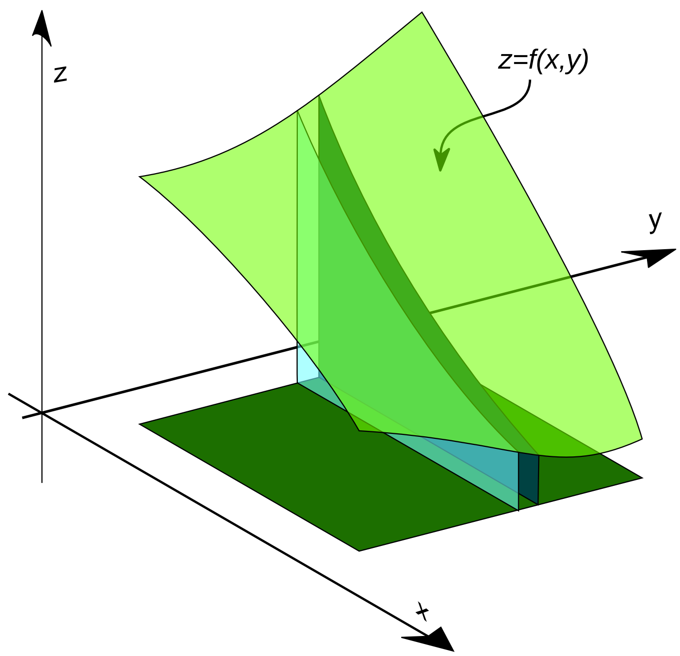


MATH 234
THIRD SEMESTER
CALCULUS

Fall 2009

Math 234 – 3rd Semester Calculus
Lecture notes version 0.9(Fall 2009)

This is a self contained set of lecture notes for Math 234. The notes were written by Sigurd Angenent, many problems and parts of the text were taken from Guichard's open calculus text which is available at http://www.whitman.edu/mathematics/multivariable/src/

The LaTeX files, as well as the PYTHON and INKSCAPE-SVG files which were used to produce the notes before you can be obtained from the following web site

http://www.math.wisc.edu/~angenent/Free-Lecture-Notes

They are meant to be freely available for non-commercial use, in the sense that "free software" is free. More precisely:

Contents

CHAPTER 1

Functions of two and more variables.

1. n-dimensional space

The line is one-dimensional, the plane is two dimensional, and the space around us is three dimensional[1]

A point on the line is specified by one coordinate "x", a point in the plane by two coordinates, "(x, y)", and a point in three dimensional space can be specified by three coordinates (x, y, z). Going on like that, a point in 56-dimensional space is specified by 56 coordinates, $(x_1, x_2, \ldots, x_{55}, x_{56})$. Instead of getting philosophical about what n-dimensional space really is ("does it exist?"), we simply say that a point in n-dimensional space is a list of n-real numbers, $(x_1, \ldots, x_n)$ and that, as far as mathematics is concerned, n-dimensional space is just the collection of all possible lists $(x_1, \cdots, x_n)$ of n numbers. If $n = 1, 2$, or 3, then we can visualize such a point by drawing one, two or three axes; if $n = 4$ or more, then we can't, but it doesn't matter.

The symbol $\mathbb{R}^n$ is used to stand for n-dimensional space, meaning the collection of all such lists of n numbers $(x_1, \ldots, x_n)$.

In this course we will mostly deal with $\mathbb{R}^2$ and $\mathbb{R}^3$, although much of what we do works (and gets used) without modification in $\mathbb{R}^n$.

2. Functions of two or more variables

2.1. The graph of a function. In first-year calculus we were concerned with functions of one variable, meaning the "input" is a single real number and the "output" is likewise a single real number. At the end of math 222 we considered functions taking a real number to a vector: for each input value we get a position in space. Now we turn to functions of several variables, meaning several input variables, functions. While we will deal primarily with $n = 2$ and to a lesser extent $n = 3$, many of the techniques we discuss can be applied to larger values of n as well.

A function of two variables maps a pair of values (x, y) to a single real number. The three-dimensional xyz-coordinate system is a convenient aid in visualizing such functions: above each point (x, y) in the xy-plane we graph the point (x, y, z), where of course $z = f(x, y)$.

2.2. Vector notation. We will use vectors all the time in this course. If $\vec{\boldsymbol{x}}$ is the position vector of the point (x, y) in the plane, i.e. if $\vec{\boldsymbol{x}} = \left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right)$, then one writes

$$f(x, y) = f(\vec{\boldsymbol{x}}).$$

2.3. Example. Consider $f(x, y) = 3x + 4y - 5$. Writing this as $z = 3x + 4y - 5$ and then $3x + 4y - z = 5$ we recognize the equation of a plane. In the form $f(x, y) = 3x + 4y - 5$ the emphasis has shifted: we now think of x and y as independent variables and z as a variable dependent on them, but the geometry is unchanged.

[1]Although some physicists will tell you it's really 11 or 24 dimensional.

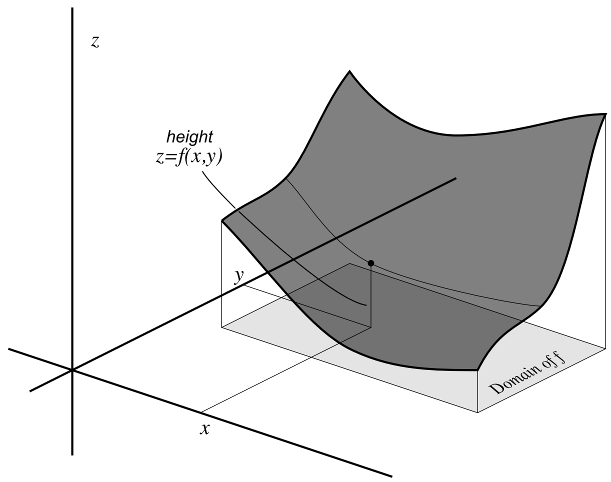
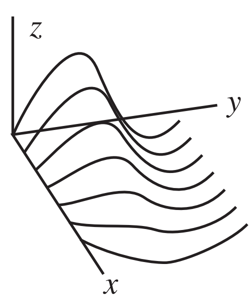


Figure 1: The graph of some function, and its domain (a rectangle in this example).

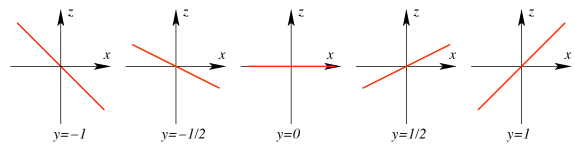
2.4. Example. You know that $x^2+y^2+z^2=4$ represents a sphere of radius 2. We cannot write this in the form $z=f(x,y)$, since for each x and y in the disk $x^2+y^2<4$ there are two corresponding points on the sphere. As with the equation of a circle, we can resolve this equation into two functions, $f(x,y)=\sqrt{4-x^2-y^2}$ and $f(x,y)=-\sqrt{4-x^2-y^2}$, representing the upper and lower hemispheres. Each of these is an example of a function with a restricted domain: only certain values of x and y make sense (namely, those for which $x^2+y^2\le 4$) and the graphs of these functions are limited to a small region of the plane.

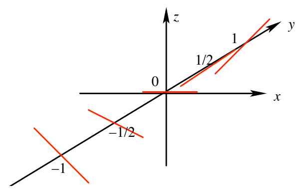
2.5. Freezing a variable. If a function isn't familiar, then a good strategy for drawing its graph is to "***freeze a variable.***" In other words, to analyze a function $z=f(x,y)$ you pretend y is a constant: then x is the only independent variable, and you can try to draw the graph of the function $z=f(x,y)$, now thinking of this as a function of only one variable. This graph is a curve in the xz plane. You get one such curve for each choice of y. Piecing these graphs together then gives you the graph of the two-variable function $z=f(x,y)$.

You could apply the same procedure with the roles of x and y switched: i.e. for each fixed x you try to graph $z=f(x,y)$ as a function of the variable y only, after which you try to fit all the graphs you get for different values of x together.

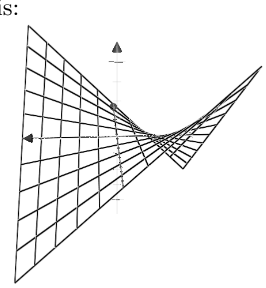


2.6. Example – draw the graph of $f(x,y)=xy$. Let's plot the graph of $z=f(x,y)=xy$. For each fixed value of y the graph of $f(x,y)=xy$ is a straight line with slope y. For positive y the line has positive slope, for negative y it has negative slope. Plotting the graphs of $z = xy$ for y frozen at the values -1,$-\frac{1}{2}$, 0, $\frac{1}{2}$, and 1 gives us these drawings:

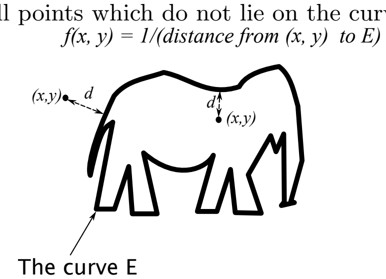




The function $z = xy$ is symmetric in the x and y variables, so you get similar pictures if you freeze x and graph $z = xy$ as a function of y. Carefully putting both pictures together gives something like this:



2.7. The domain of a function. Just as with functions of one variable, functions of two variables have a ***domain***, consisting of all the points (x, y) in the plain for which $f(x, y)$ is defined. For functions of one variable the domain is usually an interval, but for functions of two variables the domain can have more interesting shapes. In the drawing on the left here, the function $f(x, y)$ is defined to be the inverse of the distance from the point (x, y) to the curve E in the picture. This function is only defined when this distance is not zero (otherwise you can't divide by the distance...), so the domain of this function consists of all points which do not lie on the curve.



2.8. Example. What is the domain of the function

$$f(x, y) = \frac{1}{\sqrt{1 - x - y}} \quad ??$$

Clearly the function is defined if the quantity under the square root is nonnegative (otherwise you can't take the square root), and not zero (otherwise you can't divide by the resulting square root). So the domain consists of all points with $1 - x - y > 0$, or, equivalently, $y < 1 - x$. The domain consists of all points in the plane which line below the graph of $y = 1 - x$.

3. Open and closed sets in $\mathbb{R}^n$

Intervals in the real line come in four kinds, depending on whether they include their endpoints or not: you can have (a, b), $(a, b]$, $[a, b)$ and $[a, b]$, and those are all the possibilities. With domains in the plane, or in space there are many more possibilities, and it will sometimes be important to distinguish between domains which include all their "endpoints" and those that don't. In the present context one doesn't say "endpoint" but speaks of ***boundary point*** instead. To define what a boundary point is, it turns out that you need to resort to ε and δ again, or a least to ε. Here is some terminology which we will use:

- $B_r(p)$ is the ball with center p and radius r.
- $G \subset \mathbb{R}^n$ is ***open*** if for every point p in G there is an $\varepsilon > 0$ such that G contains $B_\varepsilon(p)$.
- $G \subset \mathbb{R}^n$ is ***closed*** if its complement is open.
- p is a ***boundary point*** of G if $B_r(p)$ always contains both points from G and from its complement, no matter how small you choose $r > 0$.

The following intuitive description is good enough for math 234: G is closed if it contains all its boundary points; G is open if it contains none of its boundary points.

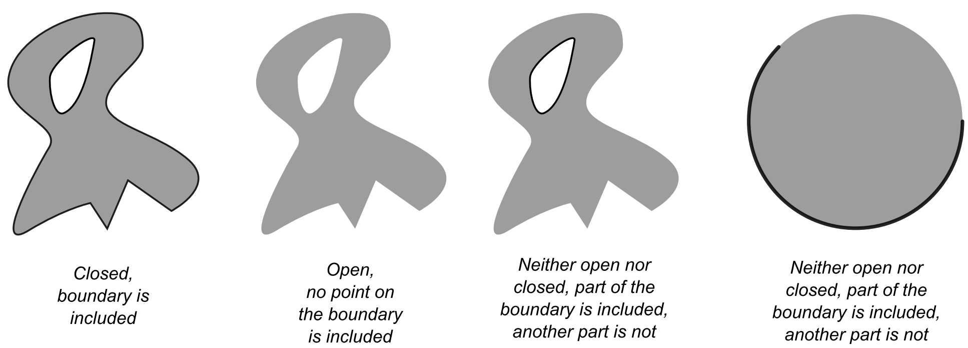


Figure 2: Some domains in the plane. Points in the domain are shaded gray. Boundary points which are included in the domain are marked in black.

3.1. Example. Consider the three domains

$$G_1 = \text{all points } (x, y) \text{ with } x^2 + y^2 < 1$$

$$G_2 = \text{all points } (x, y) \text{ with } x^2 + y^2 \leq 1$$

$$G_3 = \text{all points } (x, y) \text{ with } -1 \leq x \leq 1 \text{ and } -\sqrt{1 - x^2} < y \leq \sqrt{1 - x^2}$$

For all three domains the boundary points are the points on the unit circle. G_1 contains none of its boundary points, so it is called "open"; G_2 contains all its boundary points, so it is called "closed"; G_3 contains some but not all of its boundary points, so it is neither open nor closed.

4. More examples of visualization of Functions

You can visualize a function f of two variables by means of its graph, but this is not the only way. There are at least two alternatives. The first is in terms of ***level sets***, the other is as a ***movie of a graph of a function of one variable***.

Level sets are defined as follows. Given a function $z = f(x, y)$ and a number c, the level set at level c is the set of all points in the plane which satisfy $f(x, y) = c$; in symbols,

$$\text{“Level set of } f \text{ at level } c\text{”} \stackrel{\text{def}}{=} \{(x, y) : f(x, y) = c\}\,.$$

To describe a function in terms of its level sets, one usually picks a range of values for the constant c and draws the level sets corresponding to the chosen values of c in one figure.

While the graph is a three-dimensional object, the level set is a set of points in the plane, usually a curve. Level sets are therefore easier to draw than graphs.

4.1. Example. What are the level sets of the function $f(x, y) = 3 - x - y$?

For any given number c the level set at level c of f contains exactly those points (x, y) which satisfy $f(x, y) = c$, i.e. $3 - x - y = c$. This is a line, and it is the graph of $y = 3 - c - x$: so it is the line with slope -1 and “y-intercept” $3 - c$.

4.2. Level sets of the saddle surface. What are the level sets of the function whose graph we drew in § 2.6?

The function was given by $f(x, y) = xy$, so the level set at level c consists of all points (x, y) in the plane which satisfy $xy = c$. For instance, if $c = 1$, then you get the familiar hyperbola $y = 1/x$. For other positive values of c you get similar hyperbolas, and for negative c you get hyperbolas in the 2nd and 4th quadrants.

The level at $c = 0$ is exceptional because it is not a hyperbola, but rather consists of two crossing lines. Namely, $xy = 0$ holds when either $x = 0$ or $y = 0$ holds, so the level set at $c = 0$ is the union of the x-axis and the y-axis.

