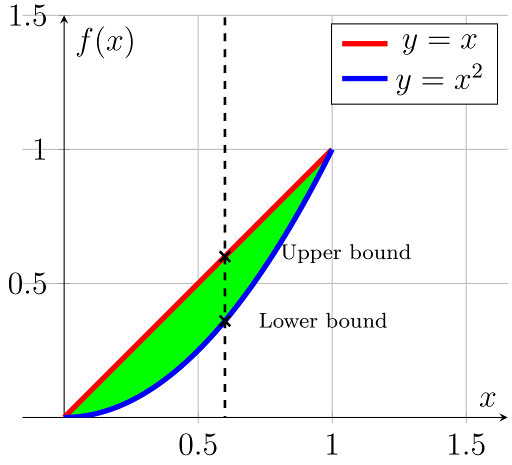



Figure 10: Region of integration for Example 15.6, $\left\{(x,y) \in \mathbb{R}^2 : \substack{x^2 \leq y \leq x \\ 0 \leq x \leq 1}\right\}$. The vertical line here is to the $x$-needle that first enters the region at the curve $y = x^2$ (lower bound) and exits the region at the curve $y = x$ (upper bound).

# 16 Lecture XVI: Papa Fubini and Double Integrals over General Regions

## 16.1 Solvable Examples of Double Integrals with Function Bounds

**Example 16.1** (Important Worked Example)**.** Consider the region depicted in Figure 9. Find the volume under the surface $f(x, y) = 2xy$ over this region.

**Solution.** We write out the domain again just for completeness.

$$D = \left\{ (x, y) \in \mathbb{R}^2 : {\scriptstyle \sqrt{x} \le y \le 1 \atop 0 \le x \le 1} \right\}.$$

Even though $f$ is separable, we CANNOT apply Baby Fubini because the region of integration is NOT a rectangle. Instead, we stick with iterated integration, and write the double integral as

$$\begin{aligned}
\iint_D f(x, y)\, dA &= \int_0^1 \int_{\sqrt{x}}^1 2xy dy dx \\
&\overset{x \text{ is a constant under } y}{=} \int_0^1 x \left( \int_{\sqrt{x}}^1 2y dy \right) dx \\
&= \int_0^1 x \left[ y^2 \right]_{y=\sqrt{x}}^{y=1} dx \\
&= \int_0^1 x \left[ 1 - x \right] dx \\
&= \int_0^1 \left( x - x^2 \right) dx \\
&= \left[ \frac{x^2}{2} - \frac{x^3}{3} \right]_{x=0}^{x=1} \\
&= \frac{1}{6}
\end{aligned}$$

## 16.2 Motivation for Double Integrals over General Regions

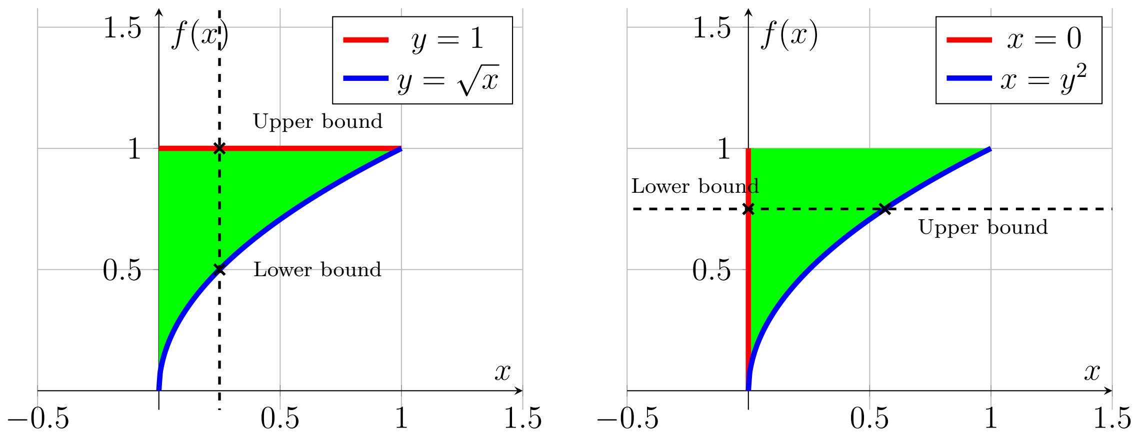
In Example 15.4 from last class and Example 16.1 above, we are blessed with the fact that the anti-derivative of $2y$ is very easy to find. What if it is impossible to anti-differentiate, such as,

$$\int_0^1 \int_{\sqrt{x}}^1 \sqrt{y^3 + 1} dy dx. \tag{16.1}$$

This is awful because there is no way to find the antiderivative of $\sqrt{y^3+1}$ against $y$. However, if we happen to find a way to integrate against $x$ first, then this difficulty goes away because in the eyes of $x$, $y$ is just a constant. Is there a general rule such that we can swap the order of integration? What price do we have to pay?

Fubini's theorem may not appear powerful when our regions are rectangles, since the order of integration did not really simplify the problem. However, it becomes a powerful tool when integrating over a general non-rectangular region. More work is to be done to convert the region of integration (see Figure 9) from a given perspective into a new perspective.

The needle method now makes its second appearance, and unleashes its full power. In order to switch the order of integration, here, from $y$ to $x$, we must now select a "$y$-needle", that is, a particular **horizontal** line parallel to the $x$-axis. Send the $y$-needle through the region, and now identify the curves at which the needle enters and exits, as the lower and upper bound, respectively.




(a) $x$-needle (b) $y$-needle

Figure 11: Region of integration for Example 16.2, $\left\{(x,y) \in \mathbb{R}^2 : {\sqrt{x} \le y \le 1 \atop 0 \le x \le 1}\right\}$. The vertical line here is to the $x$-needle that first enters the region at the curve $y = \sqrt{x}$ (lower bound) and exits the region at the curve $y = 1$ (upper bound).

**Example 16.2.** (**Identifying the region of integration and swapping the order**)

$$\int_0^1 \int_{\sqrt{x}}^1 \sqrt{y^3+1}\, dy\, dx.$$

The region described here is $0 \le x \le 1$, $\sqrt{x} \le y \le 1$—a boomerang-shaped slice above the curve $y = \sqrt{x}$ (see Figure 11a). At first glance, this integral poses a challenge: integrating $\sqrt{y^3+1}$ with respect to $y$ is not straightforward.

So, game over?

Never.

To make progress, we swap the order of integration. To do so, we reinterpret the region in terms of horizontal slices—what we call "$y$-needles." As shown in Figure 11b, the needle enters the region at $x = 0$ and exits at $x = y^2$. This reframes the region as:

$$\left\{(x,y) \in \mathbb{R}^2 : \substack{0 \le x \le y^2 \\ 0 \le y \le 1}\right\}$$

with $x$ now bounded above by the curve $x = y^2$ (formerly $y = \sqrt{x}$).

Now apply Fubini's theorem (swap the order of integration with the added care above in reframing the region of integration). Since the integrand $\sqrt{y^3+1}$ is independent of $x$, it factors out of the inner integral:

$$\begin{aligned}
\int_0^1 \int_{\sqrt{x}}^1 \sqrt{y^3+1}\, dy\, dx &= \int_0^1 \int_0^{y^2} \sqrt{y^3+1}\, dx\, dy \\
&= \int_0^1 \sqrt{y^3+1} \left( \int_0^{y^2} dx \right) dy \\
&= \int_0^1 \sqrt{y^3+1} \left( [x]_{x=0}^{x=y^2} \right) dy \\
&= \int_0^1 y^2 \sqrt{y^3+1}\, dy.
\end{aligned}$$

Use the substitution $u = y^3 + 1$ to finish:

$$\frac{1}{3} \int_1^2 \sqrt{u}\, du = \frac{1}{3} \left[ \frac{2}{3} u^{3/2} \right]_1^2 = \frac{2}{9}(2\sqrt{2} - 1).$$

**Theorem 16.1.** *(Fubini, Strong Form, "Papa Fubini") Let $f(x,y)$ be continuous on a region $R$.*

1. *If $R$ is defined by $a \le x \le b$, $g_1(x) \le y \le g_2(x)$, with $g_1$ and $g_2$ continuous on $[a,b]$, then*
$$\iint_R f(x,y)\, dA = \int_a^b \int_{g_1(x)}^{g_2(x)} f(x,y)\, dy dx.$$

2. *If $R$ is defined by $c \le y \le d$, $h_1(y) \le x \le h_2(y)$, with $h_1$ and $h_2$ continuous on $[c,d]$, then*
$$\iint_R f(x,y)\, dA = \int_c^d \int_{h_1(y)}^{h_2(y)} f(x,y)\, dx dy.$$

**Takeaway 16.1. Therefore, it is instrumental to describe the region of integration in two ways**. Sometimes, integrals may be easily evaluated if we change the order of integration, **by re-describing the region of integration**.Two reminders:

1. The outer integral **ALWAYS** has constant bounds.
2. The inner integral **MAY** have function bounds.

## 16.3 The Single Most Important Procedure for the Rest of the Term: the Needle Method

Procedures to determine bounds of integration:

For an integral $\iint_R f(x, y)\, dA$ with a given region $R$,

1. Sketch the region. Sketch the region. Sketch the region.
2. Choose the **the variable of interest** you want to integrate first (which may or may not lead to an easy integral, so this is done by experience learned from careful observations).
3. Take a "needle" and put it parallel to the axis of **the variable of interest**.
4. Poke this "needle" from the "bottom" of the region through the "top" of the region.
5. Identify the curves at entry and at exit. **These curves are written using that the variable of interest as the dependent variable, expressed in terms of all other variables.**

**Remark 16.1.** *In later sections, we deal with non-Cartesian variables, such as the radial distance $r$, and the radial angle $\theta$. These variables have* ***different*** *interpretations than $x$ and $y$. The needle method still applies – you just need to know which "axis" the new variables situate. For example, the radial distance variable always starts at the origin and points straight out, so one end of the needle sits at the origin and then you send it through the region of interest. The curves at entry and exit then will be in terms of the new variables. More on this later.*

## 16.4 Examples

**Example 16.3.** Calculate

$$\iint_R \frac{\sin x}{x} dA$$

where $R$ is the triangle in the $xy$-plane bounded by the $x$-axis, the line $y = x$ and the line $x = 1$.

**Solution.** We write out the region of integration in two ways: if we treat $x$ between constants, then we have

$$0 \leq y \leq x,$$
$$0 \leq x \leq 1;$$

on the other hand, if we treat $y$ between constants, then we have

$$y \leq x \leq 1,$$
$$0 \leq y \leq 1$$

Thus, the integral is formed in two ways:

$$\int_0^1 \int_0^x \frac{\sin x}{x} dydx, \quad \int_0^1 \int_y^1 \frac{\sin x}{x} dxdy.$$

Now, which way can we actually find the integral? In the second way, we need to find the antiderivative of $\frac{\sin x}{x}$. Good luck with that!

In the first way, we just need to integrate the integrand against $y$.

$$\begin{aligned}
\int_0^1 \int_0^x \frac{\sin x}{x} dydx &= \int_0^1 \frac{\sin x}{x} [y]_{y=0}^{y=x} dx \\
&= \int_0^1 [\sin x] dx \\
&= [-\cos(x)]_{x=0}^{x=1} \\
&= [-\cos(1) - (-\cos(0))] \\
&= 1 - \cos(1).
\end{aligned}$$

**Example 16.4.** Sketch the region of integration for the integral

$$\int_0^2 \int_{x^2}^{2x} (4x+2)\, dydx$$

and write an equivalent integral with the order of integration reversed.

**Example 16.5.** Find the volume of the wedgelike solid that lies beneath the surface $z = 16 - x^2 - y^2$ and above the region $R$ bounded by the curve $y = 2\sqrt{x}$, the line $y = 4x - 2$ and the $x$-axis.

**Remark 16.2.** *This example brings us back to what we are doing when computing a double integral – volume! The integrand (i.e., the function you are integrating) represents the height of the solid underneath the surface. The region of integration is the "base" of this solid (which can be of arbitrary shape, characterized by curves).*

## 16.5 Area of Irregular Regions Between Curves by Double Integration

In the previous section, we have developed ways to compute the double integral over general regions via Papa Fubini,

$$\iint_D f(x,y)\,dA$$

where $D$ can be described by

$$\begin{gathered} a \le x \le b, \\ h_1(x) \le y \le h_2(x), \end{gathered}$$

or

$$\begin{gathered} c \le y \le d, \\ g_1(y) \le x \le g_2(y). \end{gathered}$$

If $f(x,y) = 1$, **then we are merely computing the area of the region** $D$. It is an important application of double integrals. In fact, we can verify a lot of formulas for the area of certain shapes using this application.

**Example 16.6.** Consider a circle of radius $R$. We here show that the area of this circle is $\pi R^2$ using a double integral.

**Solution.** With radius $R$, we know

$$-R \le x \le R.$$

The circle has a parametric formula $x^2 + y^2 = R^2$. Thus, writing $y$ as a function of $x$, we just solve for it. Note that the upper semicircle is

$$y = \sqrt{R^2 - x^2}$$

while the lower semicircle is

$$y = -\sqrt{R^2 - x^2}.$$

One can also use **The Needle Method** – to find the variable bounds for $y$, send a needle parallel to the $y$-axis through the shape from the bottom. The needle enters the circle at the bottom semicircle and exits at the top.

Thus, the area of the circle is

$$\iint_R dA = \int_{-R}^{R} \int_{-\sqrt{R^2-x^2}}^{\sqrt{R^2-x^2}} dy\,dx$$

$$
\begin{aligned}
&= \int_{-R}^{R}\left(\int_{-\sqrt{R^2-x^2}}^{\sqrt{R^2-x^2}} 1dy\right)dx \\
&= \int_{-R}^{R}\left(\sqrt{R^2-x^2}-\left(-\sqrt{R^2-x^2}\right)\right)dx \\
&= 2\int_{-R}^{R}\sqrt{R^2-x^2}dx \\
&= 4\int_{0}^{R}\sqrt{R^2-x^2}dx
\end{aligned}
$$

where we used the fact that $\sqrt{R^2-x^2}$ is an even function.