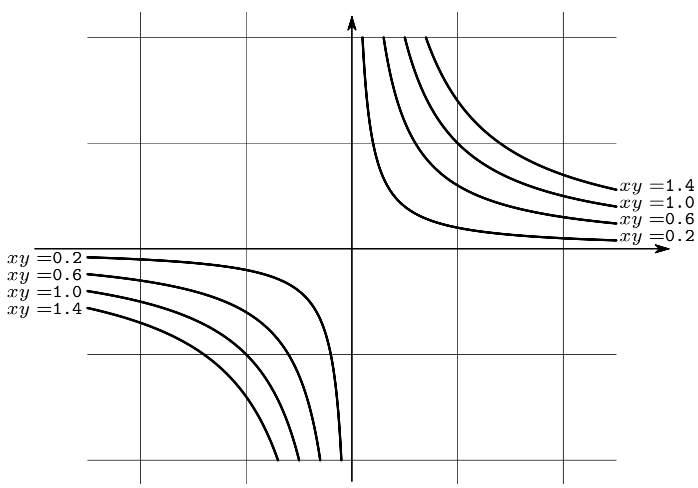
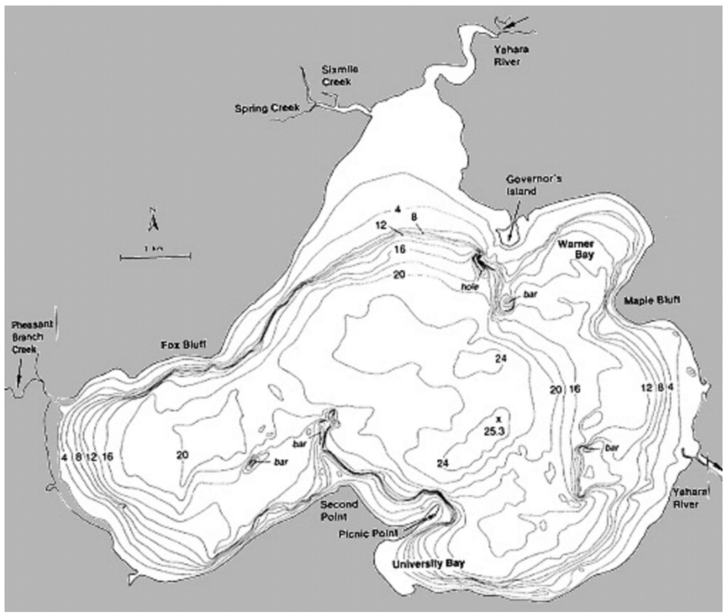


Figure 3: A few level sets of the function $f(x, y) = xy$. Only positive levels are shown.

4.3. An example from the "real" world. Here is a function of local interest. The domain of the function is the water surface of Lake Mendota (let's pretend this is a plane domain), and the function, which I'll call d instead of f, is given by $d(x, y) =$ the depth of the lake at location (x, y). There's no formula for this function, but the limnology department of the UW has measured the depth and presented the results in terms of the level sets of the function d.

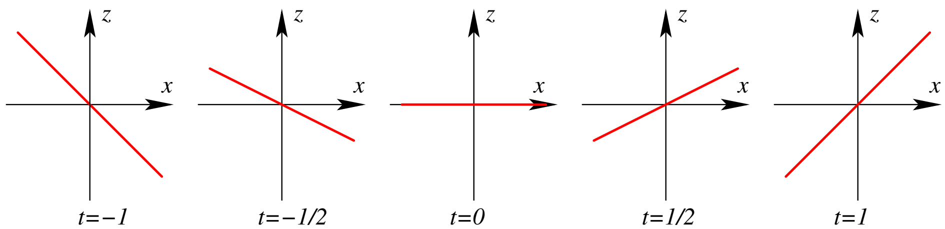


The level curves of a function $z = d(x, y)$. The domain of this function is the lake, and $d(x, y)$ is the depth in meters of Lake Mendota at (x, y).
See http://limnology.wisc.edu/lake_information/mendota/mendota.html

4.4. Moving graphs. There's another way of visualizing a function $z = f(x, y)$ of two variables where you think of one of the independent variables (e.g. y) as "time." The final picture is not one static picture of a three dimensional surface, but rather a movie of a graph which is moving around in the xz plane.

If you have a function $z = f(x, y)$, then let us think of y as time, and let us relabel it as t, so that we are looking at the function $z = f(x, t)$. Now at each moment in time t we have a function $z = f(x, t)$ of one variable x whose graph you can try to draw. Think of this graph as a still-image. Then as you let time t vary, putting the still images in a sequence, you get a movie of a graph of a changing function of one variable.

For instance, if the function is once again the saddle surface function $z = xy$, then we would be considering the function $z = xt$. At each moment t the graph of $z = xt$ is a line with slope t. Putting together these graphs gives a movie of a line which begins with a line of rather negative slope; during the movie the slope increases, and in the middle our line has achieved horizontality; finally, the closing shot presents us with a line with a very positive slope. Here are some stills from the movie:



So you see that this interpretation is not very different from the procedure of "freezing the y variable." The only real difference lies in what you do with all the separate graphs you get after you freeze a variable. In one case you try to piece them together to make a bigger drawing of a three-dimensional object, in the other you put them together to make a motion picture.

Problems

In the problems in this stage of the course, you will be asked to "sketch the graph of a function." From math 221 you remember that this meant you had to find minima, maxima, inflection points, and other features of the graph. In 234 you will learn to do the same for functions of two (and more) variables, but for now you should try to use the method of "freezing a variable" or other similar tricks to get an idea of what the graph of f looks like.

You can use a graphing program (such as `Grapher.app` on the Mac, and `GraphCalc` on Windows) to check your answer.

> Note: very often students try to fit their drawings into a region the size of a post-it. In this course, whenever you make a drawing, especially if it's a three-dimensional drawing, **make it large!** Use half a page for a drawing. Make sure you have enough paper, try to find lots of cheap scrap paper.

1. Make careful drawings of the graphs of the three functions in the examples in §2.3, and §2.4.

Find the domain of these functions. Also, label the axes in every drawing you make.

2. Which functions of two variables $z = f(x, y)$ are defined by the following formulae? Find the domain of each function. Then draw the domain. Try to sketch their graphs.

(i) $z - x^2 = 0$	**(ii)** $z^2 - x = 0$	**(iii)** $z - x^2 - y^2 = 0$
(iv) $z^2 - x^2 - y^2 = 0$	**(v)** $xyz = 1$	**(vi)** $xy/z^2 = 1$
(vii) $x + y + z^2 = 0$	**(viii)** $x + y + z^2 = 1$	

3. Figure 3 only presents level sets $f(x, y) = c$ of the function $f(x, y) = xy$ for some **positive** values of c. What does the zero set look like, and what do the level sets $f(x, y) = c$ with $c < 0$ look like?

4. Let Q be the square in the plane consisting of all points (x, y) with $|x| \le 1$, $|y| \le 1$. This problem is about the so-called **distance function** to Q. This function is defined as follows: $f(x, y)$ is the distance from the point (x, y) to the point in Q nearest to (x, y).

(i) Which point in Q is nearest to $(0, \frac{1}{2})$? Which is closest to $(0, 2)$? Which is closest to $(3, 4)$?

(ii) Compute $f(0, \frac{1}{2})$, $f(0, 2)$ and $f(3, 4)$).

(iii) What is the zero set of f?

(iv) Draw the level sets of f at levels -1, 1, 2, and 3. Describe the general level set $f(x, y) = c$ where c is an arbitrary number.

(v) Give a formula for $f(x, y)$. (It turns out too be hard to capture the distance function in one formula. You will have to split the plane into different regions and describe $f(x, y)$ by different formulas, according to which region (x, y) belongs to.)

5. If $d(x, y)$ is the depth function of Lake Mendota (see §4.3), then what are the level sets $d(x, y) = c$ for $c = 0$, $c = +10$ and for $c = -10$ (meters)? What is the level set $d(x, y) = 400$ (meter)?

6. For each of the functions in problem 2 draw the level sets at level $z = c$ for a few values of c (as was done in Figure 3 and § 4.3). What does the level set for an arbitrary c look like? Are they familiar curves?

7. Describe and explain the relation between the graph of the function $y = g(x)$ of one variable, and the corresponding function $f(x, y) = g\big(\sqrt{x^2 + y^2}\big)$ of two variables.

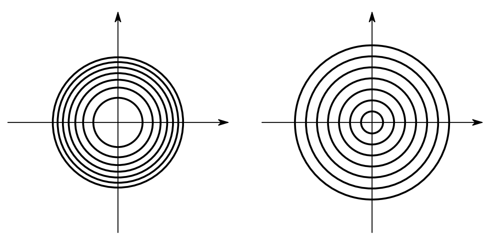
What do the level sets of $f(x, y)$ look like?

For instance, if $g(x) = x$, then $f(x, y) = \sqrt{x^2 + y^2}$: what is the relation between the graphs of g and f?

8. Find the domain of the following functions of two (or occasionally three) variables:

(**i**) $f(x, y) = \sqrt{9 - x^2} + \sqrt{y^2 - 4}$
(**ii**) $f(x, y) = \arcsin(x^2 + y^2 - 2)$
(**iii**) $f(x, y) = \sqrt{x} \cdot \sqrt{y}$
(**iv**) $f(x, y) = \sqrt{xy}$
(**v**) $f(x, y, z) = 1/\sqrt{xyz}$
(**vi**) $f(x, y) = \sqrt{16 - x^2 - 4y^2}$

9. Here are two sets of level curves with levels $z = 0.2, 0.4, 0.6, 0.8, 1.0, 1.2, 1.4$. One is for a function whose graph is a cone ($z = \sqrt{x^2 + y^2}$), the other is for a paraboloid ($z = x^2 + y^2$). Which is which? Explain.



Problems about movies

10. Describe the "movie" that goes with each of the following functions.

(**i**) $f(x, t) = x \sin t$
(**ii**) $f(x, t) = x \sin 2t$
(**iii**) $f(x, t) = t \sin x$
(**iv**) $f(x, t) = 2t \sin x$
(**v**) $f(x, t) = t \sin 2x$
(**vi**) $f(x, t) = (x - t)^2$
(**vii**) $f(x, t) = (x - \sin t)^2$
(**viii**) $f(x, t) = (x - t^2)^2$
(**ix**) $f(x, t) = \dfrac{t^2}{1 + x^2}$
(**x**) $f(x, t) = \dfrac{1}{(1 + x^2)(1 + t^2)}$

11. If $y = g(x)$ is any function of one variable, then a function of the form $f(x, t) = g(x - ct)$ is often called a **traveling wave** with wave speed c and profile g. Let g be any non constant function of your choice and describe the movie presented by the function $f(x, t) = g(x - ct)$ (can't choose? Then try "Agnesi's witch" $g(x) = \frac{1}{1+x^2}$.)

The number c is called the wave speed. If $c > 0$ is the motion to the left or to the right? Explain.

12. If $y = g(x)$ is any function of one variable, then a function of the form $f(x, t) = \cos(\omega t) g(x)$ is often called a **standing wave.** Let g be any non constant function of your choice and describe the movie presented by the function $f(x, t) = \cos(\omega t) g(x)$ (can't choose? Then try "Agnesi's witch" $g(x) = \frac{1}{1+x^2}$ again, or for this example, try $g(x) = \sin x$.)

The number $\frac{\omega}{2\pi}$ is called the frequency of the standing wave. The function $g(x)$ is called its profile. How long does it take before the standing wave returns to its original position, i.e. what is the smallest $T > 0$ for which $f(x, T) = f(x, 0)$ for all x? Explain.

About open and closed sets

13. Draw the sets G_1, G_2, G_3 from section 3.1 in the same style as figure 2 (i.e. shade the points in the region and mark the boundary points which are included in the region).

14. Using the intuitive description of when a set is open, closed, or neither of those, discuss which of the intervals $(0, 1)$, $[0, 1]$, $[0, 1)$, and $(0, 1]$ are open/closed/neither.

15. (for discussion) Can you split the plane into two sets, both of which are open?

5. Continuity and Limits

5.1. The limit of a function of two variables. Just as with functions of one variable we need to define the limit of $f(x, y)$ as (x, y) approaches some given point (a, b). There is again a precise definition involving epsilons and deltas, and it is in many ways pretty much the same definition as in math 221. Here it is:

5.2. Definition. *Let $f(x, y)$ be a function of two variables. Then we say that*

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

if for every $\varepsilon > 0$ you can find a $\delta > 0$ such that for all points (x, y) one has

$$(x, y) \text{ lies in } B_\delta(a, b) \implies |f(x, y) - L| < \varepsilon.$$

Remember that $B_\delta(a, b)$ is the disc with radius δ and center (a, b). The last line of the definition therefore says that you can be sure that $f(x, y)$ will be approximately equal to L with an error of no more than ε, provided you choose (x, y) so close to (a, b) that the distance between (x, y) and (a, b) is less than δ. The first part of the definition will say that, no matter which $\varepsilon > 0$ you come up with, a $\delta > 0$ can be found for which the second part is true.

In this course we will hardly ever use the above definition. When we have to compute limits we will use the limit properties, such as

$$\lim_{(x,y)\to(a,b)} f(x, y) \pm g(x, y) = \Big\{\lim_{(x,y)\to(a,b)} f(x, y)\Big\} \pm \Big\{\lim_{(x,y)\to(a,b)} g(x, y)\Big\}, \tag{1}$$

$$\lim_{(x,y)\to(a,b)} f(x, y)g(x, y) = \Big\{\lim_{(x,y)\to(a,b)} f(x, y)\Big\} \cdot \Big\{\lim_{(x,y)\to(a,b)} g(x, y)\Big\} \tag{2}$$

$$\lim_{(x,y)\to(a,b)} \frac{f(x, y)}{g(x, y)} = \frac{\lim_{(x,y)\to(a,b)} f(x, y)}{\lim_{(x,y)\to(a,b)} g(x, y)} \tag{3}$$

where the latter holds only if $\lim_{(x,y)\to(a,b)} g(x, y) \neq 0$, and the interpretation of these formulas is that ***if*** the expression on the right exists, ***then*** the limit on the left also exists, and both are equal.

5.3. Definition of Continuity. *A function $f(x, y)$ is called **continuous** at a point (a, b) in its domain if*

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

The precise meaning of continuity is expressed in terms of ε's and δ's, using definition 5.2, but the more important interpretation (for this course) of the definition is that if f is continuous at $(x = a, y = b)$, then the function value $f(x, y)$ will be close to $f(a, b)$ if x and y are both sufficiently close to a and b, respectively.

In math 234 we do not study the techniques that can be used to prove continuity of a function of two variables. While there are many discontinuous functions, most of these involve division by zero (see examples below), or "definition by parts" (see problem 18), or more complicated constructions.

Iterated Limits

Along path 1 you first send $x \to a$, and then $y \to b$, and this corresponds to the iterated limit

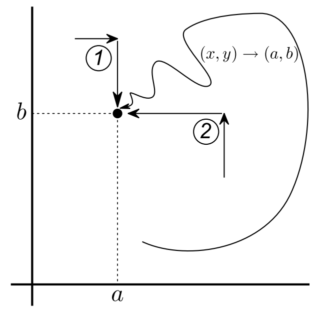
$$\lim_{y\to b}\lim_{x\to a} f(x,y).$$

If you first let $y \to b$ and then let $x \to a$, you get path 2, which corresponds to the other iterated integral.