Then, we need to revert back to Calculus II material – trigonometric substitution. Let

$$
x = R\sin(\theta), \quad dx = R\cos\theta d\theta
$$
$$
x = 0 \iff \theta = 0, \quad x = R \iff \theta = \frac{\pi}{2}
$$

and furthermore, the integrand is simplified to

$$
\sqrt{R^2-x^2} = \sqrt{R^2-r^2\sin^2\theta} = \sqrt{R^2\left(1-\sin^2\theta\right)} = \sqrt{R^2\cos^2\theta} = R\cos\theta.
$$

Then

$$
\begin{aligned}
\int_0^R \sqrt{R^2-x^2}dx &= \int_0^{\frac{\pi}{2}} \underbrace{R\cos\theta}_{\sqrt{R^2-x^2}} \cdot \underbrace{R\cos\theta d\theta}_{dx} \\
&= R^2\int_0^{\frac{\pi}{2}} \cos^2\theta d\theta.
\end{aligned}
$$

Now, we use double-angle formula

$$
\cos 2\theta = 2\cos^2\theta - 1 \implies \cos^2\theta = \frac{1}{2}\cos 2\theta + \frac{1}{2}.
$$

Altogether,

$$
\begin{aligned}
\int_0^{\frac{\pi}{2}} \cos^2\theta d\theta &= \frac{1}{2}\int_0^{\frac{\pi}{2}} (\cos 2\theta + 1)\, d\theta \\
&= \frac{1}{2}\left[\frac{\sin 2\theta}{2} + \theta\right]_0^{\frac{\pi}{2}}
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{1}{2}\left[\frac{\sin\pi}{2} + \frac{\pi}{2} - 0 - 0\right] \\
&= \frac{\pi}{4}
\end{aligned}
$$

Plugging this back into the integral, we have

$$
\begin{aligned}
\text{area of a circle of radius } R = \iint_R dA &= 4\int_0^r \sqrt{R^2 - x^2}dx \\
&= 4R^2 \int_0^{\frac{\pi}{2}} \cos^2\theta d\theta \\
&= 4R^2 \frac{\pi}{4} \\
&= \pi R^2
\end{aligned}
$$

**Example 16.7.** Find the area between two parabolas $y = 8 - x^2$ and $y = x^2$.

This is a problem from Calculus 2 that can be set up using double integrals.

**Solution.** A quick sketch informs us that $y = 8 - x^2$ is the upper parabola, while $y = x^2$ is the lower one. They intersect at

$$
8 - x^2 = x^2 \implies x = \pm 2, y = 4.
$$

Therefore, the region of integration is

$$
\begin{aligned}
x^2 \leq \ & y \leq 8 - x^2 \\
-2 \leq \ & x \leq 2.
\end{aligned}
$$

The area of a region $R$ is given by the double integral

$$
\begin{aligned}
\iint_R dA &= \int_{-2}^{2}\int_{x^2}^{8-x^2} dydx \\
&= \int_{-2}^{2}\left(\int_{x^2}^{8-x^2} 1dy\right) dx \\
&= \int_{-2}^{2}\left(8 - x^2 - x^2\right) dx \\
&= \int_{-2}^{2}\left(8 - 2x^2\right) dx \\
&\boxed{= 2\int_0^2 \left(8 - 2x^2\right) dx}
\end{aligned}
$$

$$
\begin{aligned}
&= 2\left[8x - \frac{2}{3}x^3\right]_0^2 \\
&= 2\left[16 - \frac{16}{3}\right] \\
&= \frac{64}{3}
\end{aligned}
$$

where at the boxed step we utilized the fact that the function $8 - 2x^2$ is an even function, so the integral about the symmetric interval $[-2, 2]$ is doubling the onesided integral. This is a very useful technique that makes the pesky algebra with negative numbers go away.

# 17 Lecture XVII: Double Integral in Polar Coordinates

In practice, many regions of integration follow certain symmetry, just like rectangles. For example, one may ask for an integral over a circle, which is perfectly symmetrical. Using Cartesian coordinates $x$ and $y$, we formulate the double integral

$$\iint_{circle} f(x,y)\,dA = \int_{-R}^{R}\int_{-\sqrt{R^2-x^2}}^{\sqrt{R^2-x^2}} f(x,y)\,dydx$$

where we have identified the upper and lower function as simply the upper and lower semicircles. This is not the prettiest integrals, and sometimes, with a slightly more complicated integrand $f(x,y)$, the integral is impossible to solve using standard techniques.

## 17.1 Area in Polar Coordinates: From Rectangles to Pizza Slices

Looking backwards at this integral, we didn't quite seem to use the perfect symmetry of the circle, namely, it is symmetric about **any** axis, in addition to the $x$ or $y$ axis (like the rectangle, but better). Using the Cartesian coordinates, that is, using little squares to add up areas, we can't quite add up to a circle perfectly using only rectangles. But we can definitely use pizza slices.

$$\begin{aligned} S_{\text{outer}} &= \frac{1}{2}(r+\Delta r)^2\,\Delta\theta \\ S_{\text{inner}} &= \frac{1}{2}r^2\Delta\theta \end{aligned}$$

so

$$\begin{aligned} \Delta A &= S_{\text{outer}} - S_{\text{inner}} \\ &= \frac{1}{2}\Delta\theta\left((r+\Delta r)^2 - r^2\right) \\ &= \frac{1}{2}\Delta\theta\left(r^2 + 2r\Delta r + \Delta r^2 - r^2\right) \\ &= r\Delta r\Delta\theta + \frac{1}{2}\Delta\theta\,(\Delta r)^2 \end{aligned}$$

Now, when $\Delta r$ is very small, the dominant term above is $r\Delta r\Delta\theta$. Thus, we claim that

$$\Delta A \approx r\Delta r\Delta\theta.$$

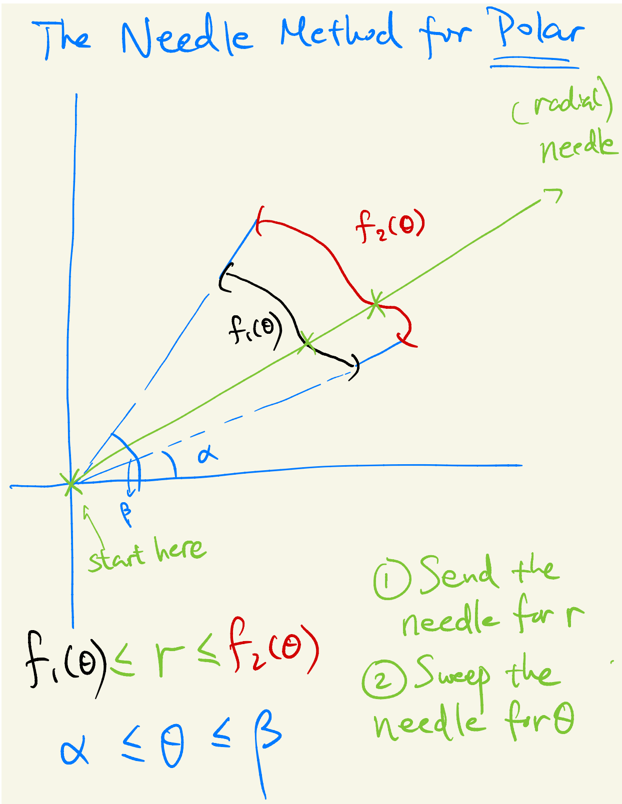
Thus, using the differentials to replace the $\Delta$'s, we have

$$\iint_R f(x,y)\,dA = \iint_R f(r,\theta)\,r dr d\theta.$$

**Takeaway 17.1** (Key Takeaway)**.** To change a double integral from Cartesian to Polar, do

1. $x = r\cos\theta$.
2. $y = r\sin\theta$.
3. $dA = dxdy = rdrd\theta$. The extra $r$ here is important. Why does it show up? Only when you internalize the derivation from a beautiful geometric argument presented above.
4. To determine the bounds of integration in the new variables, implement the needle method for polar coordinates (see sketch under Lecture Notes on Canvas: The-Needle-Method-Polar.pdf)

## 17.2 The Needle Method in Polar Coordinates




The Needle Method is to help determine the bounds of integration in the variables of interest. Here, since we have changed from Cartesian to Polar, we need to understand what exactly is meant by a needle.

The variable $r$, namely, radial distance, **always** starts at the origin. We say a "radial needle" is no other than shooting a ray from the origin outward through the region of interest. In other words, we have fixed an angle $\theta$ and let $r$ vary. The bounds of integration are determined by the same principle as in the Cartesian case: the curve at which the radial needle enters the region is the **lower bound** (resp. exits/upper bound).

**Example 17.1.** Find the area of the region $R$ that lies inside the circle $r = 2$ and outside the circle $r = 1$.

**Solution.** The region is the annulus between two concentric circles.

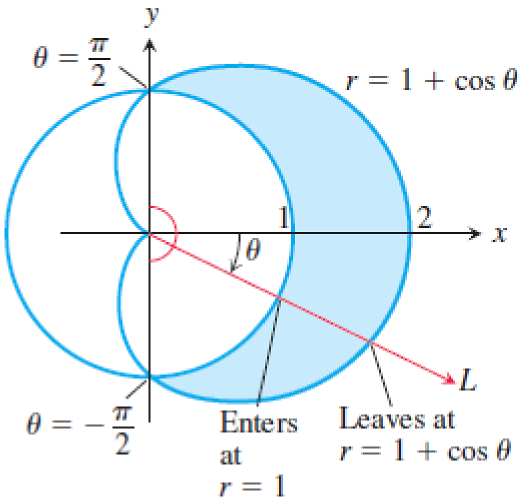
$$\begin{aligned}
\iint_R dA &= \int_0^{2\pi} \int_1^2 r dr d\theta \\
&= \frac{1}{2} \int_0^{2\pi} \left[r^2\right]_1^2 d\theta \\
&= \frac{1}{2} 2\pi \left[4 - 1\right] \\
&= 3\pi.
\end{aligned}$$

Sanity check (via common sense) tells us that the annulus has area

$$\pi \left(2\right)^2 - \pi \left(1\right)^2 = 3\pi.$$

# 18 Lecture XVIII: Examples for Double Integrals in Polar Coordinates

**Example 18.1.** Find the limits of integration for integrating $f(r,\theta)$ over the region $R$ that lies inside the cardioid $r = 1 + \cos(\theta)$ and outside the circle $r = 1$.




**Solution.**

$$\iint_R f(r,\theta)\, dA = \int_{-\pi/2}^{\pi/2} \int_1^{1+\cos(\theta)} f(r,\theta)\, r dr d\theta.$$

Once you identify the region, you now send first a radial needle (labeled as $L$ in the sketch) from the origin through the shape. This aims to obtain the bounds for the radius $r$. We note that first, the needle enters the region at the circle of radius 1, thus, the lower bound for $r$ is 1. The ray leaves the region at the cardioid $r = 1+\cos\theta$ which then naturally becomes the upper bound of $r$.

The way to determine the bounds for the polar angle $\theta$ is the following "sweeping method". Take a radial needle with $\theta = 0$, that is, a needle in the direction of the positive $x$-axis. Start sweeping counterclockwise through the $xy$-plane. Record the smallest and largest $\theta$ at which the needle still intersects the region. In this case, the lower bound is $-\pi/2$ and the upper bound is $\pi/2$.

The shaded region is described in polar coordinate

$$\begin{aligned} -\pi/2 \le \theta \le \pi/2 \\ 1 \le r \le 1 + \cos\theta \end{aligned}$$

The area of a region $R$ bounded by curves in Cartesian coordinates is given by

$$A = \iint_R dxdy,$$

where now we know the transformation

$$dxdy = rdrd\theta.$$

Therefore, in polar coordinates, we have

$$A = \iint_R rdrd\theta.$$

**Example 18.2.** Find the area of a circle with radius 2.

**Solution.** We know the answer is $4\pi$. Let's confirm it with polar coordinates.

$$\begin{aligned}
A &= \iint_R rdrd\theta \\
&= \int_0^{2\pi} \int_0^2 rdrd\theta \\
&\overset{\text{Baby Fubini}}{=} \left( \int_0^{2\pi} d\theta \right) \left( \int_0^2 rdr \right) \\
&= 2\pi \left( \frac{r^2}{2} \right) \Big|_{r=0}^{r=2} \\
&= 2\pi (2 - 0) \\
&= 4\pi.
\end{aligned}$$

**Remark 18.1.** *Note that here at the third equality, we separate the integrand, and in fact, the double integral into a product of two single-variable integrals. This is an application of the Corollary of Baby Fubini. The reason why the Theorem applies here, is because now, the variables $r$ and $\theta$ no longer have geometric meaning, and are just variables to integrate over. The fact that we have constant bounds means that we are integrating practically over a rectangular domain, which is the premise of Baby Fubini.*

The benefit of using polar coordiantes lies in not only dealing with region of integration that has special radial symmetry, but also transforming a difficult integral in Cartesian coordinates into an easy one in polar. The following problem showcases this power.

**Example 18.3.** Evaluate

$$\iint_R e^{x^2+y^2}\, dxdy$$

where $R$ is the semicircular region bounded by the $x$-axis and the curve $y = \sqrt{1-x^2}$ (that is, the upper semicircle).

**Solution.** This problem is in fact impossible since we do not have an elementary function that represents the antiderivative of $e^{x^2}$.

A good hint is that the moment we see the expression $x^2 + y^2$, we think polar. The relationship is

$$\begin{aligned} x &= r\cos\theta \\ y &= r\sin\theta \end{aligned}$$

and

$$dxdy = rdrd\theta.$$

Thus,

$$\begin{aligned} \iint_R e^{x^2+y^2}\, dxdy &= \iint_R e^{r^2} rdrd\theta \\ &= \int_0^{\pi} \int_0^1 re^{r^2}\, drd\theta \\ &= \left(\int_0^{\pi} d\theta\right)\left(\int_0^1 re^{r^2}\, dr\right) \\ &= \frac{\pi}{2}\int_0^1 e^u\, du \\ &= \frac{\pi}{2}(e-1) \end{aligned}$$

via integration by substitution.

**Example 18.4.** Evaluate

$$\int_0^1 \int_0^{\sqrt{1-x^2}} \left(x^2+y^2\right) dydx.$$

**Solution.** At first sight, we observe quite a naive integrand. It seems the logical next step is to integration the inner integral with respect to $y$ first,

$$\int_0^1 \int_0^{\sqrt{1-x^2}} \left(x^2+y^2\right) dydx = \int_0^1 \left(x^2\sqrt{1-x^2} + \frac{1}{3}\left(1-x^2\right)^{\frac{3}{2}}\right) dx.$$

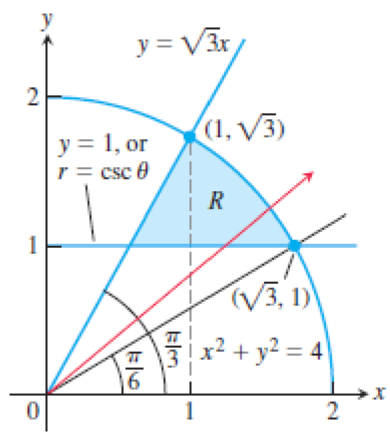
Now is the time to frown. Either integral requires a smart substitution, which is a lot of work. For those who would like some calculus II exercise, feel free to go on from here.

However, as hinted from the previous problem, the moment we see $x^2 + y^2$, we think polar. Indeed, the region of integration is no other than the sector in the first quadrant with radius 1. This sector, expressed in polar is $0 \le r \le 1$ and $0 \le \theta \le \pi/2$. Thus,

$$\begin{aligned}
\int_0^1 \int_0^{\sqrt{1-x^2}} \left(x^2 + y^2\right) dydx &= \int_0^{\pi/2} \int_0^1 r^2 r dr d\theta \\
&= \int_0^{\pi/2} \left[\frac{r^4}{4}\right]_0^1 d\theta \\
&= \frac{\pi}{2}\left(\frac{1}{4}\right) \\
&= \frac{\pi}{8}.
\end{aligned}$$

Lastly, we can even find the area of shapes bounded by radially symmetric curves and other non-symmetric curves. The idea is to be able to transform the region described by Cartesian into that expressed by polar.

**Example 18.5.** Consider the region enclosed by the circle $x^2 + y^2 = 4$, the line $y = 1$ and the line $y = \sqrt{3}x$. Find its area.




**Solution.** From the figure, we see clearly where the region is. The Cartesian description is given, but we need some help with the description in polar.

The range for polar angle $\theta$ is not terrible. If $y = 1$, then $x = \sqrt{3}$, so the lower side angle satisfies $\tan(\theta) = \frac{1}{\sqrt{3}}$ which implies $\theta = \frac{\pi}{6}$. Then, for the point at $\left(1, \sqrt{3}\right)$, the angle satisfies $\tan(\theta) = \sqrt{3}$ which implies $\theta = \pi/3$. Thus, the polar angle sweeps the range $\pi/6 \leq \theta \leq \pi/3$ (look at how the red arrow sweeps in the figure).

The range for radius is trickier. It certainly go up to 2 since the region reaches the skirt of the circle. However, the region is also bounded by the line $y = 1$. In polar coordinates, we know

$$y = r\sin\theta \implies r = \frac{y}{\sin\theta} = y\csc\theta.$$

Since we want to describe the line $y = 1$, we then must have $r = \csc(\theta)$. The range for the radius is

$$\csc\theta \leq r \leq 2.$$

Altogether,

$$\begin{aligned}
A &= \iint_R r dr d\theta \\
&= \int_{\pi/6}^{\pi/3} \int_{\csc\theta}^{2} r dr d\theta \\
&= \frac{1}{2}\int_{\pi/6}^{\pi/3} \left[4 - \csc^2(\theta)\right] d\theta \\
&= \frac{1}{2}\left[4\frac{\pi}{6} + \cot\left(\frac{\pi}{3}\right) - \cot\left(\frac{\pi}{6}\right)\right] \\
&= \frac{1}{2}\left[\frac{2}{3}\pi + \frac{1}{\sqrt{3}} - \sqrt{3}\right] \\
&= \boxed{\frac{1}{3}\pi + \frac{1}{2}\left(\frac{1}{\sqrt{3}} - \sqrt{3}\right)} \\
&= \frac{1}{3}\pi + \frac{1}{2}\frac{1-3}{\sqrt{3}} \\
&= \frac{1}{3}\pi - \frac{1}{\sqrt{3}} \\
&= \frac{1}{3}\left(\pi - \sqrt{3}\right)
\end{aligned}$$

The boxed answer is already satisfactory – the rest is algebraic brooming.

# 19 Lecture XIX: Masses and First Moments in One and Two Dimensions

In this lecture, we apply what we know about single and double integrals to find the center of mass of objects in 1D and 2D, respectively. The foundation is the fundamental understanding of how a Riemann sum is constructed. This is what Calculus is about.

## 19.1 Thin rod: Lines in 1D

**Example 19.1.** (Motivating example 1) $N$ point masses $m_1, m_2, \ldots, m_N$ sit on a weightless seesaw of length $L$. Their locations on the seesaw are described by the coordinates $0 < x_1 < x_2 < \cdots < x_N < L$. Where would you put the fulcrum so that the seesaw balances?

Suppose this fulcrum sits at $0 \leq x^* \leq L$ and everything balances. What "everything balances" means is that the seesaw experiences zero torque, or otherwise it will tip over to the side that experiences more torque.

For the first mass, it is at a displacement $x_1 - x^*$ from the fulcrum, and exerts its weight $m_1$. The amount of torque on the seesaw by the first mass is

$$\tau_1 = (x_1 - x^*)\, m_1.$$

Since the seesaw is in balance if the fulcrum is at $x^*$, we must have total torque equal to 0, that is,

$$\sum_{i=1}^{N} \tau_i = \sum_{i=1}^{N} (x_i - x^*)\, m_i = 0.$$

Now, we need to solve for $x^*$.

$$\sum_{i=1}^{N} (x_i - x^*)\, m_i = 0 \implies \sum_{i=1}^{N} x_i m_i - \sum_{i=1}^{N} x^* m_i = 0 \implies \sum_{i=1}^{N} x_i m_i - x^* \sum_{i=1}^{N} m_i = 0$$

which then ultimately yields

$$x^* = \frac{\sum_{i=1}^{N} x_i m_i}{\sum_{i=1}^{N} m_i},$$

where we observe that the denominator is **total** mass.

Now, suppose each object on the line has different density $\rho_i$ and a common with $\Delta x$, then, the center of mass changes to

$$x^* = \frac{\sum_{i=1}^{N} x_i \rho_i\,(x_i)\,\Delta x}{\sum_{i=1}^{N} \rho_i\,(x_i)\,\Delta x},$$

where we noted that the density of each object is directly related to where the object is.

Taking a limit as $N \to \infty$, we recognize that the numerator and the denominator are in fact, limit of Riemann sums,

$$x^* = \frac{\sum_{i=1}^{N} x_i \rho_i (x_i) \Delta x}{\sum_{i=1}^{N} \rho_i (x_i) \Delta x} \overset{N \to \infty}{\to} \frac{\int_a^b x\rho (x) \, dx}{\int_a^b \rho (x) \, dx}$$

This is the center of mass of a **continuous** thin rod with density $\rho(x)$. The numerator has a special name, called, the **first moment** of the rod; the denominator is called **total mass**.

**Remark 19.1.** *The derivation above is a demonstration from little discrete objects to a full continuous object. One should internalize the transformation, and take away the final formula using integrals.*

**Example 19.2.** Compute the total mass and the first moment of a thin rod with non-uniform density $\rho(x) = x(1-x)$ for $0 \leq x \leq 1$.

**Solution.**

$$\begin{aligned} x^* &= \frac{\int_0^1 x\rho (x) \, dx}{\int_0^1 \rho (x) \, dx} \\ &= \frac{\int_0^1 x \cdot x (1-x) \, dx}{\int_0^1 x(1-x) \, dx} \\ &= \frac{\int_0^1 (x^2 - x^3) \, dx}{\int_0^1 (x - x^2) \, dx} \\ &= \frac{\left[\frac{x^3}{3} - \frac{x^4}{4}\right]_{x=0}^{x=1}}{\left[\frac{x^2}{2} - \frac{x^3}{3}\right]_{x=0}^{x=1}} \\ &= \frac{\frac{1}{3} - \frac{1}{4}}{\frac{1}{2} - \frac{1}{3}} \\ &= \frac{\frac{1}{12}}{\frac{1}{6}} \\ &= \frac{1}{2} \end{aligned}$$

Does this answer make sense? Even though the rod has non-uniform density, the density function $\rho(x)$ is in fact symmetric about the vertical line $x = \frac{1}{2}$. Therefore, it makes sense that the center of mass is still at the center of the rod.

## 19.2 Lamina: Flat Plate in 2D

Now, we upgrade to two dimensions, where we consider flat plates. This is simply to assume that the masses in the last example are of the **area** $\Delta A$ (as opposed to length in 1D). Each element now is taken as the location $(x_{ij}, y_{ij})$, with which we associate it a density $\rho(x_{ij}, y_{ij})$.

Then, we can rewrite the fulcrum location as

$$x^* = \frac{\sum_{i=1}^{N}\sum_{j=1}^{M} x_{ij}\rho(x_{ij}, y_{ij})\,\Delta A}{\sum_{i=1}^{N}\sum_{j=1}^{M} \rho(x_{ij}, y_{ij})\,\Delta A} \overset{N,M\to\infty}{\to} \frac{\iint_R x\rho(x,y)\,dA}{\iint_R \rho(x,y)\,dA} = \frac{M_y}{M},$$

$$y^* = \frac{\sum_{i=1}^{N}\sum_{j=1}^{M} y_{ij}\rho(x_{ij}, y_{ij})\,\Delta A}{\sum_{i=1}^{N}\sum_{j=1}^{M} \rho(x_{ij}, y_{ij})\,\Delta A} \overset{N,M\to\infty}{\to} \frac{\iint_R y\rho(x,y)\,dA}{\iint_R \rho(x,y)\,dA} = \frac{M_x}{M}.$$

Each numerator still has the meaning of a **first moment**, but now since we have two variables, the first moment needs to be understood against which variable.

$$M_y = \iint_R x\rho(x,y)\,dA$$

$$M_x = \iint_R y\rho(x,y)\,dA$$

**Remark 19.2** (Important Remark)**.** *This seems a little confusing because $M_y$ is used in the center of mass in the $x$-direction, i.e. $x^*$.* ***A good way to remember this is based on the physical understanding of what the integral actually means.*** *Take $M_y$ for example, it is simply measuring the "torque" of a mass if you were to rotate it about the $x$-**axis**, which depends on the distance from the point to the $x$-axis – this is precisely the $y$-coordinate.*

In other words, $M_y$ should be fully understood as the first moment of the lamina about the $x$-axis. Similarly, $M_x$ is the first moment of the lamina about the $y$-axis.

If the symbols $M_y$ and $M_x$ are too confusing to think about, still, then just remember that the coordinate of the center of mass is the coordinate which you integrate with the density (see the red highlighted $x^*$ has $x\rho(x,y)$ as the integrand).

**Example 19.3.** [Thin Plate in 2D] Consider the thin plate formed by the $y$-axis, $y = 8$ and the curve $y = x^3$, with density $\rho(x,y) = xy$. Find the center of mass of this plate.

**Solution.** The first goal is to setup the integral for total mass and the first moments.

We first compute the total mass.

$$M = \iint_R \rho(x,y)\,dA$$

$$
\begin{aligned}
&= \int_0^2 \int_{x^3}^8 xydydx \\
&= \int_0^2 x \boxed{\left[\frac{y^2}{2}\right]_{y=x^3}^{y=8}} dx \\
&= \int_0^2 x \boxed{\left[32 - \frac{x^6}{2}\right]} dx \\
&= \int_0^2 \left[32x - \frac{x^7}{2}\right] dx \\
&= \left[16x^2 - \frac{x^8}{16}\right]_0^2 \\
&= \left[64 - \frac{256}{16}\right] \\
&= 48.
\end{aligned}
$$

We can reuse the boxed item from the mass calculation. The first moment about $x$-axis is

$$
\begin{aligned}
M_y &= \iint_R x\rho(x,y)\, dA \\
&= \boxed{\int_0^2 \int_{x^3}^8 x(xy)\, dydx} \\
&= \int_0^2 x^2 \left[32 - \frac{x^6}{2}\right] dx \\
&= \int_0^2 \left[32x^2 - \frac{x^8}{2}\right] dx \\
&= \left[\frac{32}{3}x^3 - \frac{x^9}{18}\right]_0^2 \\
&= \frac{512}{9};
\end{aligned}
$$

the first moment about $y$-axis is

$$
\begin{aligned}
M_x &= \iint_R y\rho(x,y)\, dA \\
&= \boxed{\int_0^2 \int_{x^3}^8 y(xy)\, dydx}
\end{aligned}
$$

$$
\begin{aligned}
&= \int_0^2 x \left[\frac{y^3}{3}\right]_{y=x^3}^{y=8} dx \\
&= \int_0^2 x \left[\frac{512}{3} - \frac{x^9}{3}\right] dx \\
&= \left[\frac{256}{3}x^2 - \frac{x^{11}}{33}\right]_0^2 \\
&= \frac{1024}{3} - \frac{2048}{33} \\
&= \frac{3072}{11}.
\end{aligned}
$$

Therefore,

$$
(x^*, y^*) = \left(\frac{M_y}{M}, \frac{M_x}{M}\right) = \left(\frac{512}{9 \cdot 48}, \frac{3072}{11 \cdot 48}\right) = \left(\frac{32}{27}, \frac{64}{11}\right).
$$

Just to give you a sense of where it is, $\left(\frac{32}{27}, \frac{64}{11}\right) \approx (1.185, 5.818)$, which is still in the region of interest. The density function

$$
\rho(x, y) = xy
$$

suggests that the density increases linearly in the $x$ and $y$ directions respectively, so the shape is "heavier" towards north and east. It makes sense then that each component of the center of mass is northeast of the center point $(2, 4)$ of the region.

We can also pose a problem that uses polar coordinates.

**Example 19.4.** Consider the annulus in the 1st quadrant between a circle of radius 1 and 2, respectively. The density of the annulus is $\rho(x, y) = xy$. Find the total mass and first moments of the annulus.

**Solution.**

$$
\begin{aligned}
M_x &= \iint_R y\rho(x, y)\, dA \\
&= \int_0^{\frac{\pi}{2}} \int_1^2 r \sin\theta \left(r \cos\theta r \sin\theta\right) r dr d\theta \\
&= \left(\int_0^{\frac{\pi}{2}} \sin^2(\theta) \cos\theta d\theta\right) \left(\int_1^2 r^4 dr\right) \\
&= \left(\int_0^1 u^2 du\right) \left(\frac{r^5}{5} \Big|_{r=1}^{r=2}\right)
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{1}{3} \cdot \frac{1}{5} \left[2^5 - 1\right] \\
&= \frac{31}{15}.
\end{aligned}
$$

$$
\begin{aligned}
M_y &= \iint_R x\rho\,(x, y)\,dA \\
&= \int_0^{\frac{\pi}{2}} \int_1^2 r\cos\theta\,(r\cos\theta r\sin\theta)\,rdrd\theta \\
&= \left(\int_0^{\frac{\pi}{2}} \cos^2\theta\sin\theta d\theta\right)\left(\int_1^2 r^4 dr\right) \\
&= \left(-\int_1^0 u^2 du\right)\left(\frac{r^5}{5}\,|_{r=1}^{r=2}\right) \\
&= \frac{1}{3} \cdot \frac{1}{5} \left[2^5 - 1\right] \\
&= \frac{31}{15}.
\end{aligned}
$$

$$
\begin{aligned}
M &= \iint_R \rho\,(x, y)\,dA \\
&= \int_0^{\frac{\pi}{2}} \int_1^2 (r\cos\theta r\sin\theta)\,rdrd\theta \\
&= \left(\int_0^{\frac{\pi}{2}} \cos\theta\sin\theta d\theta\right)\left(\int_1^2 r^3 dr\right) \\
&= \left(\int_0^1 udu\right)\left(\frac{r^4}{4}\,|_{r=1}^{r=2}\right) \\
&= \frac{1}{2} \cdot \frac{1}{4} \left[2^4 - 1\right] \\
&= \frac{15}{8}.
\end{aligned}
$$

Thus,

$$
\begin{aligned}
x^* &= \frac{M_y}{M} = \frac{\frac{31}{15}}{\frac{15}{8}} = \frac{248}{225} \\
y^* &= \frac{M_x}{M} = \frac{\frac{31}{15}}{\frac{15}{8}} = \frac{248}{225}
\end{aligned}
$$

# 20 Lecture XX: Triple Integrals in Cartesian Coordinates

## 20.1 Motivation

Recall that when we deal with double integrals, the region of integration lives in the $x$-$y$ plane – a 2D shape. A classical example is to find the area of the region $R$ between two curves, $f_2(x)$ (as the upper) and $f_1(x)$ (as the lower). This area can be expressed using a double integral

$$A = \iint_R dA = \int_a^b \int_{f_2(x)}^{f_1(x)} dydx.$$

Now, what about the volume between two surfaces? The simplest example would be between two spherical caps, that is, the volume of a sphere. Does "volume between surfaces" follow a similar formulation to "area between curves"?

## 20.2 Riemann Sum: Triple Integral

Motivated by the double integral formulation of the area of a region between two curves, consider now the volume between two surfaces.