There are many other paths along which (x, y) can approach (a, b), and the limit

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

equals some number L if f approaches this value no matter which path (x, y) follows as it approaches (a, b).



5.4. Iterated limits. Instead of introducing a brand new definition of "limit" you could try to recycle the old one-variable definition of limit. Thus, in order to find the limit of $f(x, y)$ as (x, y) approaches some point (a, b), you could first forget about y and just let x approach a. This leads to

$$\lim_{x\to a} f(x,y) = L(y).$$

This is a limit of one variable, because we're freezing the y variable for the moment. The result is some quantity which will depend on the value at which we froze y. Next you could let y approach b, and compute

$$\lim_{y\to b} L(y) = \lim_{y\to b}\bigl\{\lim_{x\to a} f(x,y)\bigr\}.$$

The result of this computation would then be our answer to the question "what happens to $f(x, y)$ when (x, y) goes to (a, b)?"

The problem here is that there are at least two versions of this approach, depending on which limit you take first. You could compute

$$\lim_{y\to b}\bigl\{\lim_{x\to a} f(x,y)\bigr\} \text{ and } \lim_{x\to a}\bigl\{\lim_{y\to b} f(x,y)\bigr\}.$$

Do these always give the same result? And do they give the same result as the limit which we defined above in §5.3. The answer to these questions is "yes, most of the time, but not always."

5.5. Theorem on Switching Limits. ***If*** *$\lim_{(x,y)\to(a,b)} f(x,y) = L$ exists,* ***then*** *the two iterated limits exist, and they are the same:*

$$\lim_{x\to a}\lim_{y\to b} f(x,y) = \lim_{y\to b}\lim_{x\to a} f(x,y) = L.$$

Also, if $\lim_{(x,y)\to(a,b)} f(x,y) = L$ exists, and if $x(t)$ and $y(t)$ are any two functions with

$$\lim_{t\to t_0} x(t) = a, \ \textit{and} \ \lim_{t\to t_0} y(t) = b,$$

(so that $(x(t), y(t))$ represents a path which approaches the point (a, b) as $t \to t_0$) then

$$\lim_{t\to t_0} f(x(t), y(t)) = L.$$

5.6. Limit examples. The function $f(x,y) = (x^2 - y^2)/(x^2 + y^2)$ is defined everywhere on the plane, except at the origin. You could try to assign a value to $f(0,0)$ by taking the limit of $f(x,y)$ as x and y go to zero. This is what you find :

Consider the limits

$$A = \lim_{x\to 0}\lim_{y\to 0}\frac{x^2-y^2}{x^2+y^2} \text{ and } B = \lim_{y\to 0}\lim_{x\to 0}\frac{x^2-y^2}{x^2+y^2}.$$

Then you can easily compute that $A = 1$ and $B = -1$. So here is an example where switching the order of limits changes the outcome. The theorem tells us that the limit

$$\lim_{(x,y)\to(0,0)}\frac{x^2-y^2}{x^2+y^2}$$

does not exist.

Note that to make this example we had to divide by zero at $(0,0)$.

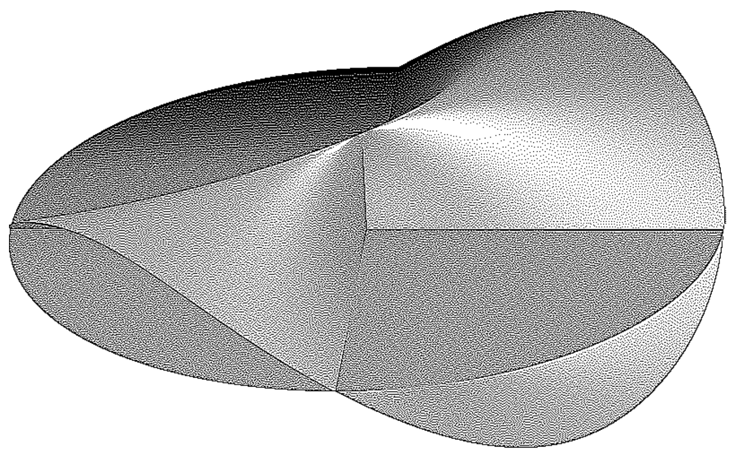


Figure 4: The graph of a function which is discontinuous at the origin. (See Problem 19.)

Here is another example: consider the function

$$g(x,y) = \frac{2xy}{x^2+y^2}.$$

Its domain is the whole plane, except the origin, where we once again would have to divide by zero.

The iterated limits exist for this example. If you try to compute them you will find

$$\lim_{x\to 0}\lim_{y\to 0}\frac{2xy}{x^2+y^2} = 0, \text{ and } \lim_{y\to 0}\lim_{x\to 0}\frac{2xy}{x^2+y^2} = 0.$$

Nevertheless, the limit $\lim_{(x,y)\to(0,0)} g(x,y)$ ***does not exist***. One way to see that is to let (x,y) approach the origin along a straight line, say the line with equation $y = x$. (What happens along other lines is one of the exercises). You get

$$\lim_{x\to 0,\ y=x} g(x,y) = \lim_{x\to 0} g(x,x) = \lim_{x\to 0}\frac{2x\cdot x}{x^2+x^2} = 1.$$

Conclusion: along the x-axis and along the y-axis g remains 0, but along the diagonal the function has the value 1, so that its limit along the diagonal is 1.

6. Problems

16. Find the level sets of the functions f and g from §5.6.

17. Compute the limits of the functions f and g from §5.6 along the lines $y = mx$, where m is a constant. Does the result depend on m?

18. Consider the function

$$f(x, y) = \begin{cases} 1 & \text{if } y \geq |x| \\ 0 & \text{if } y < |x|. \end{cases}$$

(**i**) Draw the graph of f. What is its domain?

(**ii**) Compute the two iterated limits

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

and

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

(**iii**) Compute $\lim_{(x,y)\to(0,0)} f(x, y)$ if it exists.

(**iv**) At which points (a, b) in the plane is the function continuous?

(**v**) Answer the same questions for the function

$$g(x, y) = \begin{cases} 1 & \text{if } |x| \leq y \leq 2|x| \\ 0 & \text{otherwise.} \end{cases}$$

19. (**i**) Figure 4 shows the graph of

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

and the xy-plane (the plane $z = 0$). The axes are missing. Draw the x and y axes in the figure.

(**ii**) It turns out that the graph of

$$g(x, y) = 2xy/(x^2 + y^2)$$

also looks like Figure 4. Assuming that Figure 4 is in fact the graph of g, draw the x and y axes in Figure 4.

20. Let

$$h(x, y) = \frac{x^4 - y^2}{x^4 + y^2}.$$

(**i**) Compute the limit of $h(x, y)$ as (x, y) approaches the origin along the line $y = mx$. Does the result depend on m?

(**ii**) Compute the limit of $h(x, y)$ as (x, y) approaches the origin along the parabola $y = mx^2$. Does the result depend on m?

(**iii**) Does the limit $\lim_{(x,y)\to(0,0)} h(x, y)$ exist?

(**iv**) Answer the same questions for the function

$$k(x, y) = \frac{yx^2}{y^2 + x^4}.$$

21. The following function plays an important role in the theory of heat conduction, the theory of diffusion, and in probability theory. It is called the "heat kernel" or "Gauss kernel".

$$H(x, t) = \frac{1}{\sqrt{t}} e^{-x^2/t}?$$

Does the limit of $H(x, t)$ at $(0, 0)$ exist? Do any of the iterated limits exist? More precisely,

(**i**) Find $\lim_{x\to 0} \lim_{t\searrow 0} H(x, t)$.

(**ii**) Find $\lim_{t\searrow 0} \lim_{x\to 0} H(x, t)$.

(The domain of this function is all points (x, t) with $t > 0$ – why?)

A hint: How do you find the limit $\lim_{s\searrow 0} \frac{1}{\sqrt{s}} e^{-1/s}$? You substitute $s = 1/z$, so when $s \searrow 0$ you have $z \to +\infty$, and $\lim_{s\searrow 0} \frac{1}{\sqrt{s}} e^{-1/s} = \lim_{z\to\infty} \sqrt{z} e^{-z}$. Now use your math 221 limits.

CHAPTER 2

Derivatives

1. Partial Derivatives

The derivative $f'(x)$ of a function of one variable, $y = f(x)$, measures a rate of change: if you increase x by a small amount Δx then $y = f(x)$ also increases by a small amount Δy. The ratio between these two changes is the derivative: $f'(x) \approx \frac{\Delta y}{\Delta x}$.

For a function $z = f(x, y)$ of two variables there is a similar concept: if you change x and/or y by a small amount then z will also change by a small amount, and there are formulas relating the changes Δx, Δy and Δz. Because there are many different ways in which you can change x and y there are a few different formulas. We will encounter the following versions of "the derivative of $f(x, y)$":

– Freeze y and change x, or freeze x and change y: this leads to the so-called ***partial derivatives***.

– Simultaneously vary both x and y: the resulting change turns out to be the sum of the changes you would get if you only varied x or only varied y, respectively. This will follow from the ***chain rule***, and the resulting formula is called the ***total derivative***.

We begin with the partial derivatives.

1.1. Definition of Partial Derivatives. *If $z = f(x, y)$ is a function of two variables which is defined on an open domain G, then at any point (x, y) in that domain the partial derivatives of f with respect to x and with respect to y are*

$$\frac{\partial f}{\partial x}(x,y) = \lim_{\Delta x \to 0} \frac{f(x+\Delta x, y) - f(x,y)}{\Delta x} \tag{4}$$

and

$$\frac{\partial f}{\partial y}(x,y) = \lim_{\Delta y \to 0} \frac{f(x, y+\Delta y) - f(x,y)}{\Delta y} \tag{5}$$

The following more convenient notation is used very often (because it's so much shorter):

$$f_x(x,y) = \frac{\partial f}{\partial x}(x,y), \qquad f_y(x,y) = \frac{\partial f}{\partial y}(x,y). \tag{6}$$

When we are in a hurry we drop the "(x, y)" from our notation for derivatives.

1.2. Examples. Computing partial derivatives not harder than computing ordinary derivatives. To find the partial derivative of a function with respect to x you just pretend all other variables are constants and differentiate. Or, in other words, you could think of the partial derivative of $f(x, y)$ with respect to x as the ordinary derivative of the function f in which you have frozen the variable y at some particular value.

For instance, the partial derivatives of the function $f(x, y, z) = x^2 \sin \pi y$ of ***three*** variables x, y, and z, are

$$f_x = 2x \sin \pi y, \quad f_y = \pi x^2 \cos \pi y \text{ and } f_z = 0.$$

The function we chose doesn't actually depend on z so the derivative with respect to z vanishes.

2. Problems

22. Find the partial derivatives of the following functions:

(i) $f(x, y) = x^2y^3 - x^3y^2$.

(ii) $f(x, y) = \cos(x^2y) + y^3$.

(iii) $f(x, y) = \dfrac{xy}{x^2 + y}$.

(iv) $f(x, t) = (x + t)^4$.

(v) $f(x, t) = (x - t)^4$.

(vi) $f(x, t) = \sin \omega t \cos \dfrac{2\pi x}{L}$.

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

(viii) $f(x, y) = xy \ln(xy)$.

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

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

(xi) $f(u, v) = e^{u+v}$

(xii) $f(x, y) = x \tan(y)$.

(xiii) $f(x, y) = \dfrac{1}{xy}$.

23. Let $f(x, y) =$ the distance from (x, y) to the origin.

Find a formula for f, and compute

$$f_x, \quad f_y, \text{ and } \sqrt{f_x^2 + f_y^2}.$$

24. Suppose $f(t)$ and $g(t)$ are single variable differentiable functions. Find $\partial z/\partial x$ and $\partial z/\partial y$ for each of the following two variable functions.

(i) $z = f(x)g(y)$ **(ii)** $z = f(xy)$
(iii) $z = f(x/y)$

25. Let f be the **distance to the square Q function** from problem 4. Find the partial derivatives f_x and f_y of f. (You will need your answer to problem 4, in particular the description of f as a "piecewise defined function".)

3. The Chain Rule and friends

When you compute the partial derivative of a function with respect to a variable x you pretend all other variables are constants, and just differentiate with respect to x, just as you would in first semester calculus. There is therefore no need to state a product rule or quotient rule, because these are *exactly the same* as for functions of one variable. The chain rule on the other hand is different: there is a chain rule for functions of several variables, but it has more terms than the chain rule from one-variable calculus. There are several related topics which fit together in a discussion of the chain rule, namely ***Linear Approximation***, ***Tangent Planes to a Graph***, and ***The Total Derivative***. We'll go through these one at a time in the section.

Throughout this whole section we will assume that

$$\begin{cases} z = f(x, y) \text{ is a function on some domain whose partial derivatives} \\ \qquad f_x(x, y) \text{ and } f_y(x, y) \text{ are continuous on this domain.} \end{cases} \tag{7}$$

3.1. Linear approximation of a graph. The key to the chain rule is the linear approximation formula. This formula tells us approximately how much a function $z = f(x, y)$ of two variables changes if both variables are subjected to a small change.