Analogous to how we sum up areas in two dimensions, the fundamental approach here is to think of infinitesimal volumes,

$$\Delta V_{ijk} = \Delta x_i \Delta y_j \Delta z_k$$

(recall an area element is $\Delta A_{ij} = \Delta x_i \Delta y_j$).

To estimate the volume of an object in 3D we consider the Riemann sum

$$S_n = \sum_{i=1}^{n} \sum_{k=1}^{n} \sum_{j=1}^{n} \Delta V_{ijk}.$$

As we make $\Delta V_k$ smaller and smaller (while $n$, the number of volume elements, increases), we retrieve

$$\lim_{n\to\infty} S_n = \iiint_D dV = \iiint_D dxdydz$$

should this limit exist.

The volume of a closed, bounded region in space then follows very analogous to the area formula,

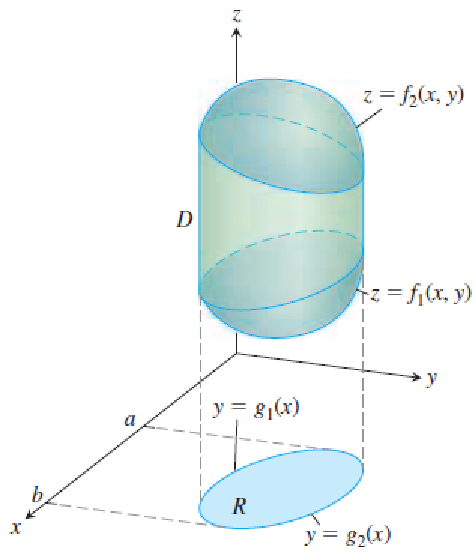
$$V = \iiint_D dV.$$

Next, we illustrate the needle method for Cartesian coordinates in 3D, **to determine the bounds of integration**.

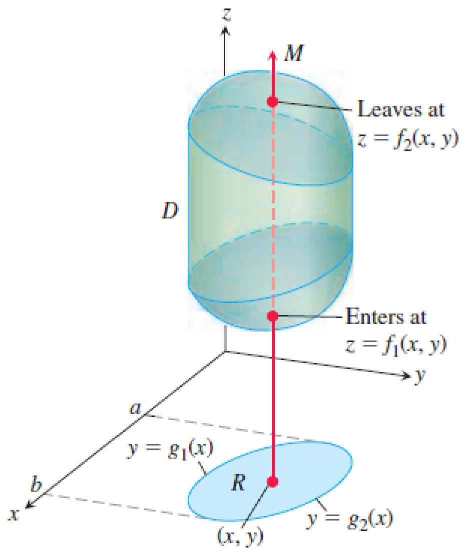
## 20.3 The Needle Method for 3D objects

When we perform a triple integral, our region now becomes a 3D shape, which now becomes increasingly hard to visualize. Our next job is to consider the bounds of integration using the information of $D$. Let us focus on the case where we integrate against $z$, $y$ and $x$, in this particular order. Other ordering leads to similar interpretations.

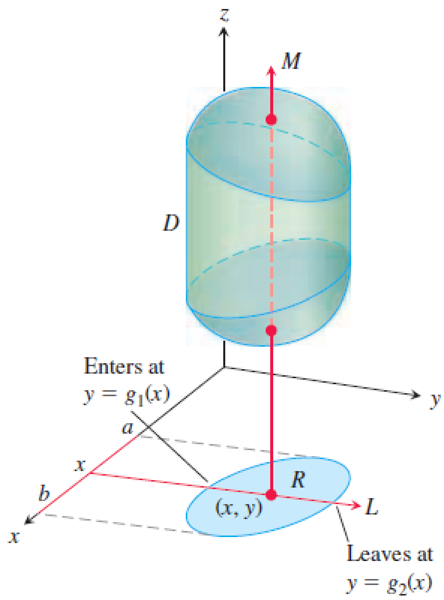
1. Sketch $D$ along with its shadow $R$ (vertical projection) in the $x$-$y$ plane.




2. Find the $z$-limits of integration by drawing a line $M$ passing through a typical point $(x, y)$ in $R$ parallel to the $z$-axis. This line will encounter two surfaces that outline $D$, namely, $z = f_1(x, y)$ upon entering, and $z = f_2(x, y)$ upon leaving.




3. Find the $y$-limits of integration then by drawing a line $L$ through $(x, y)$ parallel to $y$-axis. This line encounters $R$ at two places, $y = g_1(x)$ upon entering and $y = g_2(x)$ upon leaving.




**Remark 20.1.** *This is a quite important step. One way to remember this is to understand that the direction in z is dealt with in the previous step. Therefore, we no longer need to think about "elevation" but just a problem in the xy-plane.*

*In other words, you have solved the $z$-problem, thus detaching from it. Since the problem now is in the $xy$-plane, use the needle for 2D Cartesian, as we have learned before.*

4. Find the $x$-limits of integration by finding the $x$-coordinates of all the lines $L$ (like the one in part 3) that can encounter $R$. This yields $x = a$ and $x = b$.

Altogether, the triple integral for a general region is

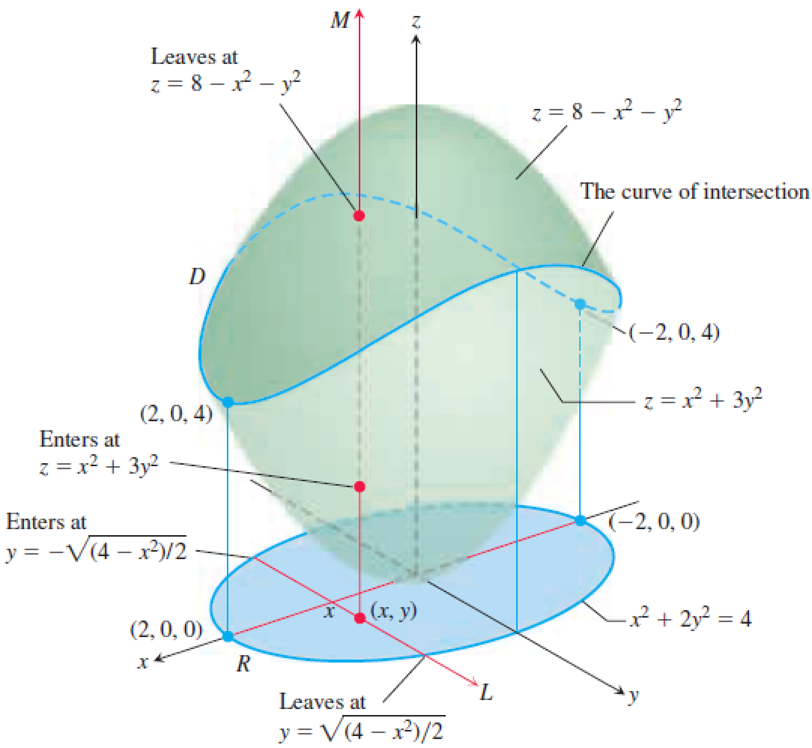
$$\int_{x=a}^{x=b} \int_{y=g_1(x)}^{y=g_2(x)} \int_{z=f_2(x,y)}^{z=f_1(x,y)} F\left(x, y, z\right) dzdydx.$$

**Example 20.1.** Find the volume of the spherical cap bounded by the sphere $x^2 + y^2 + z^2 = 25$ and the horizontal plane $z = 3$.

$$\begin{aligned}
\int_{-4}^{4} \int_{-\sqrt{16-x^2}}^{\sqrt{16-x^2}} \int_{3}^{\sqrt{25-x^2-y^2}} dzdydx &= \int_{-4}^{4} \int_{-\sqrt{16-x^2}}^{\sqrt{16-x^2}} \left(\sqrt{25 - x^2 - y^2} - 3\right) dydx \\
&= \int_{0}^{2\pi} \int_{0}^{4} \left(\sqrt{25 - r^2} - 3\right) rdrd\theta \\
&= 2\pi \left[\int_{0}^{4} r\sqrt{25 - r^2}dr - \int_{0}^{4} 3rdr\right] \\
&= 2\pi \left[\frac{1}{2} \int_{3}^{5} \sqrt{u}du - \frac{3}{2} \left[r^2\right]_{0}^{4}\right] \\
&= \pi \left[\frac{2}{3}u^{\frac{3}{2}}\right]_{3}^{5} - 3\pi \left[4^2 - 0^2\right] \\
&= \frac{2}{3}\pi \left(5^{3/2} - 3^{3/2}\right) - 48\pi
\end{aligned}$$

**Remark 20.2.** *Notice that at the second equality, we used polar coordinates to simplify the integrand, AND the bounds of integration. It is simply the case of one stone for two birds. This saves a lot of time by being intelligent.*

**Example 20.2.** Set up the integral that determines the volume of the region $D$ enclosed by the surfaces $z = x^2 + 3y^2$ and $z = 8 - x^2 - y^2$.




**Solution.** We need to identify the region of integration.

First, we need to determine which surface is the "top" and which is "bottom". Both surfaces are paraboloids. $z = x^2 + 3y^2$ is opening up, and attains a minimum at the origin $(0, 0, 0)$. $z = 8 - x^2 - y^2$ is opening down, and attains a maximum at $(0, 0, 8)$. This implies that $z = 8 - x^2 - y^2$ is the top surface. This effectively determines the range of values for $z$, namely,

$$x^2 + 3y^2 \le z \le 8 - x^2 - y^2.$$

Second, we need to find where the two surfaces intersect. Setting them equal, we have

$$x^2 + 3y^2 = 8 - x^2 - y^2 \implies 2x^2 + 4y^2 = 8 \implies y = \pm\sqrt{\frac{4 - x^2}{2}}.$$

This is a closed curve in the $xy$-plane. To describe the region enclosed by the curve, we revert back to the old technique – for the variable bounds of $y$, we send a line parallel to the $y$-axis and see where this line enters and exits the region.

$$-\sqrt{\frac{4 - x^2}{2}} \le y \le \sqrt{\frac{4 - x^2}{2}}.$$

Lastly, we find the constant bounds for $x$. Note that the closed curve has domain $4 - x^2 \geq 0$ which automatically implies that

$$-2 \leq x \leq 2.$$

Altogether, the volume is represented by the triple integral over $dV = dzdydx$,

$$\begin{aligned} V &= \iiint_D dzdydx \\ &= \int_{-2}^{2} \int_{-\sqrt{(4-x^2)/2}}^{\sqrt{(4-x^2)/2}} \int_{x^2+3y^2}^{8-x^2-y^2} dzdydx. \end{aligned}$$

**Example 20.3** (Spherical cap)**.** Find the volume of the spherical cap bounded by the sphere $x^2 + y^2 + z^2 = 25$ and the horizontal plane $z = 3$.

**Solution.** The part of the (northern) hemisphere can be rewritten as a function in two variables

$$z = f(x, y) = \sqrt{25 - x^2 - y^2}.$$

Therefore, taking a $z$-needle parallel to the $z$-axis, we see that the bounds for the variable $z$ is

$$3 \leq z \leq \sqrt{25 - x^2 - y^2}.$$

Now, consider the intersection of the horizontal plane and the sphere, described by the curve satisfying

$$3 = \sqrt{25 - x^2 - y^2} \implies x^2 + y^2 = 4^2$$

that is, a circle of radius 4.

Even though this region is easier to be described in polar, we do the Cartesian coordinate as a practice:

$$-\sqrt{16 - x^2} \leq y \leq \sqrt{16 - x^2}, \quad -4 \leq x \leq 4.$$

Piecing all the bounds together, we have

$$V = \int_{-4}^{4} \int_{-\sqrt{16-x^2}}^{\sqrt{16-x^2}} \int_{3}^{\sqrt{25-x^2-y^2}} dzdydx.$$

## 20.4 Application of Triple Integrals: Mass and Moments of Objects in 3D

Similar to 2D objects, we now formulate the total mass and the first moment of solid objects in 3D, with a variable density $\rho(x, y, z)$. The added dimension suggests that the orientation of a rotation should be specified towards a particular coordinate **plane** rather than axis. Indeed, we can always measure the distance from a little cube at $(x, y, z)$ inside an object in 3D to one of the coordinate planes. For example, the distance from the point $(x, y, z)$ to the $yz$-plane is $x$, etc.

Drawing from inspiration from 2D objects, we have

$$\begin{aligned}
\text{(Total Mass) } M &= \iiint_D \rho(x, y, z)\, dV \\
\text{(First Moment about } yz\text{-plane) } M_{yz} &= \iiint_D x\rho(x, y, z)\, dV \\
\text{(First Moment about } xz\text{-plane) } M_{xz} &= \iiint_D y\rho(x, y, z)\, dV \\
\text{(First Moment about } xy\text{-plane) } M_{xy} &= \iiint_D z\rho(x, y, z)\, dV
\end{aligned}$$

Note that $M_{yz}$ is found with a seesaw parallel to the $x$-axis, and thus perpendicular to the $yz$-plane. We call $M_{yz}$ the **first moment about the $yz$-plane**. The important bit here is that this integral is integrated against $x$, which is treated as the distance to the reference plane ($yz$-plane in this case).

The center of mass is then given by the ratios of the first moments against total mass. Remember, the center of mass is no other than the **weighted** average of all masses due to spatially dependent density and shape of the object.

$$\overline{x} = \frac{M_{yz}}{M}, \quad \overline{y} = \frac{M_{xz}}{M}, \quad \overline{z} = \frac{M_{xy}}{M},$$

where

$$M = \iiint_D \rho(x, y, z)\, dV.$$

In 2D,

$$\overline{x} = \frac{M_y}{M}, \quad \overline{y} = \frac{M_x}{M},$$

where

$$M = \iint_D \rho(x, y)\, dA.$$

**Example 20.4.** Find the center of mass of a solid of constant density $\rho$ bounded below by the disk $R : x^2 + y^2 \le 4$ in the plane $z = 0$ and above by the paraboloid $z = 4 - x^2 - y^2$.

**Solution.** By symmetry $\bar{x} = \bar{y} = 0$. To find $\bar{z}$, we need the first moment about the $xy$-plane, i.e.,

$$\begin{aligned}
M_{xy} &= \iint_R \int_{z=0}^{z=4-x^2-y^2} z\rho dzdydx \\
&= \frac{\rho}{2} \iint_R \left(4 - x^2 - y^2\right)^2 dydx \\
&= \frac{\rho}{2} \int_0^{2\pi} \int_0^2 \left(4 - r^2\right)^2 rdrd\theta \\
&= \rho\pi \left[-\frac{1}{6} \left(4 - r^2\right)^3 |_0^2\right] \\
&= \rho\pi \left(0 + \frac{64}{6}\right) \\
&= \frac{32}{3}\pi\rho.
\end{aligned}$$

The total mass

$$\begin{aligned}
M &= \iint_R \int_{z=0}^{z=4-x^2-y^2} \rho dzdydx \\
&= \rho \iint_R \left(4 - x^2 - y^2\right) dzdydx \\
&= \rho \int_0^{2\pi} \int_0^2 \left(4 - r^2\right) rdrd\theta \\
&= 2\pi\rho \left[2r^2 - \frac{r^4}{4}\right]_0^2 \\
&= 2\pi\rho \left[8 - 4\right] \\
&= 8\pi\rho.
\end{aligned}$$

Thus,

$$\bar{z} = \frac{M_{xy}}{M} = \frac{4}{3}.$$

**Remark 20.3.** *We see that the center of mass for objects with constant density does NOT depend on (the value of) the density.*

The center of mass of an object with constant density is called the **centroid** of the object. One may set the constant density to 1 and proceed as the calculation in the last example.

# 21 Lecture XXI: Cylindrical Coordinates

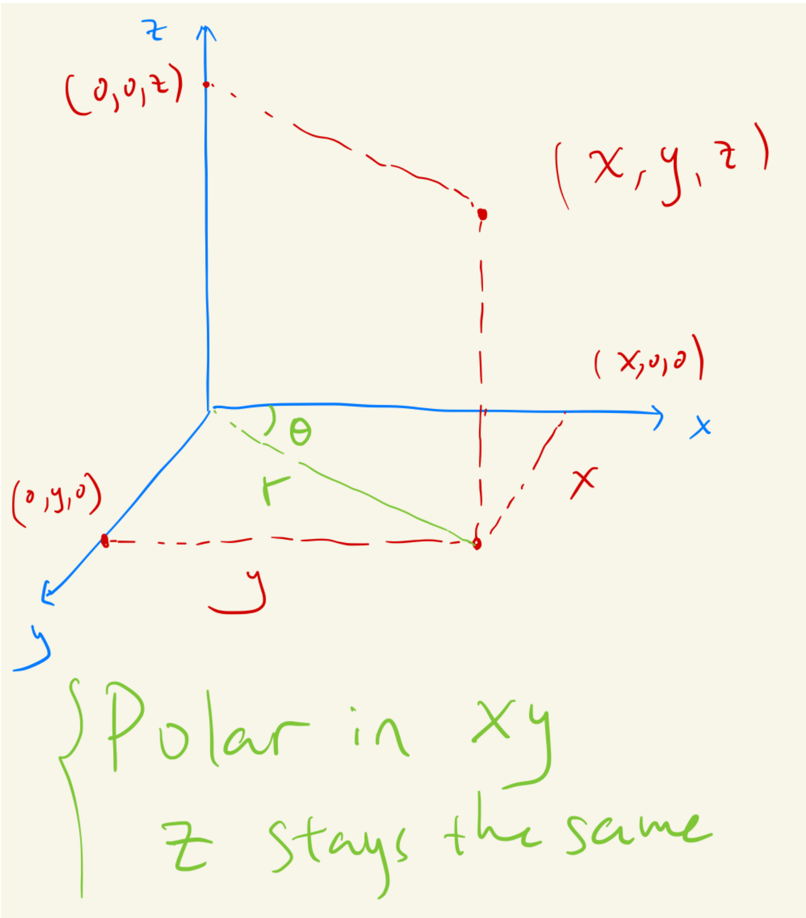
When the underlying geometry over which we integrate is a cylinder, it is beneficial to utilize the nice symmetry of the shape by introducing a new transformation.

**Definition 21.1.** ***Cylindrical coordinates*** *represent a point* $P(x, y, z)$ *in the Euclidean space by an ordered triple* $(r, \theta, z)$ *in which* $r \geq 0$,

1. $r$ *and* $\theta$ *are* ***polar coordinates*** *for the vertical projection of* $P$ *on the* $xy$*-plane.*
2. $z$ *is the rectangular vertical coordinate (i.e., nothing is changed in this direction).*

The transformation is detailed by the following forward and backward relationships

$$x = r\cos\theta, \quad y = r\sin\theta, \quad z = z \iff r^2 = x^2 + y^2, \quad \tan(\theta) = \frac{y}{x}.$$



**Example 21.1** (Common Cartesian expressions in Cylindrical)**.** First, let's convert various shapes/surfaces from Cartesian Coordinates into Cylindrical Coordinates.

1. Sphere of radius 1 centered at the origin.

$$x^2 + y^2 + z^2 = 1 \iff r^2 \cos^2\theta + r^2 \sin^2\theta + z^2 = 1$$
$$\implies \boxed{r^2 + z^2 = 1}.$$

2. Paraboloids

$$z = 100 - x^2 - y^2 \iff \boxed{z = 100 - r^2}.$$

3. Sphere of radius 1 centered at $(1, 0, 0)$

$$(x-1)^2 + y^2 + z^2 = 1 \implies x^2 - 2x + 1 + y^2 + z^2 = 1$$
$$\implies x^2 + y^2 - 2x + z^2 = 0 \implies \boxed{r^2 - 2r\cos\theta + z^2 = 0}.$$

## 21.1 The Needle Method in Cylindrical Coordinates

Now, let's study an example with a triple integral in which we must learn how to convert the region of integration into cylindrical coordinates. The main machinery is still, the needle method, now in cylindrical coordinates.

---

**Procedures of The Needle Method in Cylindrical Coordinates**

1. $z$-needle: exactly the same as in Cartesian coordinates.

2. Then, projected in the $xy$-plane, do polar coordinate **exactly**

   (a) $r$-needle in the $xy$-plane.

   (b) $\theta$ bounds are found by considering sweeping the $r$-needle, across the angles where the needle intersects with the projection.

---

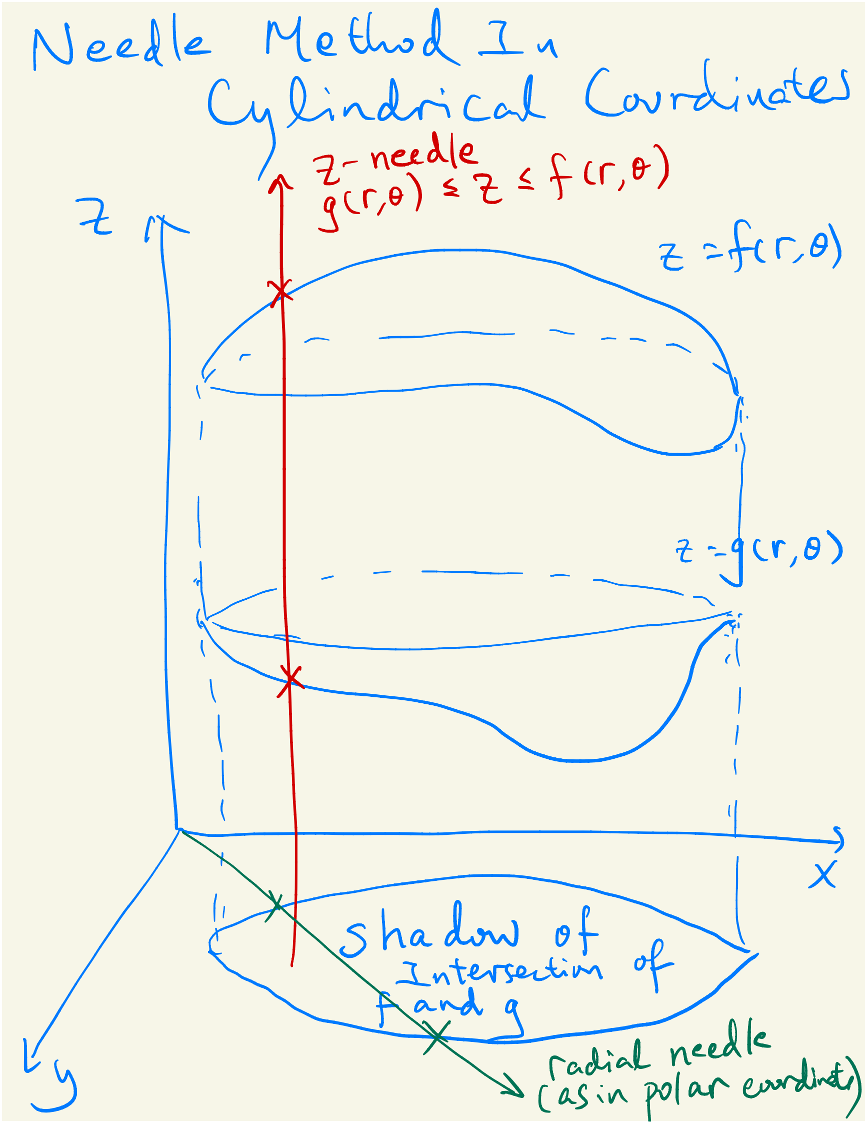



Figure 12: The Needle Method in Cylindrical Coordinates.

**Example 21.2** (Spherical cap revisited)**.** Find the volume of the spherical cap bounded by the sphere $x^2 + y^2 + z^2 = 25$ and the horizontal plane $z = 3$.

We note first that

$$[\text{plane at } z = 3] \leq z \leq [\text{northern hemisphere of radius 5}]$$

by using the needle technique. Now, to make both sides of the equation precise, first in Cartesian, and then convert to cylindrical. We note that the LHS is as straightforward as $z = 3$. For the RHS, we write in Cartesian first, and then cylindrical

$$z = \sqrt{25 - x^2 - y^2} = \sqrt{25 - r^2}$$

Thus,

$$3 \leq z \leq \sqrt{25 - r^2}.$$

Now, as we have taken care of $z$ which represents the height of the cap, we are now in the $x$-$y$ plane. The cutting cross-section is in fact a circle of radius 4 – it is found by solving $z = 3$ and $x^2 + y^2 + z^2 = 25$ simultaneously.

$$x^2 + y^2 + 3^2 = 25 \implies x^2 + y^2 = 16$$

which certainly implies that the cross-section is a circle of radius 4. Since cylindrical coordinate has exactly the same form as the polar coordinate (except the added but already resolved $z$ dimension), we have

$$0 \leq \theta \leq 2\pi, \quad 0 \leq r \leq 4$$

to describe this circle of radius 4 (centered at the $z$-axis). Thus

$$\begin{aligned}
\int_0^{2\pi} \int_0^4 \int_3^{\sqrt{25-r^2}} dz r dr d\theta &= \int_0^{2\pi} \int_0^4 \left( \sqrt{25 - r^2} - 3 \right) r dr d\theta \\
&= 2\pi \left[ \int_0^4 r\sqrt{25 - r^2} dr - \int_0^4 3r dr \right] \\
&= 2\pi \left[ \frac{1}{2} \int_3^5 \sqrt{u} du - \frac{3}{2} \left[ r^2 \right]_0^4 \right] \\
&= \pi \left[ \frac{2}{3} u^{\frac{3}{2}} \right]_3^5 - 3\pi \left[ 4^2 - 0^2 \right] \\
&= \frac{2}{3}\pi \left( 5^{3/2} - 3^{3/2} \right) - 48\pi
\end{aligned}$$

## 21.2 Cylindrical Volume Element: Pizza Crust

Now, to utilize this transformation in a triple integral, we need to identify the volume element represented by $(r, \theta, z)$. Consider a piece cut from a circular cake with thickness $\Delta z$. We know from polar coordinates that the area element now is

$$dA = r dr d\theta,$$

which is represented by the area of the base of the incremental "cake". Thus, the volume of this increment is

$$dV = r dr d\theta dz.$$

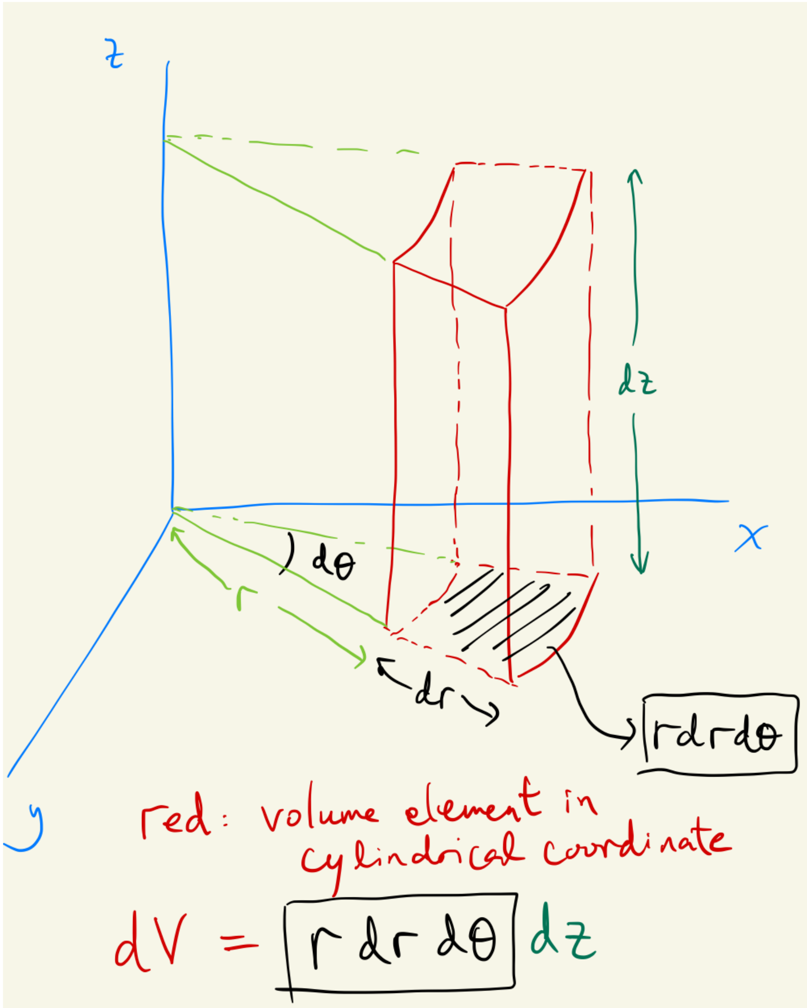



Figure 13

## 21.3 Examples

**Example 21.3** (Scooping parabolically a Cylindrical Ice Cream)**.** Find the limits of integration in cylindrical coordinates for integrating a function $f(r, \theta, z)$ over the region $D$ bounded below by the plane $z = 0$, laterally by the circle cylinder $x^2 + (y-1)^2 = 1$ and above by the paraboloid $z = x^2 + y^2$.

Note first that since the $z$-coordinate is unchanged in cylindrical coordinates, the method to find the bounds of integration follows exactly the same way as in a usual triple integral. One sends a line parallel to the $z$-axis through the shape and observe where the line enters and leaves the region. In this case, the line enters at the flat plane $z = 0$ and leaves at the paraboloid $z = x^2 + y^2$. Thus

$$\iint \int_0^{x^2+y^2} f(r, \theta, z)\, dz\, r dr d\theta.$$

Now, since eventually we need to integrate in $r$ and $\theta$, it would be worthwhile to write the upper bound here in polar, namely, $r^2$. We now have

$$\iint \int_0^{r^2} f(r, \theta, z)\, dz\, r dr d\theta.$$

To find the bounds for $r$, we look at the $xy$-plane, where we observe a circle centered at $(0, 1)$. Starting at the origin, we send a straight line through the shape. We note that it starts with $r = 0$, but exits the shape at the circle, which necessitates an expression of the circle in polar coordinates. In other words, find the polar expression for the circle

$$x^2 + (y - 1)^2 = 1.$$

We use polar transformation $x = r\cos\theta$ and $y = r\sin\theta$ and $x^2 + y^2 = r^2$, expand the above expression as

$$x^2 + y^2 - 2y + 1 = 1 \implies x^2 + y^2 = 2y \implies r^2 = 2r\sin\theta \implies r = 2\sin\theta.$$

This means $r = 2\sin\theta$ describes the circle centered at $(x, y) = (0, 1)$ with radius 1. Thus, using the radial needle in the $xy$-plane, we find

$$0 \le r \le 2\sin\theta.$$

Lastly, for $\theta$, we observe that this circle sits above the $x$-axis. Starting at the origin, we send a line (in the radial direction, as we call it) through the shape – in order to cover the line, this line must span an angle from 0 to $\pi$ to cover the circle. Thus,

$$0 \le \theta \le \pi.$$

Altogether, we have

$$\iiint_D f(r, \theta, z)\, dV = \int_0^{\pi} \int_0^{2\sin\theta} \int_0^{r^2} f(r, \theta, z)\, dz\, r dr d\theta$$

# 22 Lecture XXII: Spherical Coordinates

We have seen two ways of describing a point in 3D space – rectangular coordinate $(x, y, z)$ and cylindrical coordinates $(r, \theta, z)$. In practice, we sometimes deal with objects that have more symmetry than cylinders (only symmetric about the axis perpendicular to the cylinder base). Spherical (or similar) objects enjoy symmetry along three principal axes ($x$, $y$ and $z$).