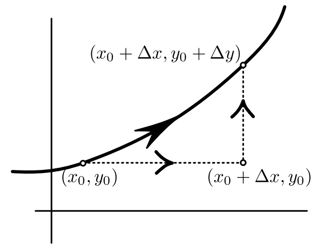


Figure 1: A picture of the calculations in (8)

To arrive at the formula assume that x is increased from x_0 to $x_0 + \Delta x$, and that y is similarly increased from y_0 to $y_0 + \Delta y$. Then the change in $f(x, y)$ is given by

$$
\begin{aligned}
(8)\qquad \Delta f &= f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0, y_0) \\
&= \underbrace{f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0 + \Delta x, y_0)}_{\text{only } y \text{ changes}} + \underbrace{f(x_0 + \Delta x, y_0) - f(x_0, y_0)}_{\text{only } x \text{ changes}} \\
&= f_y(x_0 + \Delta x, \tilde{y})\Delta y + f_x(\tilde{x}, y_0)\Delta x \qquad \text{(use Mean Value Theorem twice)} \\
&= f_x(\tilde{x}, y_0)\Delta x + f_y(x_0 + \Delta x, \tilde{y})\Delta y \qquad \text{(write } x \text{ terms first)}
\end{aligned}
$$

The numbers $\tilde{x}$, $\tilde{y}$ are provided by the Mean Value Theorem, so, $\tilde{x}$ lies between x_0 and $x_0 + \Delta x$, and $\tilde{y}$ lies between y_0 and $y_0 + \Delta y$. The numbers $\tilde{x}$ and $\tilde{y}$ are otherwise unknown, but the assumption (7) that the partial derivatives f_x and f_y are continuous allows us to get rid of $\tilde{x}$ and $\tilde{y}$ if we assume that Δx and Δy are small. So assume that Δx and Δy are indeed "small." Then, since $\tilde{x}$ lies between x_0 and $x_0 + \Delta x$ we will have $f_x(\tilde{x}, y_0 + \Delta y) \approx f_x(x_0, y_0)$ and similarly, we will have $f_y(x_0, \tilde{y}) \approx f_y(x_0, y_0)$. We can make this a bit more precise by saying that there are small numbers e_x and e_y such that

$$f_x(\tilde{x}, y_0) = f_x(x_0, y_0) + e_x, \text{ and } f_y(x_0 + \Delta x, \tilde{y}) = f_y(x_0, y_0) + e_y.$$

Putting this in (8) we get the linear approximation formula:

$$(9)\qquad f(x_0 + \Delta x, y_0 + \Delta y) = \underbrace{f(x_0, y_0) + f_x(x_0, y_0)\Delta x + f_y(x_0, y_0)\Delta y}_{\text{linear approximation}} + \underbrace{e_x\Delta x + e_y\Delta y}_{\text{error}}$$

in which e_x and e_y depend on $\Delta x, \Delta y$, but they satisfy

$$\lim_{(\Delta x, \Delta y)\to(0,0)} e_x = \lim_{(\Delta x, \Delta y)\to(0,0)} e_y = 0.$$

If we ignore the "error term" then we find the following more commonly used form of the linear approximation formula:

$$(10)\qquad f(x_0 + \Delta x, y_0 + \Delta y) \approx f(x_0, y_0) + f_x(x_0, y_0)\Delta x + f_y(x_0, y_0)\Delta y$$

Another way of writing this equation appears if you let Δf stand for the change in f, i.e. $\Delta f = f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0, y_0)$. You then get

$$(11)\qquad \Delta f \approx f_x(x_0, y_0)\Delta x + f_y(x_0, y_0)\Delta y = \frac{\partial f}{\partial x}\Delta x + \frac{\partial f}{\partial y}\Delta y.$$

It is important to realize that this is only an approximate equation, and that according to (9) the error (difference between left and right hand sides) is given by $e_x\Delta x + e_y\Delta y =$ "$o(\Delta x) + o(\Delta y)$"; the error is "small " compared to Δx and Δy. The smaller one chooses

Δx and Δy, the better the approximation. This leads many to say that there is an exact equation when Δx and Δy are "infinitely small," and in this case one writes

$$df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy. \tag{12}$$

The meaning of this equation is that infinitesimally small changes in x and y, of magnitudes dx and dy, respectively, lead to an infinitesimally small change in f of magnitude df, and that df, dx, and dy are related by (12). Even though it is very difficult to make sense of the "infinitely small" quantities dx, dy, df, in (12), this notation is widely used, because the make-belief it entails allows one to ignore the more awkward error terms in (9).

3.2. The tangent plane to a graph. We return to the linear approximation formula (10). With

$$z = f(x, y), \quad x = x_0 + \Delta x, \quad y = y_0 + \Delta y$$

this is the same as

$$z = f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0). \tag{13}$$

This is the equation for a plane which we call the ***tangent plane*** to the graph of f at the point $(x_0, y_0, f(x_0, y_0))$.

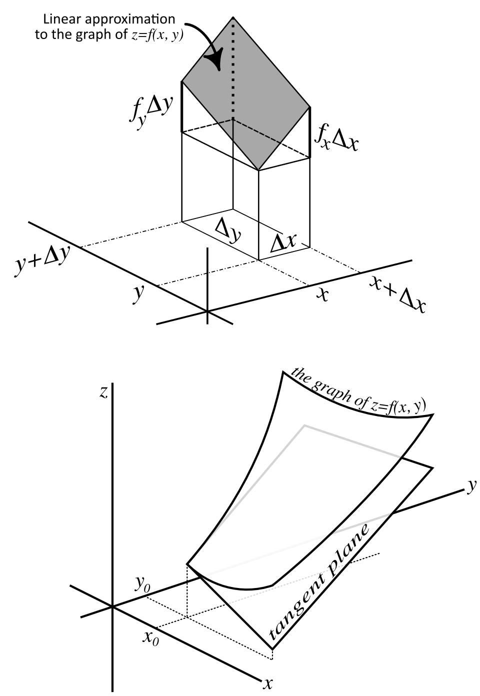


Figure 2: Top: The graph of the linear approximation of f (graph of f itself is not shown – see the bottom figure). If you increase x by Δx, then f will increase by approximately $f_x\Delta x$, and if you increase y by Δy, then f increases by approximately $f_y\Delta y$. If you increase x and y by Δx and Δy at the same time, then f increases by roughly $f_x\Delta x + f_y\Delta y$. The vertical dotted line behind the parallelogram represents this increase in f. **Bottom:** The graph of a function, and of its tangent plane at some point (x_0, y_0, z_0). The tangent plane is the graph of the linear approximation to f.

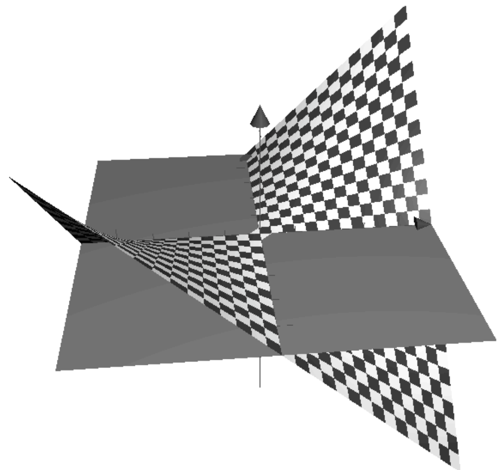


Figure 3: The graph of $z = xy$ and the tangent plane at the origin.

3.3. Example: tangent plane to the sphere. *The point* (x_0, y_0, z_0) *lies on the upper half of the sphere with radius 4 centered at the origin. Find an equation for the tangent plane to the sphere at that point, if* $x_0 = 1$ *and* $y_0 = 3$*.*

Solution: The equation for the sphere is $x^2 + y^2 + z^2 = 4^2 = 16$, so the upper half is the graph of the function $f(x, y) = \sqrt{16 - x^2 - y^2}$. The z coordinate of the given point is therefore $z_0 = \sqrt{16 - 1^2 - 3^2} = \sqrt{6}$. The partial derivatives of f at $(x_0, y_0) = (1, 3)$ are

$$\frac{\partial f}{\partial x} = \frac{-x_0}{\sqrt{16 - x_0^2 - y_0^2}} = -\frac{1}{\sqrt{6}}, \qquad \frac{\partial f}{\partial y} = \frac{-y_0}{\sqrt{16 - x_0^2 - y_0^2}} = -\frac{3}{\sqrt{6}}.$$

The equation for the tangent plane is then

$$z = \sqrt{6} - \frac{1}{\sqrt{6}}(x - 1) - \frac{3}{\sqrt{6}}(y - 3) = \frac{16}{\sqrt{6}} - \frac{x}{\sqrt{6}} - \frac{3y}{\sqrt{6}}.$$

3.4. Example: tangent planes to the saddle surface. *Find the equation for the tangent plane to the saddle surface* $z = xy$ *at the origin.*

Solution: The saddle surface is the graph of the function $f(x, y) = xy$ whose partial derivatives are $f_x(x, y) = y$ and $f_y(x, y) = x$. By Eq. (13) the tangent plane to any point (x_0, y_0, x_0y_0) on the graph is given by

$$z = x_0y_0 + y_0(x - x_0) + x_0(y - y_0). \tag{14}$$

At the origin we have $x_0 = y_0 = 0$, so the tangent plane there is given by

$$z = 0,$$

i.e. it is just the xy-plane.

3.5. Example: another tangent plane to the saddle surface. *Find the equation for the tangent plane to the saddle surface $z = xy$ at the point $(2, 1, 2)$. Where does this plane intersect the coordinate axes?*

Solution: This is almost the same problem as before. The only difference is that we are trying to find the tangent plane at a point other than the origin. To get the tangent plane at the point with $x_0 = 2$, $y_0 = 1$ we substitute and find

$$z = 2 + 1 \cdot (x - 2) + 2 \cdot (y - 1) = -2 + x + 2y.$$

The intersections with the x, y and z axes are, respectively, $(2, 0, 0)$, $(0, 1, 0)$, and $(0, 0, -2)$.

3.6. Follow-up problem – intersection of tangent plane and graph. *Find those points at which the tangent plane to the graph at $(2, 1, 2)$ intersects the saddle surface itself.*

Solution: We have just found that the tangent plane is the graph of $z = -2+x+2y$, while we are given that the saddle surface is the graph of $z = xy$. Any point (x, y, z) lies in the intersection exactly when its coordinates satisfy both equations. Eliminating z we see that (x, y) must satisfy

$$xy = -2 + x + 2y, \text{ or, equivalently, } xy - x - 2y + 2 = 0.$$

This is a quadratic equation, so you would normally expect a circle, ellipse, or hyperbola, but in this case the right hand side can be factored:

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

So we see that (x, y, z) lies on the intersection of the tangent plane if and only if either

$$\boxed{x = 2,\ z = 2y, \text{ and } y \text{ is arbitrary,}} \quad \text{or} \quad \boxed{y = 1,\ z = x, \text{ and } x \text{ is arbitrary.}} \tag{15}$$

You can describe the points we found in vector form, which leads to

$$\vec{x} = \begin{pmatrix} 2 \\ y \\ 2y \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix} + y \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} \text{ and } \vec{x} = \begin{pmatrix} x \\ 1 \\ x \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + x \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}. \tag{16}$$

From this you see that the intersection consists of two straight lines. [1]

3.7. The Chain Rule. Given two functions $x = x(t)$, $y = y(t)$ of one variable, and a function $z = f(x, y)$ of two variables, what is the derivative of the function $g(t) = f(x(t), y(t))$?

If t increases by an amount Δt from t_0 to $t_0 + \Delta t$, then x and y will increase by amounts Δx and Δy,

$$\Delta x = x(t_0 + \Delta t) - x_0, \qquad \Delta y = y(t_0 + \Delta t) - y_0,$$

where $x_0 = x(t_0)$ and $y_0 = y(t_0)$. By the linear approximation formula (8) one then has

$$\frac{\Delta f}{\Delta t} = f_x(x_0, y_0)\frac{\Delta x}{\Delta t} + f_y(x_0, y_0)\frac{\Delta y}{\Delta t} + e_x\frac{\Delta x}{\Delta t} + e_y\frac{\Delta x}{\Delta t}$$

As we let $\Delta t \to 0$ the quotients $\Delta x/\Delta t$ and $\Delta y/\Delta t$ converge to $x'(t_0)$ and $y'(t_0)$, while the errors e_x and e_y converge to zero, so we get

$$\frac{df(x(t), y(t))}{dt} = f_x(x_0, y_0)x'(t_0) + f_y(x_0, y_0)y'(t_0). \tag{17}$$

since $\tilde{x}$ tends to x_0 and $\tilde{y}$ tends to y_0 as $\Delta t \to 0$.

This formula is often also written as

$$\frac{df}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt}. \tag{18}$$

[1] If the last calculation (going from (15) to (16)) is a mystery, then this would be a **very** good time to review vectors and parametric representations of lines from math 222.

This formula becomes easy to remember if you interpret the first term as "the change in f caused by the change in x" and the second term as "the change in f caused by the change in y."

In the way (18) is written a number of details are swept under the rug: the two derivatives $\frac{dx}{dt}$ and $\frac{dy}{dt}$ are ordinary (math 221) derivatives of the two functions $x(t)$ and $y(t)$; the two partial derivatives $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ are the partial derivatives of f ***in which one has substituted*** $x(t)$ ***and*** $y(t)$***.*** A more correct way of writing the equation would be

$$\frac{df(x(t), y(t))}{dt} = \frac{\partial f}{\partial x}(x(t), y(t))x'(t) + \frac{\partial f}{\partial y}(x(t), y(t))y'(t).$$

Many people find (18) easier on the eyes, so that is what we will write.

3.8. The difference between d and ∂. Compare (18) with the linear approximation formula (12) with infinitesimal small quantities. Equation (18) is just (12) in which one has divided both sides by dt. In contrast to equation (12) which contains the strange "infinitely small quantities" dx, dy, df, equation (18) contains the derivatives $\frac{dx}{dt}$, etc. which ***are*** well-defined.

Note that we have a breakdown of Leibniz's notation: If you ignore the distinction between "d" and "∂", and just cancel dx and ∂x, and also dy and ∂y on the right then you end up with

$$\frac{df}{dt} = \frac{\partial f}{dt} + \frac{\partial f}{dt} = 2\frac{\partial f}{dt},$$

which doesn't make a lot of sense. The moral: don't cancel dx against ∂x!

4. Problems

26. Find the linear approximation to $f(x, y)$ at the point (a, b) in the following cases:

(i) $f(x, y) = xy^2$, $(a, b) = (3, 1)$.

(ii) $f(x, y) = x/y^2$, $(a, b) = (3, 1)$.

(iii) $f(x, y) = \sin x + \cos y$, $(a, b) = (\pi, \pi)$.

(iv) $f(x, y) = xy/(x + y)$, $(a, b) = (3, 1)$.

27. Find an equation for the plane tangent to the graph of $f(x, y) = \sin(xy)$ at $(\pi, 1/2, 1)$.

28. Find an equation for the plane tangent to the graph of $f(x, y) = x^2 + y^3$ at $(3, 1, 10)$.

29. Find an equation for the plane tangent to the graph of $f(x, y) = x \ln(xy)$ at $(2, 1/2, 0)$.

30. Find an equation for the tangent plane to the graph of $f(x, y) = x^2 - 2xy$ at the point with $x = 2$, $y = 1$.

Find the intersection of the graph of f and the tangent plane you found. Show that it consists of two lines. (Hint: compare with the example in §3.6).

31. **(i)** Find an equation for the tangent plane to the graph of $f(x, y) = xy$ at the point (a, b, ab). Here a and b are constants which will appear in your answer.

(ii) Show that the intersection of the tangent plane and the graph contains two straight lines.

32. **(i)** Find an equation for the plane tangent to the surface defined by $2x^2 + 3y^2 - z^2 = 4$ at $(1, 1, -1)$. (Hint: first write the surface as a graph $z = f(x, y)$).

(ii) The same question at the point $(1, 1, +1)$.

33. **(i)** Suppose you have computed the two partial derivatives of a function $z = f(x_0, y_0)$, and you found $f_x(x_0, y_0) = A$ and $f_y(x_0, y_0) = B$. Find a normal vector to the tangent plane of the graph of $z = f(x, y)$ at (x_0, y_0, z_0).

(Hint: If you know the equation for a plane, then how do you find a normal vector to this plane? Review math 222 for the answer.)

(ii) Find an equation in vector form for the line normal to $x^2 + 4y^2 = 2z$ at $(2, 1, 4)$. (A line is normal to the graph of a function at some point P, if it passes to through P, and if it is perpendicular to the tangent plane to the graph at P.)

34. Imagine a differentiable function, $f(x, y)$. Make a good drawing of the function f and show how $f_x(a, b)$ and $f_y(a, b)$ are the slopes of two lines which are tangent to the graph at (a, b). Indicate clearly which two lines you mean, and describe how they are defined.

(Can't think of a nice graph? Take something like the bottom drawing in Figure 2.)

35. A bug is crawling on the surface of a hot plate, the temperature of which at the point x units to the right of the lower left corner and y units up from the lower left corner is given by $T(x, y) = 100 - x^2 - 3y^3$.