Consider a point $(x, y, z)$ that lives in 3D space. We associate its distance to the origin as "radius", namely,

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

Draw a line $L$ that links $(x, y, z)$ and the origin. We label the angle between $L$ and the $z$-axis as $\theta$, called **polar angle**. This angle satisfies

$$z = \rho \cos\theta$$

by looking at the triangle formed by $L$, the $z$-axis and the shortest line connecting the point and the $z$-axis, call it $L_z$.

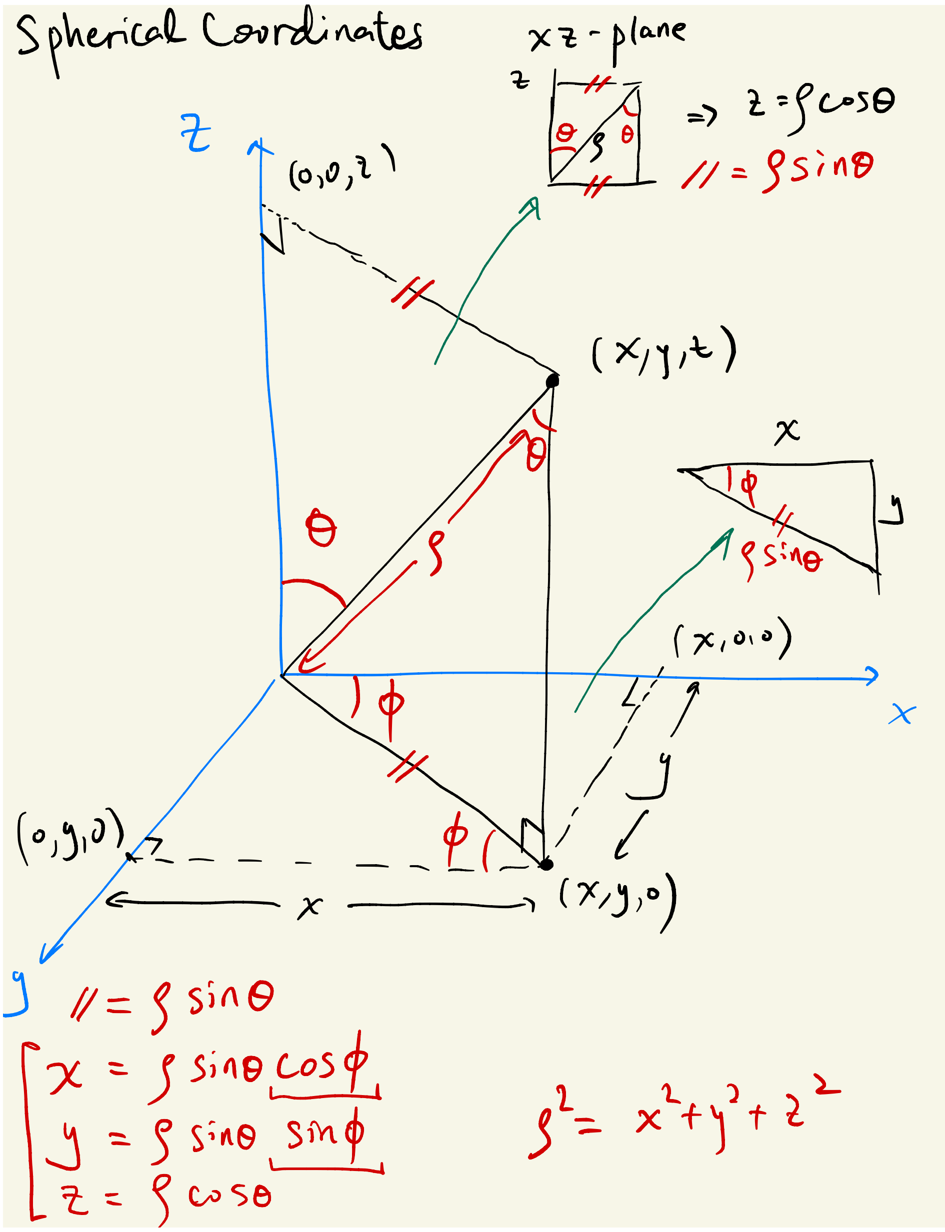
**Remark 22.1** (Lexical semantics). *The polar angle is called **polar** because changing this angle takes us to the north **pole**.*

If we were to stop now, we have obtained a collection of points that live on the circle formed by rotating the line $L$ at fixed polar angle $\theta$. To really pinpoint a point on this circle, we need to specify the angle the projection of this point makes with the $x$-axis, just like in polar coordinates. When we project this line onto the $xy$-plane, and obtain now a new line $L_{xy}$, whose length is equal to $\rho \sin\theta$. $L_{xy}$ is in fact the parallel line to $L_z$ (the red double bars in the sketch below).

Knowing that the length of $L_{xy}$ is $\rho \sin\theta$, we can then relate the $x$ and $y$ coordinates, since now we are only working in the $xy$-plane. Label the angle between $L_{xy}$ and the $x$-axis by $\phi$, called the **azimuthal angle**. Denote $|L_{xy}|$ as the length of $L_{xy}$. We observe that

$$x = |L_{xy}| \cos\phi = \rho \sin\theta \cos\phi;$$
$$y = |L_{xy}| \sin\phi = \rho \sin\theta \sin\phi.$$

Altogether, given $(x, y, z)$, we express it in terms of $(\rho, \theta, \phi)$.




we can translate it to spherical coordinates using

$$x = \rho \sin\theta \cos\phi; \tag{22.1a}$$

$$y = \rho \sin\theta \sin\phi; \tag{22.1b}$$

$$z = \rho \cos\theta. \tag{22.1c}$$

Let's also identify the range of values that $\theta$ and $\phi$ can take. $\theta$ is the angle between $L$ and the $z$-axis – that is, if $L$ IS parallel to the $z$-axis, then $\theta = 0$, whereas if $L$ is parallel to the negative $z$-axis, then $\theta = \pi$ (turned only 180 degrees). Thus,

$$0 \leq \theta \leq \pi.$$

$\phi$ is the same as in polar coordinates since it sweeps the entire $xy$-plane, namely, $0 \leq \phi \leq 2\pi$.

**Remark 22.2.** *The Spherical Coordinate system is the true 3D upgrade of polar coordinates in 2D. The Cylindrical Coordinate system is a superficial 3D coordinate system with $z$ as an added dimension, like a car (polar) with wings, but still a car (intrinsically 2D polar). The Spherical Coordinate System is like a spaceship that serves an equivalent function as a car in 2D, so it is truly "3D polar".*

**Takeaway 22.1.** Every measurement in Spherical Coordinates is relative to the origin, just like in polar in 2D.

## 22.1 Expressions of Objects in Spherical Coordinates

**Example 22.1.** Spherical coordinates, by the name, should really simplify the expression of a sphere. Let's use the spherical coordinate transformation given in Equation (22.1) and write out all the details

$$\begin{aligned}
x^2 + y^2 + z^2 &= 1 \\
\implies (\rho \sin\theta\cos\phi)^2 + (\rho\sin\theta\sin\phi)^2 + (\rho\cos\theta)^2 &= 1 \\
\implies \rho^2\sin^2\theta\cos^2\phi + \rho^2\sin^2\theta\sin^2\phi + \rho^2\cos^2\theta &= 1 \\
\implies \rho^2\sin^2\theta\big(\cancelto{1}{\cos^2\phi + \sin^2\phi}\big) + \rho^2\cos^2\theta &= 1 \\
\implies \rho^2\sin^2\theta + \rho^2\cos^2\theta &= 1 \\
\implies \rho^2\left(\sin^2\theta + \cos^2\theta\right) &= 1 \\
\implies \rho^2 &= 1 \\
\implies \rho &= 1
\end{aligned}$$

In other words, a sphere in spherical coordinates is a constant, just like $r = 1$ represents a circle of radius 1 centered at the origin in polar. It is **really** simplified.

**Example 22.2.** Express the cone $z = \sqrt{x^2 + y^2}$ in spherical coordinates.

$$z = \sqrt{x^2 + y^2}$$

$$\implies \rho\cos\theta = \sqrt{\rho^2\sin^2\theta} = \rho\sin\theta$$
$$\implies \theta = \frac{\pi}{4}$$

**Example 22.3.** Express the sphere $x^2 + y^2 + (z-1)^2 = 1$ in spherical coordinates.

$$x^2 + y^2 + (z-1)^2 = 1$$
$$\implies x^2 + y^2 + z^2 - 2z + 1 = 1$$
$$\implies \rho^2 - 2\rho\cos\theta = 0$$
$$\implies \rho = 2\cos\theta.$$

**Example 22.4.** The horizontal plane $z = 3$.

$$\rho\cos\theta = 3.$$

No surprise here. However, in triple integrals, we tend to separate the variables, and thus,

$$\rho = 3\sec\theta$$

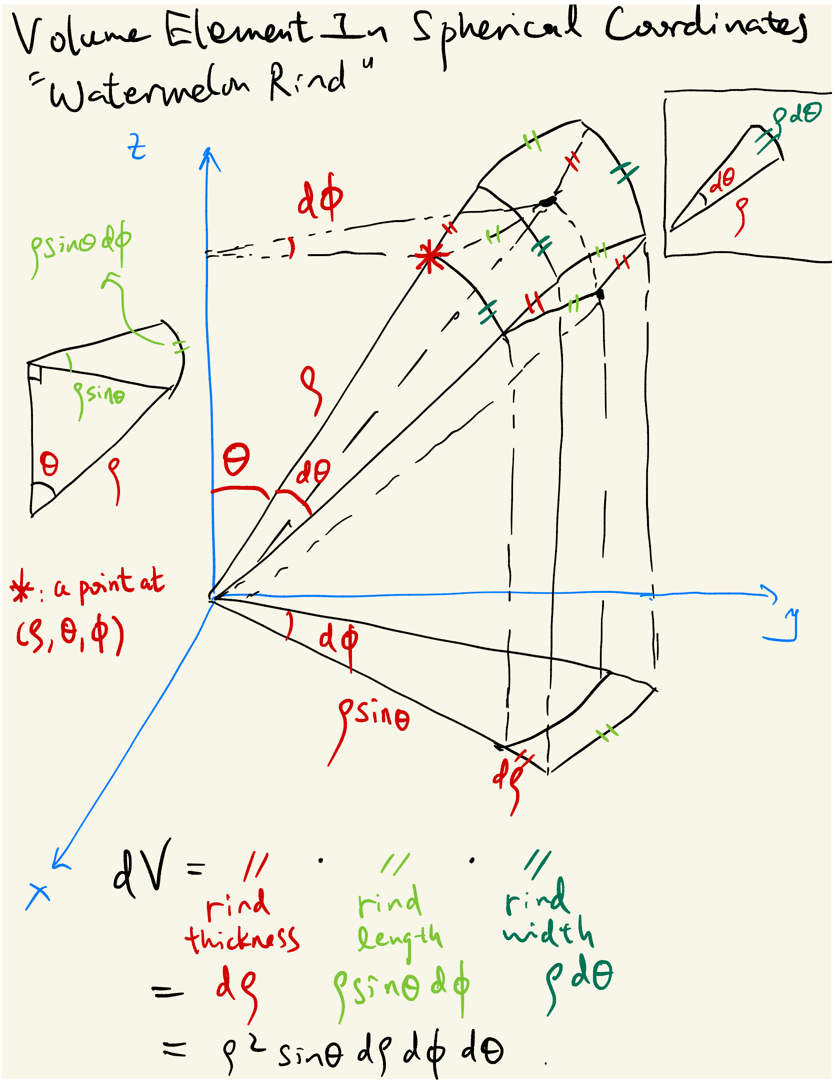
would be a more apt spherical representation of the plane $z = 3$.

**Takeaway 22.2.** $x^2 + y^2 + z^2 = \rho^2$ is your good friend. When converting the expression from Cartesian to Spherical, try to "cook up" $x^2 + y^2 + z^2$ from algebra to start the simplification.

# 23 Lecture XXIII: Triple Integrals in Spherical Coordinates

## 23.1 Volume Element in Spherical Coordinates

We want to estimate the volume element $dV$ in spherical coordinates. The idea is to cut out a wedge from a spherical watermelon, and estimate the volume of its rind.




$$dV = \underbrace{d\rho}_{\text{rind thickness}} \underbrace{\rho\sin\theta d\phi}_{\text{rind length}} \underbrace{\rho d\theta}_{\text{rind width}} = \rho^2 \sin\theta\, d\rho\, d\phi\, d\theta$$

# 24 Lecture XIV: Examples of Triple Integrals in Spherical Coordinates

**Example 24.1** (Spherical cap re-revisited)**.** Set up a triple integral in spherical coordinates that compute the volume of a spherical cap, obtained by cutting the sphere of radius 2 centered at the origin with a horizontal plane $z = 1$.

**Solution.** We have to determine the bounds of integration for three variables, so three steps.

1. Bounds for $\rho$.

   Send a radial needle in 3D – start at the origin, and shoot towards to the region of interest. The needle enters the region at the plane $z = 1$ and leaves the region at the sphere. The sphere here corresponds to $\rho = 2$. So, we just need to rewrite $z = 1$ in spherical coordinates, which we have at our disposal per example in the last section.

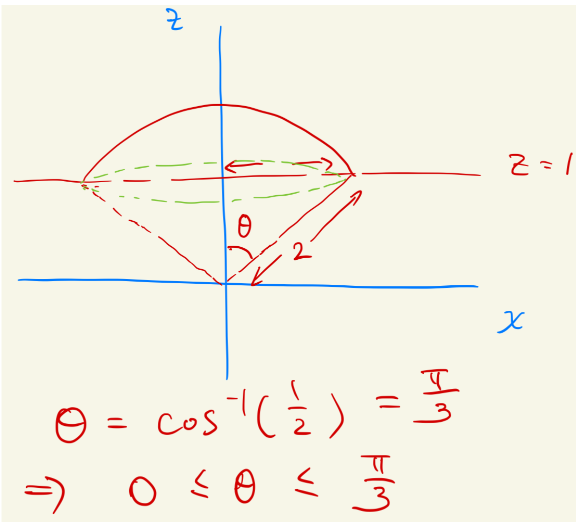
   $$z = 1 \iff \rho = \sec\theta.$$

   Therefore,

   $$\sec\theta \leq \rho \leq 2.$$

2. Bounds for $\theta$, the polar angle.

   The polar angle has nothing to do with what happens in the $xy$-plane, so we look at the $xz$ or $yz$-cross-section of this spherical cap.



Noting that $\theta$ starts at 0 along the positive $z$-axis, we find

$$0 \leq \theta \leq \frac{\pi}{3}.$$

3. Bounds for $\phi$, the azimuthal angle.

   The cap revolves around the $z$-axis fully, so

$$0 \leq \phi \leq 2\pi.$$

Altogther,

$$V = \int_0^{2\pi} \int_0^{\frac{\pi}{3}} \int_{\sec(\theta)}^{2} \rho^2 \sin\theta \, d\rho \, d\theta \, d\phi.$$

To evaluate, just do it patiently (and independently first), making shortcuts when you know what you are doing

$$\begin{aligned}
V &= \int_0^{2\pi} \int_0^{\frac{\pi}{3}} \int_{\sec(\theta)}^{2} \rho^2 \sin\theta d\rho d\theta d\phi \\
&= \left( \int_0^{2\pi} d\phi \right) \left( \int_0^{\frac{\pi}{3}} \sin\theta \int_{\sec(\theta)}^{2} \rho^2 d\rho d\theta \right) \\
&= 2\pi \int_0^{\frac{\pi}{3}} \sin\theta \left[ \frac{\rho^3}{3} \right]_{\rho=\sec\theta}^{\rho=2} d\theta \\
&= 2\pi \int_0^{\frac{\pi}{3}} \sin\theta \left[ \frac{8}{3} - \frac{\sec^3\theta}{3} \right] d\theta \\
&= 2\pi \left[ \frac{8}{3} \int_0^{\frac{\pi}{3}} \sin\theta d\theta - \frac{1}{3} \int_0^{\frac{\pi}{3}} \frac{\sin\theta}{\cos^3\theta} d\theta \right] \\
&= 2\pi \left[ \frac{8}{3} \left( \cos(0) - \cos\left( \frac{\pi}{3} \right) \right) - \frac{1}{3} \int_{\frac{1}{2}}^{1} \frac{1}{u^3} du \right] \\
&= 2\pi \left[ \frac{8}{3} \cdot \frac{1}{2} + \frac{1}{3} \left[ \frac{1}{2u^2} \right]_{u=\frac{1}{2}}^{u=1} \right] \\
&= 2\pi \left[ \frac{4}{3} + \frac{1}{3} \left[ \frac{1}{2} - 2 \right] \right] \\
&= \frac{5\pi}{3}
\end{aligned}$$

**Example 24.2.** Find the volume of a sphere of radius 1.

$$\begin{aligned}
\iiint_D dV &= \int_0^{2\pi}\int_0^{\pi}\int_0^1 \rho^2 \sin\theta d\rho d\theta d\phi \\
&= \left(\int_0^{2\pi} d\phi\right)\left(\int_0^{\pi} \sin\theta d\theta\right)\left(\int_0^1 \rho^2 d\rho\right) \\
&= 2\pi\left(-\cos\theta\right)_0^{\pi}\left(\frac{\rho^3}{3}\right)_0^1 \\
&= -2\pi\left(-1-(1)\right)\frac{1}{3} \\
&= \frac{4\pi}{3}
\end{aligned}$$

as expected.

**Example 24.3.** Find the volume of the "ice cream cone" $D$ cut from the solid sphere $\rho \leq 1$ by the cone $\theta = \frac{\pi}{3}$.

$$\begin{aligned}
\iiint_D dV &= \int_0^{2\pi}\int_0^{\pi/3}\int_0^1 \rho^2 \sin\theta d\rho d\theta d\phi \\
&= \left(\int_0^{2\pi} d\phi\right)\left(\int_0^{\pi/3} \sin\theta d\theta\right)\left(\int_0^1 \rho^2 d\rho\right) \\
&= 2\pi\left(-\cos\theta\right)_0^{\pi/3}\left(\frac{\rho^3}{3}\right)_0^1 \\
&= 2\pi\left(-\frac{1}{2}-(-1)\right)\frac{1}{3} \\
&= \frac{\pi}{3}
\end{aligned}$$

**Example 24.4.** Consider a paraboloid $z = x^2+y^2$ and a sphere of radius $\sqrt{2}$ centered at the origin. Set up a triple integral in spherical coordinates that calculate the volume of the region bounded above by the sphere and below by the paraboloid.

**Solution.** In spherical coordinates, the lower paraboloid reads

$$z = x^2 + y^2 \implies \rho\cos\theta = \rho^2\sin^2\theta \implies \rho = \frac{\cos\theta}{\sin^2\theta}.$$

Noting that the origin is contained in this region, we know the lower bound for $\rho$ is 0.

Here is one careful analysis:

If the needle is sent to exit the region below the intersection of the paraboloid and the sphere, then the upper bound of $\rho$ is the paraboloid; otherwise, the upper bound is the sphere. So it is of paramount important that we find the point of intersection (see sketch on the next page).

To find the intersection, we solve the simultaneous equations

$$\begin{aligned} x^2+y^2+z^2 &= 2; \\ z &= x^2+y^2. \end{aligned}$$

Substituting the second in the first, we have

$$z^2+z=2 \implies z^2+z-2=0 \implies (z+2)(z-1)=0 \implies z=1.$$

We don't care about the negative root because the paraboloid is always positive (there is actually no intersection at $z=-2$). So, the intersection of the paraboloid and the sphere is the horizontal plane $z=1$. We then draw the sideview, and found that the critical angle separating the cases we spoke of is $\frac{\pi}{4}$ (hypotenuse $\sqrt{2}$ and adjacent 1).

Spelling out the cases, we have

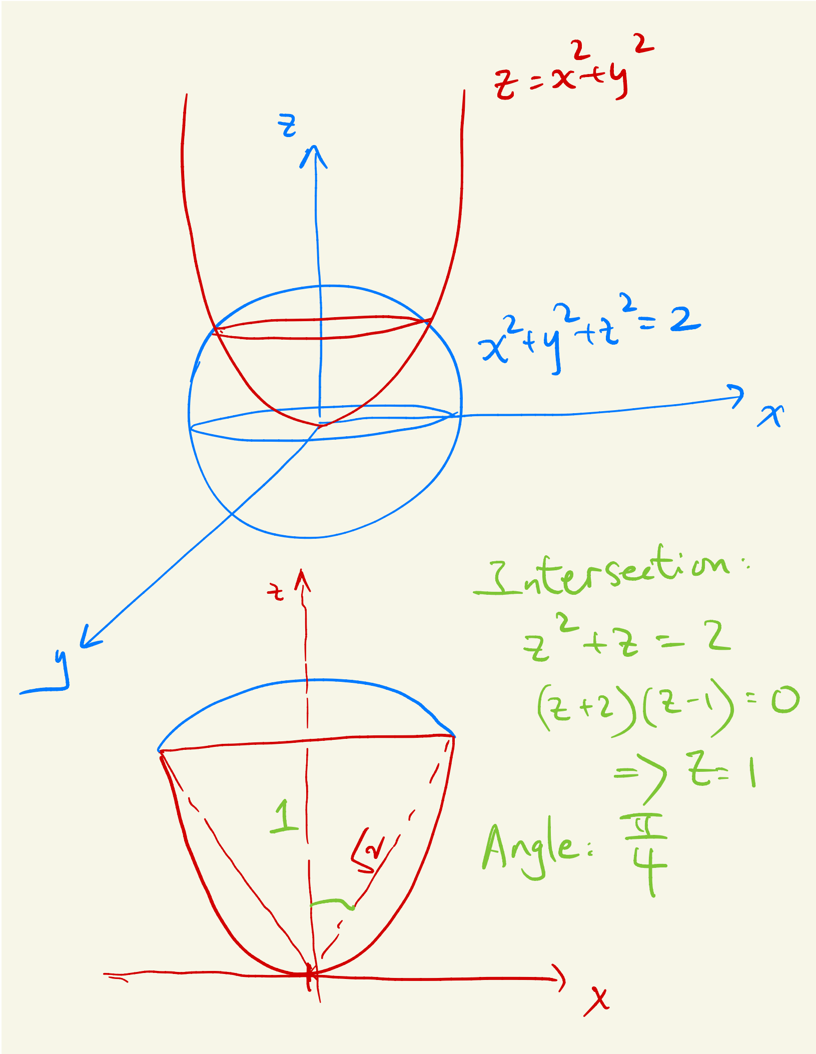
$$\begin{aligned} &\text{(needle leaves at the sphere)} \quad 0 \le \theta \le \frac{\pi}{4}, \quad 0 \le \rho \le \sqrt{2} \\ &\text{(needle leaves at the paraboloid)} \quad \frac{\pi}{4} \le \theta \le \frac{\pi}{2}, \quad 0 \le \rho \le \frac{\cos\theta}{\sin^2\theta} \end{aligned}$$

The reason why the second case has $\theta$ up to $\pi/2$ (rather than $\pi$) is that the intersection does NOT go below $z=0$ into the bottom four octants.

The bounds for $\phi$, the azimuthal angle, can be found by looking at the bird's-eye view of the intersection, which is a full circle about the origin – $0 \le \phi \le 2\pi$.

Altogether, the volume of the region of interest is

$$\begin{aligned} V &= \iiint dV \\ &= \int_0^{2\pi}\int_0^{\frac{\pi}{4}}\int_0^{\sqrt{2}} \rho^2\sin\theta d\rho d\theta d\phi + \int_0^{2\pi}\int_{\frac{\pi}{4}}^{\frac{\pi}{2}}\int_0^{\frac{\cos\theta}{\sin^2\theta}} \rho^2\sin\theta d\rho d\theta d\phi \end{aligned}$$


$z = x^2 + y^2$
$x^2+y^2+z^2=2$
Intersection:
$z^2+z=2$
$(z+2)(z-1)=0$
$\Rightarrow z=1$
Angle: $\frac{\pi}{4}$
1
$\sqrt{2}$

# 25 Lecture XV: Integrals in General Coordinates

## 25.1 Motivation from Single Variable Integrals

In 1D, we have made use of substitutions for integrals such as

$$\int_a^b f(g(t))\,g'(t)\,dt = \int_{g(a)}^{g(b)} f(u)\,du.$$

Note that the transform made here is

$$u = g(t), \quad du = g'(t)\,dt.$$

**The goal of making such a transformation is to simplify the integrand so that an antiderivative is more easily identified.**

**Remark 25.1** (Physical interpretation)**.** *In physical terms, we are saying that the variable $u$ represents the displacement of a particle in space (tracing its trajectory). It is parametrized by the function $g(t)$. Thus, $g'(t)$ represents the velocity of the particle.*

*You can consider the LHS integral as if you are observing at a fixed location, so you see the particle flying around with velocity $g'(t)$. The integral on the RHS then is viewed as if you are traveling with the particle, so the velocity (relative to you) is irrelevant (does not explicitly show up in the integrand, but instead, in the bounds).*

The idea behind a substitution or change of variable is a mapping between two intervals. In the example above, we start with $t \in [a,b]$ and ended with $g(t) \in [g(a), g(b)]$. This implies that $g$ is a function that relates values from $[a,b]$ to $[g(a), g(b)]$. The price to pay is the additional factor $g'(t)$.

**Example 25.1.** Consider the integral $\int_0^{10} e^{-x/5}dx$. One may consider the substitution

$$x = g(u) = 5u \implies g'(u) = 5.$$

Thus,

$$e^{-x/5} = e^u, \quad dx = 5du.$$

What did $g$ do to the interval $[0,10]$? $g : \underbrace{[0,2]}_{u\text{-space}} \to \underbrace{[0,10]}_{x\text{-space}}$, that is, $g$ maps from the $u$-space to the $x$-space. The problem is made **easier** in $u$-space. So,

$$\int_0^{10} e^{-x/5}dx = \int_0^2 e^{-u}5du.$$

**Remark 25.2.** *The workflow is*

$$\begin{array}{ccc} u\text{-space} & \overset{x=g(u)}{\rightarrow} & x\text{-space} \\ \text{easier problem} & & \text{harder problem} \end{array}$$

*Solve the easier problem in u-space, and via g, we know how to translate the solution to x-space (if we have to).*

## 25.2 Change of Variables in Two Dimensions

In two dimensions and above, we now map a region $R_1$ to another region $R_2$ (rather than from one interval to another) using a multivariable function $\Phi$ whose domain and range have the same dimensions. See the following example on polar transformation, which we are already quite familiar with.

**Example 25.2.** Polar change of variable

$$\begin{aligned} x\left(r,\theta\right) &= r\cos\theta \\ y\left(r,\theta\right) &= r\sin\theta \end{aligned}$$

where we identify the (map of) transformation $\Phi\left(r,\theta\right)=\left(x\left(r,\theta\right),y\left(r,\theta\right)\right)=\left(r\cos\theta,r\sin\theta\right)$. However, we used geometry to identify the extra factor we needed for the differential, $dxdy=rdrd\theta$.

We first show the results. The extra factor for an **arbitrary** transformation

$$\Phi\left(u,v\right)=\left(x\left(u,v\right),y\left(u,v\right)\right)=x\left(u,v\right)\mathbf{i}+y\left(u,v\right)\mathbf{j}$$

that makes between $(x,y)$ and $(u,v)$ satisfies the following relationship

$$dA=dxdy=\left|J\left(u,v\right)\right|dudv$$

where

$$J\left(u,v\right)=\frac{\partial\left(x,y\right)}{\partial\left(u,v\right)}=\left[\begin{array}{cc} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{array}\right],$$

is called the **Jacobian** of the change of variables, and

$$\left|J\left(u,v\right)\right|=\frac{\partial x}{\partial u}\frac{\partial y}{\partial v}-\frac{\partial x}{\partial v}\frac{\partial y}{\partial u},$$

where the $|\cdot|$ operation is called **the determinant**.

Let's check whether this formula holds for polar change of variables.

**Example 25.3.** The polar transformation is

$$\begin{aligned} x(r,\theta) &= r\cos\theta \\ y(r,\theta) &= r\sin\theta \end{aligned}$$

and therefore,

$$\begin{aligned} |J(r,\theta)| = \left|\frac{\partial(x,y)}{\partial(r,\theta)}\right| = \left|\begin{array}{cc} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{array}\right| = \frac{\partial x}{\partial r}\frac{\partial y}{\partial \theta} - \frac{\partial x}{\partial \theta}\frac{\partial y}{\partial r} \\ = (\cos\theta)(r\cos\theta) - (-r\sin\theta)(\sin\theta) = r\left(\cos^2\theta + \sin^2\theta\right) = r \end{aligned}$$

and thus

$$dA = dxdy = rdrd\theta,$$

as expected.

## 25.3 Why this extra factor? (Optional)

There is a beautiful result from geometry and linear algebra.