(i) If the bug is at the point $(2, 1)$, in what direction should it move to cool off the fastest?

(ii) If the bug is at the point $(1, 3)$, in what direction should it move in order to maintain its temperature?

36. Let f be as in problem 29. Use linear approximation to approximate $f(1.98, 0.4)$ by hand. Compare your answer with the actual value of $f(1.98, 0.4)$ (you'll need a calculator).

5. Gradients

5.1. The gradient vector of a function. The right hand side in the chain rule (17) can be written as a dot-product of two vectors, namely

$$\frac{df}{dt} = \begin{pmatrix} f_x(x, y) \\ f_y(x, y) \end{pmatrix} \cdot \begin{pmatrix} x'(t) \\ y'(t) \end{pmatrix} \tag{19}$$

It often turns out to be useful to do this, so the vector containing the derivatives of f has been given a name. It is called the ***gradient of*** f, and it is written as

$$\vec{\boldsymbol{\nabla}} f(x, y) = \begin{pmatrix} f_x(x, y) \\ f_y(x, y) \end{pmatrix} \tag{20}$$

The symbol $\vec{\boldsymbol{\nabla}}$ is pronounced "nabla."

The chain rule, written in vector form, looks like this:

$$\frac{df(\vec{\boldsymbol{x}}(t))}{dt} = \vec{\boldsymbol{\nabla}} f(x(t)) \cdot \vec{\boldsymbol{x}}'(t) \tag{21}$$

The linear approximation formula (10) can be rewritten more compactly using the gradient vector:

$$f(\vec{\boldsymbol{x}}_0 + \Delta\vec{\boldsymbol{x}}) \approx f(\vec{\boldsymbol{x}}_0) + \vec{\boldsymbol{\nabla}} f(\vec{\boldsymbol{x}}_0) \cdot \Delta\vec{\boldsymbol{x}}. \tag{22}$$

5.2. The gradient as the "direction of greatest increase" for a function f. The formula

$$\vec{\boldsymbol{a}} \cdot \vec{\boldsymbol{b}} = \|\vec{\boldsymbol{a}}\|\, \|\vec{\boldsymbol{b}}\| \cos\angle(\vec{\boldsymbol{a}}, \vec{\boldsymbol{b}}) \tag{23}$$

for the dot product leads us to a very useful interpretation of the gradient.

If you are at a point $\vec{\boldsymbol{x}}_0$ (P in figure 4) and you are allowed to make a small step $\Delta\vec{\boldsymbol{x}}$ in any direction you like, but of prescribed length, then which way do you go if you want to increase f as much as possible? And where do you go if, instead, you want to decrease f as much as possible? What if you want to keep f the same?

From (22) we see that the change in f is (approximately) given by

$$\Delta f \stackrel{\text{def}}{=} f(\vec{\boldsymbol{x}} + \Delta\vec{\boldsymbol{x}}) - f(\vec{\boldsymbol{x}}) \stackrel{(22)}{\approx} \vec{\boldsymbol{\nabla}} f \cdot \Delta\vec{\boldsymbol{x}} \stackrel{(23)}{=} \|\vec{\boldsymbol{\nabla}} f\|\, \|\Delta\vec{\boldsymbol{x}}\| \cos\theta$$

where θ is the angle between the gradient $\vec{\boldsymbol{\nabla}} f$ and the vector $\Delta\vec{\boldsymbol{x}}$ which represents the step we take. In this formula the lengths $\vec{\boldsymbol{\nabla}} f$ and $\|\Delta\vec{\boldsymbol{x}}\|$ are fixed, and the angle θ is the only thing we can change. Therefore the largest change in f results if $\cos\theta = +1$, the smallest when $\cos\theta = -1$, and no change will result if $\cos\theta = 0$. So we conclude

- To increase f as much as possible choose $\Delta\vec{\boldsymbol{x}}$ in the direction of the gradient $\vec{\boldsymbol{\nabla}} f$,

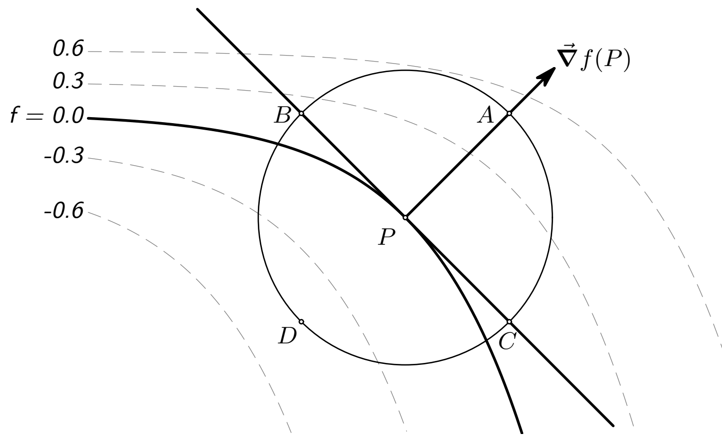


Figure 4: The gradient as direction of fastest increase: if you are at a point P, and you are allowed to jump to any point at a given fixed distance from P, and if you only know $\vec{\nabla} f(P)$, then the linear approximation formula tells you that (i) to maximize f you follow the gradient (choose A); to minimize f you go in the direction opposite to $\vec{\nabla} f(P)$ (choose D); to keep f fixed you move perpendicular to the gradient (choose B or C).

- To decrease f as much as possible choose $\Delta\vec{x}$ in the direction opposite to the gradient $\vec{\nabla} f$, i.e. in the direction of $-\vec{\nabla} f$,
- To keep f constant choose $\Delta\vec{x}$ perpendicular to the gradient.

5.3. The gradient is perpendicular to the level curve. Suppose that some level set of a function $y = f(x, y)$ is a curve, and suppose that we have a parametric representation $\vec{x}(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}$ of this curve. This means that $x(t)$ and $y(t)$ satisfy $f(x(t), y(t)) = C$ for some constant C. By the chain rule we then get

$$0 = \frac{df(\vec{x}(t))}{dt} = \vec{\nabla} f(\vec{x}(t)) \cdot \vec{x}\,'(t),$$

which tells us that the tangent vector $\vec{x}\,'(t)$ to the level set is perpendicular to the gradient $\vec{\nabla} f(\vec{x}(t))$ of the function.

Add: the equation for the tangent line to a level curve of a function $f(x, y) = C$ at a given point $\vec{x}_0 = \left(\begin{smallmatrix} x_0 \\ y_0 \end{smallmatrix}\right)$ is given by

$$\vec{\nabla} f(\vec{x}_0) \cdot (\vec{x} - \vec{x}_0) = 0,$$

or, equivalently,

$$\frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0) = 0.$$

5.4. The chain rule and the gradient of a function of three variables. So far we have only looked at the gradient of a function of two variables. But for a function of three variables there is a very similar definition, and the facts we have discovered have similar counterparts. Let me summarize these definitions and facts, going into as few details as possible.

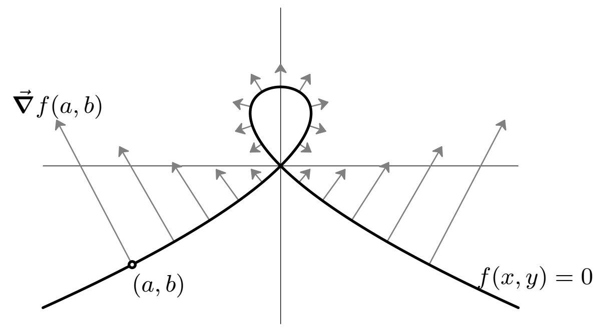


Figure 5: The zero set of the function $f(x,y)=x^2-y^2+y^3$, and its gradient at various points on this zero set.

If $u=f(x,y,z)$ is a function of three variables, then its gradient is defined to be the vector

$$\vec{\boldsymbol\nabla} f(x,y,z)=\begin{pmatrix} f_x(x,y,z)\\ f_y(x,y,z)\\ f_z(x,y,z)\end{pmatrix}, \quad \text{or} \quad \vec{\boldsymbol\nabla} f(\vec{\boldsymbol x})=\begin{pmatrix} f_x(\vec{\boldsymbol x})\\ f_y(\vec{\boldsymbol x})\\ f_z(\vec{\boldsymbol x})\end{pmatrix}.$$

The ***chain rule*** in this context says that, if $x=x(t)$, $y=y(t)$, and $z=z(t)$ are functions of one variable, then the derivative of the function you get by substituting $x(t), y(t), z(t)$ in f is given by any of the following three equivalent formulas

$$\begin{aligned}\frac{df(x(t),y(t),z(t))}{dt} &= f_x(x(t),y(t),z(t))x'(t)+f_y(x(t),y(t),z(t))y'(t)+f_z(x(t),y(t),z(t))z'(t)\\ &= \frac{\partial f}{\partial x}\frac{dx}{dt}+\frac{\partial f}{\partial y}\frac{dy}{dt}+\frac{\partial f}{\partial y}\frac{dy}{dt}\\ &= \vec{\boldsymbol\nabla} f(\vec{\boldsymbol x}(t))\cdot\vec{\boldsymbol x}'(t), \text{ where } \vec{\boldsymbol x}(t)=\begin{pmatrix} x(t)\\ y(t)\\ z(t)\end{pmatrix}.\end{aligned}$$

The linear approximation formula of the function f at some point (x_0,y_0,z_0), which gives you an approximation of the amount by which f increases if you go from (x_0,y_0,z_0) to $(x,y,z)=(x_0+\Delta x, y_0+\Delta y, z_0+\Delta z)$, is as follows:

$$\Delta f = f(x,y,z)-f(x_0,y_0,z_0)\approx \frac{\partial f}{\partial x}\cdot(x-x_0)+\frac{\partial f}{\partial y}\cdot(y-y_0)+\frac{\partial f}{\partial z}\cdot(z-z_0), \tag{24}$$

in which the partial derivatives are to be evaluated at (x_0,y_0,z_0). Compare this with the two variable version (9). In vector form we have

$$\Delta f = f(\vec{\boldsymbol x}_0+\Delta\vec{\boldsymbol x})-f(\vec{\boldsymbol x}_0)\approx \vec{\boldsymbol\nabla} f(\vec{\boldsymbol x}_0)\cdot\Delta\vec{\boldsymbol x}, \text{ where } \vec{\boldsymbol x}_0=\begin{pmatrix} x_0\\ y_0\\ z_0\end{pmatrix}, \ \Delta\vec{\boldsymbol x}=\begin{pmatrix}\Delta x\\ \Delta y\\ \Delta x\end{pmatrix}. \tag{25}$$

This is the same formula as in the two-variable case, where we had (22). The discussion about "direction of steepest increase" applies to the three variable case without change. Thus, if you are at a point $\vec{\boldsymbol x}_0$, and you are allowed to change your position by a small vector $\Delta\vec{\boldsymbol x}$ of a prescribed length, then you choose $\Delta\vec{\boldsymbol x}$ in the direction of the gradient $\vec{\boldsymbol\nabla} f(\vec{\boldsymbol x})$ if you want to increase f as much as possible; you choose $\Delta\vec{\boldsymbol x}$ in the direction of

$-\vec{\nabla} f(\vec{x})$ if you want to *decrease* f as much as possible; and you choose $\Delta\vec{x}$ perpendicular to $\vec{\nabla} f(\vec{x})$ if you want to keep f constant.

5.5. Tangent plane to a level set. If $t = f(x, y, z)$ is a function of three variables then it is hard to visualize its graph, since you would need to draw *four* mutually perpendicular axes, something we, three dimensional creatures, cannot do. However, you can try to visualize the level sets of the function. The level set at level C consists, by definition, of all points in three dimensional space whose coordinates satisfy the equation $f(x, y, z) = C$.

For instance, the unit sphere is given by the equation $x^2 + y^2 + z^2 = 1$, so it is the level set at level 1 of the function $f(x, y, z) = x^2 + y^2 + z^2$. The sphere with radius R is the level set at level R^2.

Consider any function of three variables with continuous partial derivatives, and let (x_0, y_0, z_0) be some point on the level set with level C (thus $f(x_0, y_0, z_0) = C$.) Near this point we can use the linear approximation to f to approximate the equation for the level set of f. We get

$$0 = f(x, y, z) - f(x_0, y_0, z_0) \approx \frac{\partial f}{\partial x} \cdot (x - x_0) + \frac{\partial f}{\partial y} \cdot (y - y_0) + \frac{\partial f}{\partial z} \cdot (z - z_0),$$

where, as in (24), the partial derivatives are to be computed at the given point (x_0, y_0, z_0). They are, in particular, constants (they depend on (x_0, y_0, z_0) but not on (x, y, z).) Thus we see that near any particular point on the level set of a function we can approximate the equation for the level set by

$$\frac{\partial f}{\partial x} \cdot (x - x_0) + \frac{\partial f}{\partial y} \cdot (y - y_0) + \frac{\partial f}{\partial z} \cdot (z - z_0) = 0. \tag{26}$$

If at least one of the partial derivatives at (x_0, y_0, z_0) is non zero, then this is the equation of a plane. We call this plane the tangent plane to the level set.

In vector form the equation for the tangent plane to a level set of f at a point with position vector $\vec{x}_0$ can be written as

$$\vec{\nabla} f(\vec{x}_0) \cdot (\vec{x} - \vec{x}_0) = 0. \tag{27}$$

From this equation you see that, just as in the case (§5.3) of level curves of a function of two variables, ***the gradient $\vec{\nabla} f(\vec{x}_0)$ is perpendicular to the tangent plane of the level set of the function f at the point $\vec{x}_0$.***

5.6. Example. *Find the linear approximation of $F(x, u, v) = e^{-u}(x - v)^2$ and tangent plane to its level set at $x = 1, u = 2, v = 5$*

Solution: At the given values of x, u, v on has $F(1, 2, 5) = e^{-2}(1 - 5)^2 = 16/e^2$. The partial derivatives of F are

$$F_x = 2(x - v)e^{-u}, \quad F_u = -e^{-u}(x - v)^2, \quad F_v = -2(x - v)e^{-u},$$

which at $(x, u, v) = (1, 2, 5)$ reduces to $F_x = -8/e^2$, $F_u = -16/e^2$ and $F_v = +8/e^2$. If (x, u, v) is close to $(1, 2, 5)$, then the linear approximation formula tells us that

$$F(x, u, v) \approx F(1, 2, 5) - \frac{8}{e^2}(x - 1) - \frac{16}{e^2}(u - 2) + \frac{8}{e^2}(v - 5)$$

or, in "Δx" notation,

$$F(1 + \Delta x, 2 + \Delta u, 5 + \Delta v) \approx F(1, 2, 5) - \frac{8}{e^2}\Delta x - \frac{16}{e^2}\Delta u + \frac{8}{e^2}\Delta v.$$

The equation for the tangent plane to the level set of F at the point $(1, 2, 5)$ is therefore

$$-\frac{8}{e^2}(x - 1) - \frac{16}{e^2}(u - 2) + \frac{8}{e^2}(v - 5) = 0,$$

or, after cancelling e^2's and 8's: $(x-1)+2(u-2)-(v-5)=0$. Further simplification shows that the equation for the tangent plane is

$$x+2u-v=0.$$

6. Implicit Functions

In first semester calculus you learned a procedure for finding derivatives of implicitly defined functions. If some function $y=f(x)$ was not given by an explicit formula, but rather by an implicit equation

$$F(x,y)=0 \tag{28}$$

then there was a way to find the derivative of $y=f(x)$ from the above equation only. But there was no formula for $f'(x)$. The reason is that the formula for the derivative $f'(x)$ involves the partial derivatives of F.

In this section we review implicit differentiation again. The following theorem is about the zero set of the function F. One usually thinks of the zero set of a function of two variables as a curve ("an equation defines a curve") but this is not always so. The theorem below gives you a way to find out if the zero set is really a curve, at least near any given point on the zero set which you happen to know.

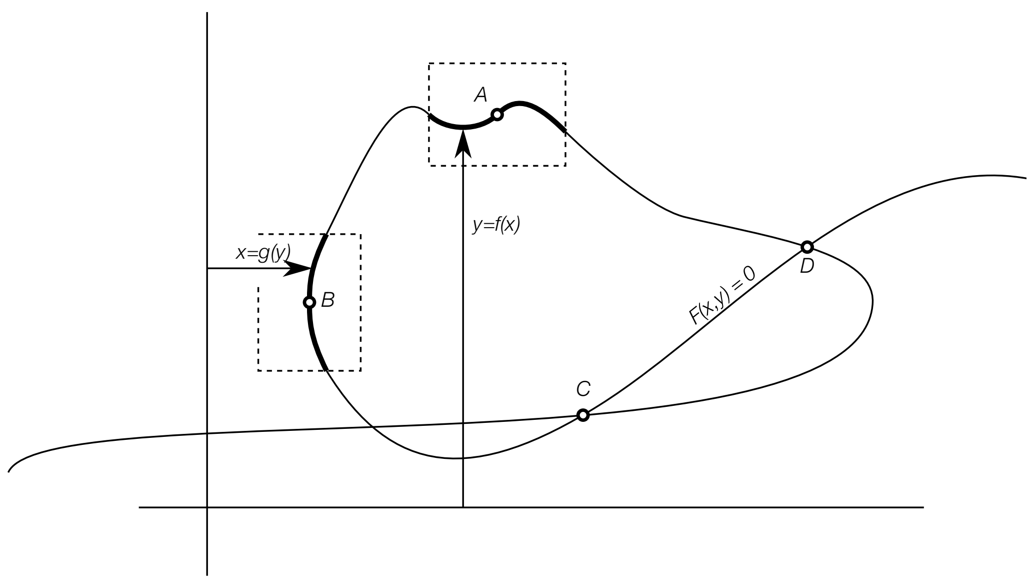


Figure 6: The Implicit Function Theorem. The zero set of a function $F(x,y)$ does not have to be the graph of a function, but if at some point (A) on the zero set you have $F_y \neq 0$, then, near that point A, the zero set is the graph of a function $y=f(x)$. If $F_x \neq 0$ at some point (B), then near B the zero set is also the graph of a function, provided you let x be a function of y: $x=g(y)$. **Exceptional points:** At some points, like C and D in this figure, the level set of F cannot be represented as the graph of a function $y=f(x)$, nor can it be represented as a graph of the type $x=g(y)$. At such points the Implicit Function Theorem implies that both $F_x=0$ and $F_y=0$.

6.1. The Implicit Function Theorem. *Let $F(x, y)$ be a function defined on some plane domain with continuous partial derivatives in that domain, and suppose that a point (x_0, y_0) in the zero set of F is given.*

If $\frac{\partial F}{\partial y}(x_0, y_0) \neq 0$ then there is a small rectangle centered at (x_0, y_0) such that within this rectangle the zero set of F is the graph of a function $y = f(x)$. The derivative of this function is

$$f'(x) = \frac{dy}{dx} = -\frac{F_x(x, f(x))}{F_y(x, f(x))}. \tag{29}$$

If $\frac{\partial F}{\partial x}(x_0, y_0) \neq 0$ then there is a small rectangle centered at (x_0, y_0) such that within this rectangle the zero set of F is the graph of a function $x = g(y)$. The derivative of this function is

$$g'(y) = \frac{dx}{dy} = -\frac{F_y(g(y), y)}{F_x(g(y), y)}. \tag{30}$$

A proof, which will help in understanding the theorem, will be given in class. There is no need to memorize the formulas (29) and (30). You can get them by using the method of implicit differentiation which you learned in math 221. For instance, suppose that the graph of the function $y = f(x)$ gives you a piece of the zero set of F. This means that $F(x, f(x)) = 0$ for all x. Differentiating both sides of this equation leads you, via the chain rule, to

$$0 = \frac{dF(x, f(x))}{dx} = F_x(x, f(x)) + F_y(x, f(x))f'(x). \tag{31}$$

Solve this for $f'(x)$ and you get

$$f'(x) = \frac{dy}{dx} = -\frac{F_x(x, f(x))}{F_y(x, f(x))},$$

which is what the theorem claims.

6.2. The Implicit Function Theorem with more variables. There are many variations and extensions of Theorem 6.1. The simplest is to consider the level set of a function of three rather than two variables. Suppose F is a function of three variables, with continuous partial derivatives, and consider the set of points defined by the equation

$$F(x, y, z) = C.$$

This is the level set of F at level C.

If

$$\frac{\partial F}{\partial y}(x_0, y_0, z_0) \neq 0,$$

then near (x_0, y_0, z_0) the level set of F is the graph of a function $y = g(x, z)$, meaning that the function $y = g(x, z)$ satisfies

$$G(x, g(x, z), z) = 0.$$

Hence you can find the partial derivatives of this function by implicit differentiation. The result is

$$\frac{\partial y}{\partial x} = g_x(x, z) = -\frac{F_x(x, y, z)}{F_y(x, y, z)}, \qquad \frac{\partial y}{\partial z} = g_z(x, z) = -\frac{F_z(x, y, z)}{F_y(x, y, z)}, \tag{32}$$

where $y = g(x, z)$.

6.3. Example – The saddle surface again. The saddle surface is the graph of the function $z = xy$, which we can think of as the zero set of the function

$$F(x, y, z) = z - xy.$$

The point $(2, 3, 6)$ lies on the saddle surface, and at this point the partial derivatives of F are

$$F_x = \frac{\partial(z - xy)}{\partial x} = y = 3, \quad F_y = \frac{\partial(z - xy)}{\partial y} = x = 2, \quad F_z = \frac{\partial(z - xy)}{\partial z} = 1.$$

Since $F_x(2, 3, 6) = y = 3$ is non zero, the Implicit Function Theorem tells us that near this point the zero set of F is the graph of a function $x = g(y, z)$. Solving $F = 0$ for x we see that his function is in fact

$$x = g(y, z) = \frac{z}{y}.$$

The partial derivatives of g are easy to compute in this example, but even if we couldn't find them directly, the Implicit Function Theorem tells us that

$$g_y(3, 6) = -\frac{F_y(2, 3, 6)}{F_x(2, 3, 6)} = -\frac{2}{3}, \qquad g_z(3, 6) = -\frac{F_z(2, 3, 6)}{F_x(2, 3, 6)} = -\frac{1}{3}.$$

7. The Chain Rule with more Independent Variables; Coordinate Transformations

The chain rule we have seen so far tells us how to differentiate expressions of the form $f(x(t), y(t))$. Such expressions are the result of substituting two functions $x(t), y(t)$ of one variable t in one function of two variables $z = f(x, y)$. What do you do if the functions x, y that get substituted in $f(x, y)$ depend on not one but two (or more) variables? The answer is easy: ***you do exactly the same.***

For instance, suppose you want to substitute $x = x(u, v)$ and $y = y(u, v)$ in a function $z = f(x, y)$, resulting in a function $F(u, v) = f(x(u, v), y(u, v))$, and suppose you want find the partial derivatives of F with respect to u. To compute this you *keep v fixed* and regard u as the variable – then $x(u, v)$ and $y(u, v)$ are functions of one variable u and you apply the chain rule you already know. This leads to

$$\frac{\partial F}{\partial u} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial u}$$

The only difference with (18) is that we have written the derivatives of x and y as partial derivatives. We do this to indicate that in computing this derivative we momentarily consider x as a function of u, but later we may want to vary v again.

The same considerations lead to the partial derivative of F with respect to v:

$$\frac{\partial F}{\partial v} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial v} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial v}.$$

7.1. An example without context. Suppose f is some function of two variables and we want to find the partial derivatives of

$$g(u, v, w) = f(2uv, u^2 + w^2).$$

By this we mean that g is the result of substituting $x = 2uv$ and $y = u^2 + w^2$ in f. Note that g is a function of three vairables, and f is a function of two variables.

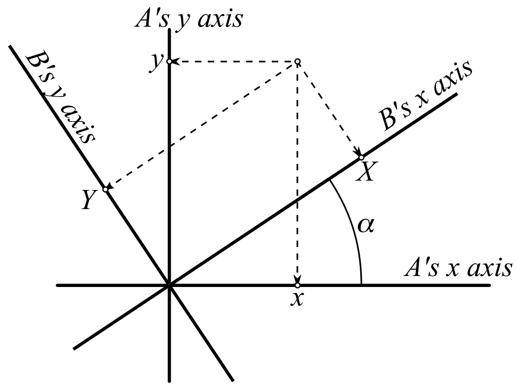


Figure 7: After choosing different x and y axes, A and B will assign different x, y coordinates to the same point in the plane. Equations (33) give the relation between these two sets of coordinates.

The chain rule tells us that the derivatives of g are

$$\begin{aligned}
\frac{\partial g}{\partial u} &= \frac{\partial f}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial u} = 2v\frac{\partial f}{\partial x} + 2u\frac{\partial f}{\partial y} \\
\frac{\partial g}{\partial v} &= \frac{\partial f}{\partial x}\frac{\partial x}{\partial v} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial v} = 2u\frac{\partial f}{\partial x} \\
\frac{\partial g}{\partial w} &= \frac{\partial f}{\partial x}\frac{\partial x}{\partial w} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial w} = 2w\frac{\partial f}{\partial y}
\end{aligned}$$

7.2. Example: a rotated coordinate system. We are used to specifying the location of points in the plane by giving their x and y coordinates. In an abstract mathematical setting there is nothing wrong with this, but in a real-world situation you have to define what you mean by x and y coordinates, and it turns out that different people will choose different but related definitions. For instance, two people A and B could have chosen the same origin, but their axes could be rotated with respect to each other. See Figure 7. If A's coordinates are called x, y and B's coordinates are X, Y then it should be possible to find A's coordinates of a point if you know what coordinates B assigns to this point – given X, Y what are x, y?

One way to derive the equations relating X, Y to x, y is to use complex numbers: the complex number $x + iy$ is obtained from the complex number $X + iY$ by rotating it through an angle α. We know that you can do this by multiplying with $e^{i\alpha}$, so

$$x + iy = e^{i\alpha}(X + iY).$$

Using Euler's formula $e^{i\alpha} = \cos\alpha + i\sin\alpha$ you find

$$\left\{\begin{aligned} x &= X\cos\alpha - Y\sin\alpha, \\ y &= X\sin\alpha + Y\cos\alpha. \end{aligned}\right. \tag{33}$$

Suppose both A and B are measuring the temperature T at various points in the plane. A predicts the temperature at various points in the plane: he says that at the point with coordinates (x, y) the temperature will be $T(x, y)$. In fact he has also found the partial derivatives $\frac{\partial T}{\partial x}$ and $\frac{\partial T}{\partial y}$.

Equipped with the $X, Y \to x, y$ conversion (33) B can now take A's formula for the temperature and express it in terms of her own X, Y coordinates. If we write $T_A(x, y)$

for the temperature at the point whose A-coordinates are (x, y) and $T_B(X, Y)$ for the temperature at the point whose B-coordinates are (X, Y), then we have

$$\begin{aligned} T_B(X,Y) &= T_A(x, y) \\ &= T_A(X\cos\alpha - Y\sin\alpha, X\sin\alpha + Y\cos\alpha). \end{aligned}$$

What is the relation between the partial derivatives of the temperatures as computed by A and by B? The chain rule gives the answer:

$$\begin{aligned} \frac{\partial T_B}{\partial X} &= \frac{\partial}{\partial X}\Big\{T_A(\underbrace{X\cos\alpha - Y\sin\alpha}_{=x}, \underbrace{X\sin\alpha + Y\cos\alpha}_{=y})\Big\} \\ &= \frac{\partial T_A}{\partial x}\cos\alpha + \frac{\partial T_A}{\partial y}\sin\alpha. \end{aligned}$$

7.3. Another example – Polar coordinates. Suppose a quantity P is given in terms of Cartesian coordinates x and y: $P = f(x, y)$. How does P change if you vary the polar coordinates r and θ, i.e. what are the partial derivatives of P with respect to r and θ?

To answer this question we must write P as a function of r and θ. Recall that the relation between Cartesian Coordinates and Polar Coordinates is

$$x = r\cos\theta, \qquad y = r\sin\theta. \tag{34}$$

Therefore $P = f(x, y) = f(r\cos\theta, r\sin\theta)$ and we get

$$\frac{\partial P}{\partial r} = \cos\theta\frac{\partial f}{\partial x} + \sin\theta\frac{\partial f}{\partial y}, \qquad \frac{\partial P}{\partial \theta} = -r\sin\theta\frac{\partial f}{\partial x} + r\cos\theta\frac{\partial f}{\partial y} \tag{35}$$

Since the function f always gives you the value of the quantity P, these relations are usually written in this way:

$$\frac{\partial P}{\partial r} = \cos\theta\frac{\partial P}{\partial x} + \sin\theta\frac{\partial P}{\partial y}, \qquad \frac{\partial P}{\partial \theta} = -r\sin\theta\frac{\partial P}{\partial x} + r\cos\theta\frac{\partial P}{\partial y} \tag{36}$$

Using the relation (34) between polar and Cartesian coordinates you can write these equations in yet another way:

$$\frac{\partial P}{\partial r} = \frac{x}{r}\frac{\partial P}{\partial x} + \frac{y}{r}\frac{\partial P}{\partial y}, \qquad \frac{\partial P}{\partial \theta} = -y\frac{\partial P}{\partial x} + x\frac{\partial P}{\partial y} \tag{37}$$

Problems about the Gradient and Level Curves

37. Compute the gradient of each function in Problem 22

38. Show that for any two differentiable functions f and g one has

$$\vec{\boldsymbol\nabla}(f \pm g) = \vec{\boldsymbol\nabla} f \pm \vec{\boldsymbol\nabla} g, \quad \vec{\boldsymbol\nabla}(fg) = f\vec{\boldsymbol\nabla} g + g\vec{\boldsymbol\nabla} f, \quad \vec{\boldsymbol\nabla}\Big(\frac{f}{g}\Big) = \frac{g\vec{\boldsymbol\nabla} f - f\vec{\boldsymbol\nabla} g}{g^2}.$$

In other words the sum-, product- and quotient rules for differentiation also apply to the gradient.