**Theorem 25.1.** *Given two vectors $\vec{u} = (u_1, u_2)$ and $\vec{v} = (v_1, v_2)$, the parallelogram it spans has area given by the determinant of the following matrix*

$$A = \begin{bmatrix} u_1 & u_2 \\ v_1 & v_2 \end{bmatrix}$$

*which is*

$$|A| = \begin{vmatrix} u_1 & u_2 \\ v_1 & v_2 \end{vmatrix} = u_1v_2 - u_2v_1.$$

*Proof.* See this page. □

So, how does the area of a parallelogram have to with the extra factor $|J(u,v)|$? Read the Change_of_variables_Jacobian_sketch.pdf.

## 25.4 Examples with Effective Change of Variables

Now, let's face several integrals where a change of variable will considerably simplify the integrand.

**Example 25.4.**

$$\int_0^1 \int_0^{1-x} \sqrt{x+y}\,(y-2x)^2\,dydx$$

**Solution.** The transformation is meant to simplify the integrand. Taking this integral head-on would be idiotic.

Here, the good choices are to eliminate the variable under the square root and the square, by choosing (lots of freedom here)

$$u = x + y, \quad v = y - 2x$$

**Step 1: Isolate $x$ and $y$ and compute $|J(u,v)|$**

Now, we need to isolate $x$ and $y$ by solving the above system for it, that is

$$x = \frac{u}{3} - \frac{v}{3}, \quad y = \frac{2u}{3} + \frac{v}{3}$$

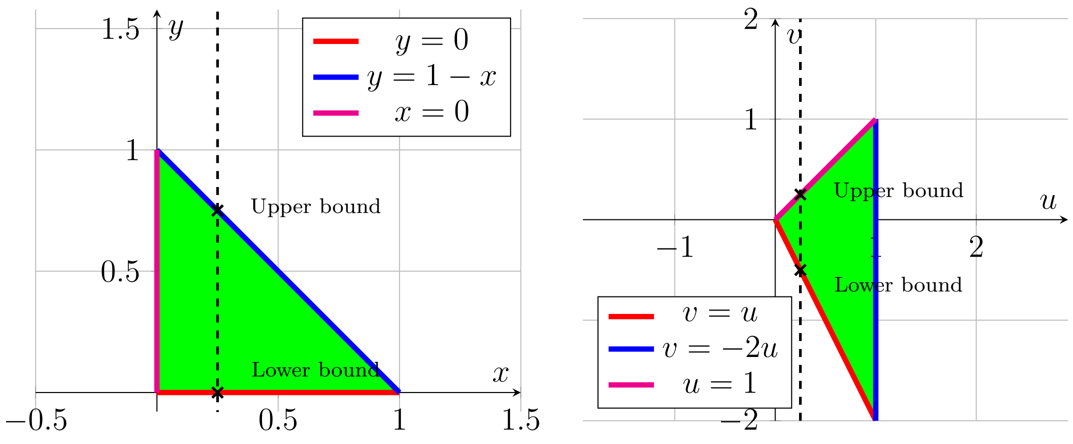
Good, we can also find the Jacobian of this transformation,

$$|J(u,v)| = \left|\frac{\partial(x,y)}{\partial(u,v)}\right| = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \frac{\partial x}{\partial u}\frac{\partial y}{\partial v} - \frac{\partial x}{\partial v}\frac{\partial y}{\partial u} = \left(\frac{1}{3}\right)\left(\frac{1}{3}\right) - \left(-\frac{1}{3}\right)\left(\frac{2}{3}\right) = \frac{1}{3}.$$

**Step 2: Rewrite the Boundary**

Always remember to change the bounds of integration as well. The original region in the $xy$-plane is a triangle, bounded by the axes $y = 0$, $x = 0$ and the line $y = 1 - x$. Using this transformation, we find that (following the same color scheme as in Figure 14)

$$\begin{aligned} y = 0 &\iff v = -2u \\ y = 1 - x &\iff u = 1 \\ x = 0 &\iff v = u \end{aligned}$$




(a) $xy$-space (b) $uv$-space

Figure 14: Region of integration. (a) Before transformation; (b) After transformation. Consistent color scheme for the transformed pieces.

We can write the bounds more concretely as

$$\begin{aligned}-\frac{u}{2} &\leq v \leq u, \\ 0 &\leq u \leq 1,\end{aligned}$$

and thus modify the integral as

$$\begin{aligned}\int_0^1 \int_0^{1-x} \sqrt{x+y}\,(y-2x)^2\,dydx &= \int_0^1 \int_{-2u}^{u} \sqrt{u}v^2\,|J(u,v)|\,dvdu \\ &= \int_0^1 \int_{-2u}^{u} \sqrt{u}v^2 \left(\frac{1}{3}\right) dvdu \\ &= \frac{1}{3}\int_0^1 \sqrt{u}\left(\int_{-2u}^{u} v^2 dv\right) du \\ &= \frac{1}{3}\int_0^1 \sqrt{u}\left[\frac{v^3}{3}\right]_{-2u}^{u} du \\ &= \frac{1}{9}\int_0^1 \sqrt{u}\left(u^3 - (-2u)^3\right) du \\ &= \frac{1}{9}\int_0^1 9u^{7/2} du \\ &= \frac{2}{9}u^{9/2}\,\Big|_0^1 \\ &= \frac{2}{9}.\end{aligned}$$

This integral would otherwise be impossible.

**Example 25.5.**

$$\int_1^2 \int_{1/y}^{y} \sqrt{\frac{y}{x}} e^{\sqrt{xy}} dxdy$$

**Solution.** The nasty parts are definitely the functions of two variables under the square roots. The idea of the previous problem applies. We set

$$u = \sqrt{xy}, \quad v = \sqrt{\frac{y}{x}}$$

which then implies (after some algebra)

$$x = \frac{u}{v}, \quad y = uv$$

with $u > 0$ and $v > 0$ (since the square root only produces nonnegative numbers). The Jacobian is found by definition,

$$J(u,v) = \frac{\partial x}{\partial u}\frac{\partial y}{\partial v} - \frac{\partial y}{\partial u}\frac{\partial x}{\partial v} = \left(\frac{1}{v}\right)(u) - (v)\left(-\frac{u}{v^2}\right) = \frac{2u}{v}.$$

Thus, the integrand becomes

$$\sqrt{\frac{y}{x}}e^{\sqrt{xy}}dxdy = ve^u\left(\frac{2u}{v}\right)dudv = \boxed{2ue^u}dudv.$$

Notice that the "price to pay" here is no longer a constant, but a function $2ue^u$.

We need to check how the boundaries have changed by this change of variables. The original boundary involves three pieces, which we investigate each below:

1. $y = 2$.

   Plugging this in the definition of $u$ and $v$ gives

   $$u = \sqrt{2x}$$
   $$v = \sqrt{\frac{2}{x}}$$

   Multiplying gives

   $$uv = 2 \implies v = \frac{2}{u}$$

   which is a reciprocal function in the $uv$-plane.

2. $y = 1/x$.

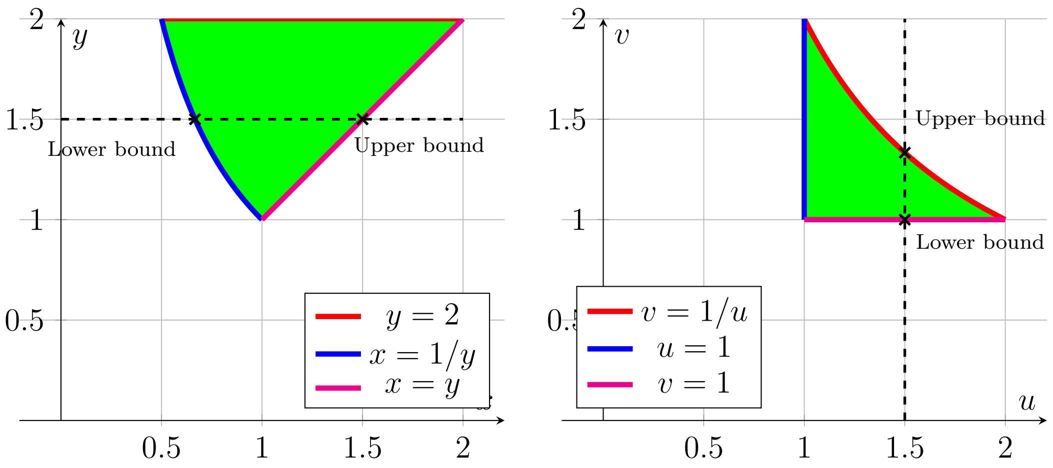
   $$u = \sqrt{x\frac{1}{x}} = 1$$
   $$v = \sqrt{\frac{1}{x^2}} = \frac{1}{x}$$

   Don't care about the second part since it is NOT a description between $x$ and $y$. However, $u = 1$ means a vertical line in the $uv$-plane. Take it.

3. $y = x$.

   $$u = \sqrt{x^2} = x$$
   $$v = \sqrt{\frac{x}{x}} = 1$$

   where now we ignore the second part. $v = 1$ is a horizontal line in the $uv$-plane.




(a) $xy$-space



(b) $uv$-space

Figure 15: Region of integration. (a) Before transformation; (b) After transformation. Consistent color scheme for the transformed pieces.

Altogether, in the $uv$-plane, we have three new pieces that enclose the new region of integration, $u = 1$, $v = 1$ and $v = \frac{2}{u}$ (how to graph $y = 2/x$?), shown in Figure 15.

To find the bounds of integral for $u$ and $v$, pick whichever variable to start. For me personally, it is easier to do $v$ first. Send a $v$needle, i.e., a vertical line (parallel to $v$-axis) through the region. It enters the region at $v = 1$ and exits at the curve $v = 2/u$ or equivalently $v = 2/u$. Then, it is clear that $1 \leq u \leq 2$.

Thus, our original integral now becomes

$$\int_1^2 \int_{1/y}^{y} \sqrt{\frac{y}{x}} e^{\sqrt{xy}} dxdy = \boxed{\int_1^2 \int_1^{2/u} 2ue^u dvdu}$$

Would it be ok to do $u$-needle first? Of course. You will end up with

$$\int_1^2 \int_{1/y}^{y} \sqrt{\frac{y}{x}} e^{\sqrt{xy}} dxdy = \int_1^2 \int_1^{2/v} 2ue^u dudv$$

But you must notice that integrating $ue^u du$ is far harder (integration by parts with variable bounds, ugh!) than doing $ue^u dv$ (the latter makes $u$ a constant, yay!). Because of this, we prefer the first formulation (in the box).

Finally, we follow the integral in the box and just integrate

$$\int_1^2 \int_1^{2/u} 2ue^u dvdu = \int_1^2 2ue^u \left( \int_1^{2/u} dv \right) du$$

$$
\begin{aligned}
&= 2\int_1^2 ue^u \left[\frac{2}{u} - 1\right] du \\
&= 2\int_1^2 \left(2e^u - ue^u\right) du \\
&= 2\int_1^2 (2-u)\, e^u du \\
&= 2\left[(2-u)\, e^u + e^u\right]_1^2 \\
&= 2e\,(e-2)
\end{aligned}
$$

where we did integration by parts once.

# 26 Lecture XVI: An Application in Solar Panels

A classic application of Integrals in General Coordinates is the integral over a rhombus. Consider the following application in solar panel.

**Example 26.1** (Total Power of a Solar Panel)**.** Consider a solar panel shaped as a rhombus, described by $|x| + |y| \leq 1$. The light intensity at each coordinate $(x, y)$ can be modeled by the function

$$I(x, y) = \cos(x - y) e^{-|x+y|}$$

i.e., the power per unit area (watt per meter squared). Compute the total power across the solar panel.

**Solution.** The total power is the integral over the rhombus

$$\iint_{|x|+|y|<1} \cos(x - y) e^{-|x+y|} dA.$$

If we work with the rhombus, sending either the $x$-needle or the $y$-needle leads to massive complications.

Instead, we make a change of variables

$$\begin{cases} u &= x - y \\ v &= x + y \end{cases} \implies \begin{cases} x &= \frac{u+v}{2} \\ y &= \frac{v-u}{2} \end{cases}$$

The Jacobian of this transformation is given by

$$J(u, v) = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix} \implies |J(u, v)| = \left(\frac{1}{2}\right)\left(\frac{1}{2}\right) - \left(-\frac{1}{2}\right)\left(\frac{1}{2}\right) = \frac{1}{2}.$$

Thus, the integrand becomes

$$\cos(x - y) e^{-|x+y|} dA = \cos(u) e^{-|v|} \frac{1}{2} du dv.$$

To transform the boundary, we consider the four separate pieces

1. $x - y = -1$

   Using the transformation, we find that

   $$-1 = x + y = \frac{u+v}{2} - \frac{v-u}{2} = u.$$

   So, this line is transformed to the vertical line $u = -1$.

2. $x + y = 1$

   Using the transformation, we find that

$$1 = x + y = \frac{u+v}{2} + \frac{v-u}{2} = v.$$

   So, this line is transformed to the horizontal line $v = 1$.

3. $x - y = 1$

   Using the transformation, we find that

$$1 = x + y = \frac{u+v}{2} - \frac{v-u}{2} = u.$$

   So, this line is transformed to the vertical line $u = 1$.

4. $x + y = -1$

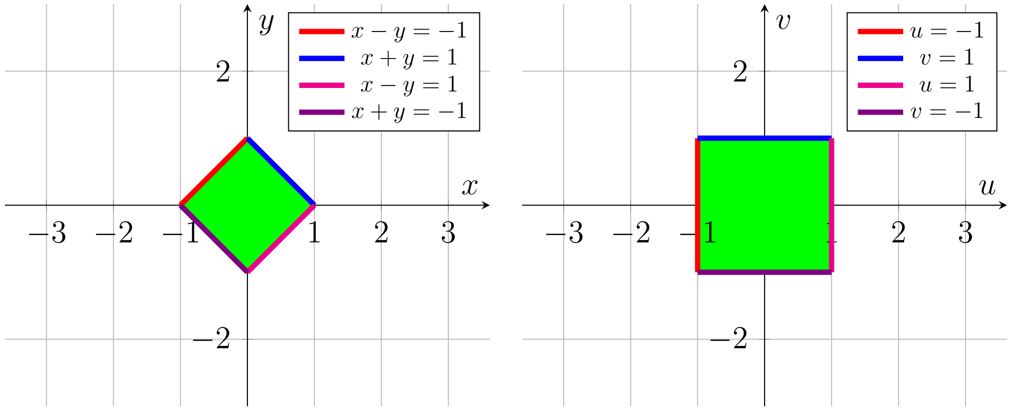
   Using the transformation, we find that

$$-1 = x + y = \frac{u+v}{2} + \frac{v-u}{2} = v.$$

   So, this line is transformed to the horizontal line $v = -1$.

Thus, the original integral is significantly simplified not only in the integrand, but also at the boundary.

$$\begin{aligned}
\iint_{|x|+|y|<1} \cos(x-y)\, e^{-|x+y|} dA &= \int_{-1}^{1}\int_{-1}^{1} \cos(u)\, e^{-|v|} du dv \\
&= \left(\int_{-1}^{1} \cos(u)\, du\right)\left(\int_{-1}^{1} e^{-|v|} dv\right) \\
&= 2\sin(1)\, 2\left(1 - \frac{1}{e}\right) \\
&= 4\sin(1)\left(1 - \frac{1}{e}\right)
\end{aligned}$$




(a) $xy$-space (b) $uv$-space

Figure 16: Region of integration. (a) Before transformation; (b) After transformation. Consistent color scheme for the transformed pieces.