39. **(i)** Draw the level sets of the function $f(x, y) = x^2 + 4y^2$ at levels 0, 4, 16.

(ii) Find the points on the level set $f(x, y) = 4$ where the gradient is parallel to the vector $\left(\begin{smallmatrix}1\\1\end{smallmatrix}\right)$. What can you say about the tangent line to the level set at those points? Draw the gradient vectors, and the tangent lines at the points you just found.

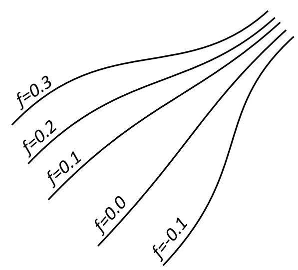
Hint: two non-zero vectors $\vec{\boldsymbol v}$ and $\vec{\boldsymbol w}$ are parallel if there is a number s such that $\vec{\boldsymbol v} = s\vec{\boldsymbol w}$.

(iii) Repeat the same two problems for the function $g(x, y) = 4xy^2$.

40. **(i)** Draw the zero set of the function $f(x, y, z) = x^2 + y^2 - 2z$.

(ii) Find all points on the zero set of the function f where the gradient is parallel to the vector $\vec{\boldsymbol v} = \left(\begin{smallmatrix}1\\1\\2\end{smallmatrix}\right)$.

41. The level sets of a function $z = f(x, y)$ are often curves. Must they always be curves? Could the zero set of a function be a solid square (e.g. all points (x, y) with $0 \le x \le 1$ and $0 \le y \le 1$)?



42. The picture above shows you some level sets of a function. On the bottom left the level sets are further apart, on the top right they are more bunched together. Where is the gradient the larger: bottom left, or top right?

43. Have a look at Figure 5. Assume the function differentiable at the origin.

(**i**) What can you say about the gradient $\vec{\nabla} f$ at the origin?

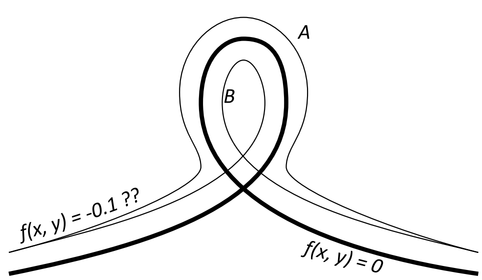
(**ii**) Where is the function positive and where is it negative (assume that the whole zero set is drawn).

44. Consider the unit circle C with equation $x^2 + y^2 = 1$. The unit circle C is a level set of the function $F(x, y) = x^2 + y^2$.

(**i**) Where on C is $F_y \neq 0$? Near which points P on C can one represent C as a graph of the form $y = f(x)$?

(**ii**) Near which points P on C can one represent C as a graph of the form $x = g(y)$?

45. Here is the zero set of a function $z = f(x, y)$ (in bold). The function is only zero on the bold curve, it is nonzero everywhere else.

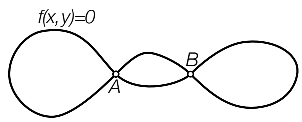


(**i**) One of the two other curves above is the level set $f(x, y) = -0.1$. Which one is it, A or B? As always, explain your answer.

(**ii**) Draw a possible level set $f(x, y) = +0.1$.

(**iii**) Draw possible gradients on the zero set (similar to Figure 5).

46. Here is the zero set of a differentiable function $z = f(x, y)$.



(i) Explain why the Implicit Function Theorem (§6.1) implies that $\vec{\nabla} f = \vec{0}$ at the two points A and B.

(ii) Consider the function $g(x,y) = f(x,y)^2$. Show that f and g have the same zero set.

(iii) Show that $\vec{\nabla} g = 2f\vec{\nabla} f$. (Hint: look at problem 38).

(iv) Show that $\vec{\nabla} g = \vec{0}$ at **all** points on the zero set of g.

47. **(i)** Compute the gradient of the "distance to the square function" f from problems 4 and 25.

(ii) How much is $|\vec{\nabla} f|$?

(iii) Make a drawing of the level sets of f, and the gradient $\vec{\nabla} f$.

48. Let $f(x,y) = \ln(2 + 2x + e^y)$.

(i) Compute the gradient of f at the point (x_0, y_0) with position vector $\vec{x}_0 = \binom{1}{0}$.

(ii) You are allowed to choose a point at a distance 0.01 from the point $(1,0)$. Where would you choose the new point if you want f to be as large as possible? (Hint: review the linear approximation formula and subsequent discussion about the gradient as direction of greatest increase in §5.2)

(iii) Is your answer to the previous the *exact* answer, or only an *approximation*? I.e., could someone else find a point at distance 0.01 from $(1,0)$ at which f has a (slightly) higher value than at the point you found?

(iv) The level set C of f through the point $(1,0)$ happens to be the graph of a function $y = g(x)$. Find that function.

(v) Find a normal vector to the tangent line to C at the point $(1,0)$. Find an equation for the tangent line to C at $(1,0)$.

(vi) How much is $g(1)$? Find two different ways to compute $g'(1)$ based on the work you have done so far.

49. Let (a,b,c) be a point on the sphere with radius R centered at the origin. Find an equation for the tangent plane to the sphere at (a,b,c). Simplify your answer as much as possible (a, b, c, and R will show up in your answer of course.)

About the chain rule and coordinate transformations

50. Use the chain rule to compute dz/dt for $z = \sin(x^2 + y^2)$, $x = t^2 + 3$, $y = t^3$.

51. Use the chain rule to compute dz/dt for $z = x^2 y$, $x = \sin(t)$, $y = t^2 + 1$.

52. Use the chain rule to compute $\partial z/\partial s$ and $\partial z/\partial t$ for $z = x^2 y$, $x = \sin(st)$, $y = t^2 + s^2$.

53. Use the chain rule to compute $\partial z/\partial s$ and $\partial z/\partial t$ for $z = x^2 y^2$, $x = st$, $y = t^2 - s^2$.

54. **(i)** Let $x = x(u,v), y = y(u,v)$ be the following set of functions of u, v:

$$x = u^2 - v^2, \quad y = 2uv.$$

If $g(u,v) = f(x(u,v), y(u,v))$ then compute $g_u(1,0)$, $g_u(1,1)$, $g_v(1,0)$, and $g_v(1,1)$, if you are given these values of the partial derivatives of f:

x	y	$f_x(x,y)$	$f_y(x,y)$
0	0	A	B
1	0	C	D
0	1	E	F
1	1	G	H
2	0	I	J
0	2	K	L

(ii) Repeat the above problem if x and y are given by $x = u$, $y = v/u$.

(iii) Repeat the problem **(i)** if x and y are given by $x = u + v$, $y = u - v$.

55. Let x, y, X, Y, T_A, and T_B be as in the example in §7.2. In that section we computed $\frac{\partial T_B}{\partial X}$.

(i) Compute $\frac{\partial T_B}{\partial Y}$.

(ii) Show that

$$\left(\frac{\partial T_A}{\partial x}\right)^2 + \left(\frac{\partial T_A}{\partial y}\right)^2 = \left(\frac{\partial T_B}{\partial X}\right)^2 + \left(\frac{\partial T_B}{\partial Y}\right)^2.$$

In other words, A and B may measure different partial derivatives, but the temperature gradients they find have the same length. $\|\vec{\boldsymbol{\nabla}} T_A\| = \|\vec{\boldsymbol{\nabla}} T_B\|$.

56. For some function f we are told that at the point with Cartesian coordinates $(2, 1)$ one has

$$\frac{\partial f}{\partial r} = 3, \quad \frac{\partial f}{\partial \theta} = 6.$$

Compute the gradient $\vec{\boldsymbol{\nabla}} f$ at $(2, 1)$.

57. (About polar coordinates). Very often a function is much easier to describe in polar coordinates (r, θ) than in Cartesian coordinates (x, y). If you are given a function in Polar coordinates and you want to know its gradient, then the chain rule gives you the answer.

(i) Show that Polar and Cartesian coordinates are related by

$$r = \sqrt{x^2 + y^2} \text{ and } \theta = \arctan \frac{y}{x},$$

at least in the region where $x > 0$.

(ii) Compute $\frac{\partial r}{\partial x}$, $\frac{\partial r}{\partial y}$, $\frac{\partial \theta}{\partial x}$, $\frac{\partial \theta}{\partial y}$. Try to simplify your answer as much as possible, by reusing the variables r and θ. For instance, the simplest way to write $\frac{\partial r}{\partial x}$ is as $\frac{\partial r}{\partial x} = \frac{x}{r}$.

(iii) Suppose a quantity P is given in terms of Polar coordinates by $P = f(r, \theta)$. Express $\frac{\partial P}{\partial x}$ and $\frac{\partial P}{\partial y}$ in terms of $\frac{\partial f}{\partial r}$ and $\frac{\partial f}{\partial \theta}$.

(iv) Show that

$$\|\vec{\boldsymbol{\nabla}} P\|^2 = \left(\frac{\partial f}{\partial r}\right)^2 + \frac{1}{r^2}\left(\frac{\partial f}{\partial \theta}\right)^2.$$

58. In physics an electric field is described by its **potential function**, $\phi = \phi(x, y)$ (in this problem we assume the world is two-dimensional; the potential ϕ is measured in Volts). Minus the gradient of the potential function is called the **electric field:**

$$\vec{\mathbf{E}} = -\vec{\boldsymbol{\nabla}} \phi.$$

The electric potential of a point charge in the plane is given in Polar coordinates by $\phi = -C \ln r$, for some constant C (the physicists will tell you that C depends on the charge that was placed at the origin; for us it is just some number, and we will in fact assume that $C = 1$.)

(i) Compute the electric field $\vec{\mathbf{E}}$ corresponding to the potential $\phi = -\ln r$.

(ii) Compute $\|\vec{\mathbf{E}}\|$ (this quantity measures the strength of the electric field, but not its direction.) Where is the electric field stronger?

(iii) Make a drawing of the level curves of the potential ϕ, and the electric field $\vec{\mathbf{E}}$.

(iv) In the three dimensional world the electric potential generated by a charged particle at the origin is not given by $-C \ln r$, but instead by the so-called *Coulomb potential*

$$\phi = \frac{C}{r}, \text{ where } r = \sqrt{x^2 + y^2 + z^2}.$$

Compute the corresponding electric field $\vec{\mathbf{E}} = -\vec{\boldsymbol{\nabla}} \phi$.

59. The *ideal gas law*, given by $PV = nRT$, relates the Pressure, Volume, and Temperature of n moles of gas. (R is the ideal gas constant). Thus, we can view pressure, volume, and temperature as variables, each one dependent on the other two.

Each of the following three questions can be answered by applying the chain rule to differentiate $z(t) = f(x(t), y(t))$ for suitable quantities x, y, and z. In each case state which variables play the role of x, y, z, and what the function f is.

(i) If pressure of a gas is increasing at a rate of $0.2Pa$/min and temperature is increasing at a rate of $1K$/min, how fast is the volume changing?

(ii) If the volume of a gas is decreasing at a rate of $0.3L$/min and temperatuere is increasing at a rate of $.5K$/min, how fast is the pressure changing?

(iii) If the pressure of a gas is decreasing at a rate of $0.4Pa$/min and the volume is increasing at a rate of $3L$/min, how fast is the temperature changing?

60. Verify the following identity in the case of the ideal gas law:

$$\frac{\partial P}{\partial V}\frac{\partial V}{\partial T}\frac{\partial T}{\partial P} = -1$$

61. The previous exercise was a special case of the following fact, which you are to verify here:

Assume that $F(x, y, z)$ is a function of 3 variables, and suppose that the relation $F(x, y, z) = 0$ defines each of the variables in terms of the other two, namely $x = f(y, z)$, $y = g(x, z)$ and $z = h(x, y)$, then

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

Hint: this is a problem about implicit differentiation.

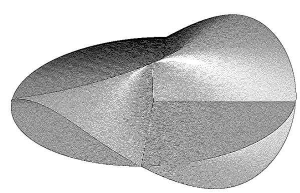
62. Four cartographers are using different coordinates to describe the same landscape. Each of them describes the landscape by specifying a the height of a point in the landscape as a function of its position above a horizontal plane.

Cartographer A uses Cartesian coordinates (x, y) in the plane, B uses Cartesian coordinates (X, Y) in the plane. The coordinates (X, Y) are rotated by $45°$ with respect to (x, y) (see §7.2).

Cartographer C works with A but uses polar coordinates (r, θ) (r is the distance to the origin, θ is the angle with A's x-axis).

Cartographer D works with B and uses polar coordinates (r, φ) (r is the distance to the origin, φ is the angle with B's X-axis).

Here is a (familiar) picture of the landscape that A, B, C, and D are looking at:



(i) If B has found that the height is given by the function $f(X, Y) = 2XY/(X^2 + Y^2)$, then what function does A find for the height?

(ii) What height function does C find?

(iii) What height function does D find?

63. Brian and Ally are using different Cartesian coordinate systems in the plane: (x, y) for Ally, (X, Y) for Brian. They have the same origin, but Brian's coordinates are rotated by an angle of $\theta = \arctan\frac{4}{3}$ ($\approx 53°$, but that's only an approximation. You can give exact answers in this problem, and you don't need a calculator.)

(i) What is the relation between (x, y) and (X, Y)?

(ii) If Ally has found that $T_A(x, y) = 32 + 0.1y$, then what formula $T_B(X, Y)$ will Brian use to describe the temperature?

(iii) On a different occasion Ally found that the temperature had changed. Now Ally measures the temperature and finds that at the point with $x = 1, y = 1$ one has $T_A(1, 1) = 35$, and also $\frac{\partial T_A}{\partial x} = 0.05$

and $\frac{\partial T_A}{\partial y} = 0.8$. Which coordinates does Brian assign to this point, which temperature T_B, and which derivatives $\frac{\partial T_B}{\partial X}$ and $\frac{\partial T_B}{\partial Y}$ does Brian compute at this point?

[Hint: before you compute anything, find $\sin\theta$ and $\cos\theta$; also draw a right triangle one of whose acute angles is θ.]

8. Higher Partials and Clairaut's Theorem

8.1. Higher partial derivatives. By definition

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial\left(\frac{\partial f}{\partial x}\right)}{\partial x}, \qquad \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial\left(\frac{\partial f}{\partial y}\right)}{\partial x}, \qquad \frac{\partial^2 f}{\partial y \partial x} = \frac{\partial\left(\frac{\partial f}{\partial x}\right)}{\partial y}, \qquad \frac{\partial^2 f}{\partial y^2} = \frac{\partial\left(\frac{\partial f}{\partial y}\right)}{\partial y} \tag{38}$$

In subscript notation one writes these higher partial derivatives as follows:

$$f_{xx}(x,y) = \frac{\partial^2 f}{\partial x^2}, \qquad f_{xy}(x,y) = \frac{\partial^2 f}{\partial y \partial x}, \qquad f_{yx}(x,y) = \frac{\partial^2 f}{\partial x \partial y}, \qquad f_{yy}(x,y) = \frac{\partial^2 f}{\partial y^2}.$$

Note the reversal in x/y order in the mixed partial derivatives!

8.2. Example. If $f(x,y) = x^2y + \cos xy$ then $f_x = 2xy - y\sin xy$, and hence

$$\begin{aligned} f_{xx} &= \frac{\partial(2xy - y\sin xy)}{\partial x} = 2y - y^2\cos xy, \\ f_{xy} &= \frac{\partial(2xy - y\sin xy)}{\partial y} = 2x - \sin xy - xy\cos xy. \end{aligned}$$

The other partial derivatives follow from $f_y = x^2 - x\sin xy$, and they are

$$f_{yx} = 2x - \sin xy - xy\cos xy, \quad f_{yy} = -x^2\cos xy.$$

Every time you take a derivative, you can choose whether you differentiate with respect to x or y. Differentiating once you have two possibilities, differentiating twice you have $2\times 2 = 4$ possibilities, etc. That's why we found four partial derivatives of second order in the above example. But if you look carefully, you also see that f_{xy} and f_{yx} are the same. This is no coincidence.

8.3. Clairaut's Theorem – mixed partials are equal. *If for a given function f of two variables the mixed partial derivative $f_{xy}(x,y)$ exists for all (x,y) in a neighborhood of a point (a,b), and if this derivative is continuous at (a,b), then the other mixed partial derivative $f_{yx}(a,b)$ also exists, and $f_{xy}(a,b) = f_{yx}(a,b)$.*

So we normally don't have to worry about the order in which we take partial derivatives.

8.4. Proof of Clairaut's theorem. With some algebra you can show that the definition of partial derivatives imply

$$\frac{\partial^2 f}{\partial x \partial y} = \lim_{\Delta x\to 0}\lim_{\Delta y\to 0} \frac{f(x+\Delta x, y+\Delta y) - f(x, y+\Delta y) - f(x+\Delta x, y) + f(x,y)}{\Delta x \Delta y} \tag{39}$$

while

$$\frac{\partial^2 f}{\partial y \partial x} = \lim_{\Delta y\to 0}\lim_{\Delta x\to 0} \frac{f(x+\Delta x, y+\Delta y) - f(x, y+\Delta y) - f(x+\Delta x, y) + f(x,y)}{\Delta x \Delta y} \tag{40}$$

So it's a matter of showing that one can switch the two limits. We won't go into the details here, but the hypothesis that f_{xy} is continuous implies that you are indeed allowed to switch the limits.

8.5. Finding a function from its derivatives. We now look at integrating the partial derivatives of a function, which looks out of place here (this being a chapter on derivatives and not on integrals), but Clairaut's Theorem actually turns out to play a role.

If you have the derivative $f'(x)$ of some function of one variable then you know how to recover the function $f(x)$: you integrate, i.e.

$$f(x) = \int f'(x)dx + C.$$

Furthermore, any (continuous) function can be the derivative of a function, because, if someone gives you a continuous function $f(x)$, then

$$F(x) \stackrel{\text{def}}{=} \int_a^x f(t)dt$$

is a differentiable function whose derivative is $F'(x) = f(x)$.

What about functions of more than one variable? Suppose you know the partial derivatives

$$\frac{\partial f}{\partial x} = P(x,y) \text{ and } \frac{\partial f}{\partial y} = Q(x,y) \tag{41}$$

of a function of two variables, can you then find the function $f(x,y)$?

The answer is "yes you can find f by integrating, but not every pair of functions P and Q can be $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$."

The following two examples are typical of what can happen.

8.6. Example. Does there exist a function $f(x,y)$ of two variables such that

$$\frac{\partial f}{\partial x} = x^3 - 2xy, \text{ and } \frac{\partial f}{\partial y} = 3y^2$$

both hold? The answer is no, such a function cannot exist, and here is the reason: if there were such a function, then we could compute

$$\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial(x^3 - 2xy)}{\partial y} = -2x, \text{ and } \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial(3y^2)}{\partial x} = 0.$$

By Clairaut's Theorem both computations should give us the same answer, but they don't. Therefore the function f whose partials are as above can't exist.

8.7. Example. Does there exist a function $f(x,y)$ of two variables whose derivatives are

$$\frac{\partial f}{\partial x} = x^3 - 2xy, \text{ and } \frac{\partial f}{\partial y} = \sin \pi y - x^2?$$

Let's check Clairaut's condition:

$$\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial(x^3 - 2xy)}{\partial y} = -2x, \text{ and } \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial(\sin \pi y - x^2)}{\partial x} = -2x.$$

This time both computations gave us the same answer, and the specified partials derivatives are well-defined and continuous for all (x,y), so there is a function f with these partial derivatives. To compute it we first integrate f_x while treating y as a constant:

$$f(x,y) = \int \{x^3 - 2xy\}\, dx = \tfrac{1}{4}x^4 - x^2 y + C(y).$$

The "constant" is only a constant in that it does not depend on x. It may depend on y, and that is why we wrote it as $C(y)$. To find $C(y)$ we differentiate this result with respect to y:

$$\sin \pi y - x^2 = f_y = \frac{\partial(\frac{1}{4}x^4 - x^2 y + C(y)i)}{\partial y} = -x^2 + C'(y).$$

So we see that $C'(y) = \sin \pi y$, and hence $C(y) = -\frac{1}{\pi}\cos \pi y + K$, where K is a real constant (K depends neither on x nor on y).

We find that the following function has the prescribed partial derivatives

$$f(x,y) = \tfrac{1}{4}x^4 - x^2 y - \tfrac{1}{\pi}\cos \pi y + K$$

where K can be any constant.

The method used in this example always works, and we summarize this fact in the following theorem.

8.8. Theorem. *Suppose $P(x,y)$ and $Q(x,y)$ are two functions which are defined on a rectangular domain $\mathcal{R} = \{(x,y) : a < x < b, c < y < d\}$, and suppose that they have continuous partial derivatives on this domain.*

If a function $f(x,y)$ exists such that (41) holds on $\mathcal{R}$, then

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x} \tag{42}$$

must hold on $\mathcal{R}$.

Conversely, if P and Q satisfy (42) then there is a function f defined on $\mathcal{R}$ which satisfies (41).

To prove this theorem we need to understand integrals of functions of several variables, and Green's theorem in particular, so this will have to wait until the end of the semester.

It should be noted that the assumption above that the functions P and Q be defined on a rectangle is important: the theorem is no longer true if the domain of P and Q "has holes." See problem 79.

9. Problems

64. Find all first and second partial derivatives of $x^3y^2 + y^5$.

65. Find all first and second partial derivatives of $4x^3 + xy^2 + 10$.

66. Find all first and second partial derivatives of $x \sin y$.

67. Find all first and second partial derivatives of $\sin(3x)\cos(2y)$.

68. Find all first and second partial derivatives of e^{x+y^2}.

69. Find all first and second partial derivatives of $\ln\sqrt{x^3 + y^4}$.

70. Find all first and second partial derivatives of z with respect to x and y if $x^2 + 4y^2 + 16z^2 - 64 = 0$. (Hint: solve for z or use implicit differentiation...)

71. Find all first and second partial derivatives of z with respect to x and y if $xy + yz + xz = 1$. (Hint: solve for z or use implicit differentiation...)

72. How many different second partial derivatives does a function of two variables have? What about a function of three variables? How many derivatives of third degree does a function of two variables have?

73. Derive the formulas (39) and (40) from the definition of partial derivatives (4) and (5).

74. The equation which describes the *vibrating string* (as in a guitar, piano, or violin string) is

$$\frac{\partial^2 f}{\partial t^2} = c^2 \frac{\partial^2 f}{\partial x^2} \tag{43}$$

where $c > 0$ is some constant. The equation is called the **wave equation**. It is an example of a **partial differential equation.**

Warning : this problem looks like a problem about differential equations, but to answer the following questions you really only have to compute partial derivatives of certain functions, and solve some (easy) algebraic equations.

(i) For which values of the constant v is a "traveling wave with velocity v and profile $F(x)$" a solution of the wave equation (43)? Does it matter which profile F is used here?

(For the terminology used here, revisit problem 11.)

(ii) Suppose the string is clamped down at its ends, and that its length is L. For which values of the constants A and α is

$$f(x,t) = A\sin(\alpha t)\sin\frac{\pi x}{L}$$

a solution of the wave equation? (Assume $A \neq 0$).

(iii) Same question for

$$g(x,t) = B\sin(\beta t)\sin\frac{2\pi x}{L}.$$

(iv) Show that $h(x,t) = f(x,t) + g(x,t)$ is again a solution of the wave equation, where f and g are as above. (Don't use the formulas for f and g: it is easier to prove a more general fact, namely, if two functions f and g satisfy (43), the so does their sum $f+g$.)

(v) Use a graphing application (`grapher.app` on Mac OS X, `graphcalc.exe` on Windows XP/Vista or Linux) to visualize the solutions f, g, and h above. (Don't have a computer? You should be able to describe f and g in words and drawings of some "stills"; h is more challenging.)

75. Suppose $P(x,y) = x^2 - 2xy^3$ and $Q(x,y) = (xy)^2$. Does there exist a function $f(x,y)$ such that $P = f_x$ and $Q = f_y$?

76. Suppose $P(x,y) = x^2 + axy^3$ and $Q(x,y) = (xy)^2$, where a is a constant. For which a does there exist a function $f(x,y)$ such that $P = f_x$ and $Q = f_y$?

77. Suppose $P(x,y) = x^2 - 2xy^3$ and $Q(x,y) = (xy)^2$. Does there exist a function $f(x,y)$ such that $P = f_x$ and $Q = f_y$?

78. Suppose $x = u+v$, $y = u-v$, and suppose $f(x,y) = g(u,v)$. Then compute

(i) $\dfrac{\partial^2 g}{\partial u^2}$ **(ii)** $\dfrac{\partial^2 g}{\partial v^2}$ **(iii)** $\dfrac{\partial^2 g}{\partial u^2} - \dfrac{\partial^2 g}{\partial v^2}$ **(iv)** $\dfrac{\partial^2 g}{\partial u^2} + \dfrac{\partial^2 g}{\partial v^2}$

79. [For discussion] Let

$$P(x,y) = \frac{-y}{x^2+y^2}, \quad Q(x,y) = \frac{x}{x^2+y^2}.$$

(i) What is the domain of P and Q?

(ii) Show that

$$P = \frac{\partial\theta}{\partial x}, \qquad Q = \frac{\partial\theta}{\partial y}$$

where θ is the angle variable from polar coordinates.

(iii) What is the domain of θ? (Careful, we want θ to be a differentiable function on the domain you specify.)

(iv) Show that P and Q satisfy the condition (42). (You don't have to compute the derivatives to check this, although you could.)

(v) Is there a function f such that (41) holds?

CHAPTER 3

Maxima and Minima

In first semester calculus you learned how to find the maximal and minimal values of a function $y = f(x)$ of one variable. The basic method was as follows: assuming the independent variable was restricted to some interval $a \leq x \leq b$, you first look for interior maxima/minima. These occur at ***critical*** or ***stationary*** points of the function, i.e. solutions x of $f'(x) = 0$. You then check the function values at the endpoints a and b of the interval, to see if they might be maxima or minima.

To see which solutions of $f'(x) = 0$ are actually local maxima or minima you can look at the sign of the derivative $f'(x)$ to see where the function is increasing or decreasing, or you can apply the ***second derivative test.***

This chapter we will see how to solve similar questions about functions of two or more variables.

1. Local and Global extrema

Let $z = f(x, y)$ be the function whose maximal or minimal values we are looking for, and let D be the domain of this function. This domain could be the largest possible domain for the given function (in case f is defined by a formula), but it could also be some smaller region which we ourselves have chosen. The question we are considering is

What are the largest and smallest values that $f(x, y)$ can have
if the point (x, y) belong to the domain D?

1.1. Definition of global extrema. *The function f has a* ***global maximum*** *or* ***absolute maximum*** *at a point (a, b) in D if $f(x, y) \leq f(a, b)$ for all points (x, y) in D.*

Similarly, *the function f has a* ***global minimum*** *or* ***absolute minimum*** *at a point (a, b) in D if $f(x, y) \geq f(a, b)$ for all points (x, y) in D.*

1.2. Definition of local extrema. *The function f has a* ***local maximum*** *at a point (a, b) in D if there is a $r > 0$ such that $f(x, y) \leq f(a, b)$ for all points (x, y) in D which also lie in a disc of radius r centered at (a, b).*

Local minima are defined analogously.

1.3. Interior extrema. Recall that a point (a, b) in a domain D is called ***interior*** if it is not a boundary point, or, more precisely, if there is some small $r > 0$ such that the disc with radius r centered at (a, b) is entirely contained in D. We will apply this distinction to the local and global maxima and minima which we find: an ***interior local minimum*** is a local minimum which occurs at an interior point of the domain D of the function.

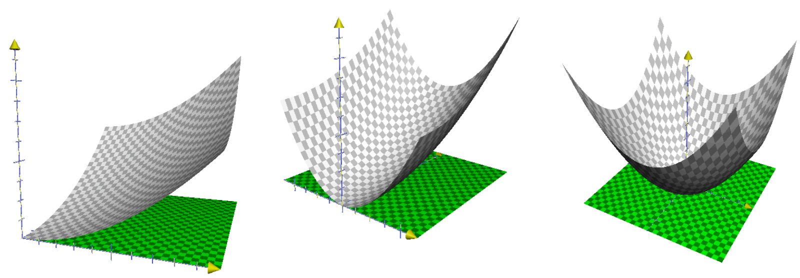


Figure 1: The graph of $f(x, y) = x^2 + y^2$ from example § 2.2 on three different rectangles. From left to right: (i) $0 \le x \le 1, 0 \le y \le 1$. Both max and min are attained at a corner point of the rectangle. (ii) $0 \le x \le 1, -1 \le y \le 1$, Two maxima, both are attained at corner points of the rectangle; the minimum is attained at an edge point. (iii) $-1 \le x \le 1, -1 \le y \le 1$, Four maxima, all attained at corner points of the rectangle; the minimum is attained at an interior point.

2. Continuous functions on closed and bounded sets

Before we go into the details of how you can actually find the maxima and minima, it is good to know the following general fact. It tells us where to expect maxima and minima.

Let $z = f(x_1, \cdots, x_n)$ be a continuous function defined on some ***closed*** and ***bounded*** region D in $\mathbb{R}^n$. Remember: "closed" means that D contains all its boundary points, and "bounded" means that all points in D are not further away from the origin than some fixed radius R (D does not "stretch all the way to infinity".)

We will also assume that f is continuous on D.

2.1. Theorem about Maxima and Minima of Continuous Functions. *A continuous function defined on a closed and bounded region $D \subset \mathbb{R}^n$ has both a maximum and minimum within that region.*

This is proved in courses like 522 (2nd semester analysis) or 561 (point set topology). The proof really doesn't belong here in math 234.

2.2. Example – The function $f(x, y) = x^2 + y^2$. This function is continuous, and the square $Q = \{(x, y) : 0 \le x \le 1, 0 \le y \le 1\}$ is bounded, and it contains all boundary points. Therefore Theorem 2.1 tells us that f attains both its highest and lowest values somewhere in the square. The theorem does not say where these max/min points are, but in this example they are easy to find. The function $f(x, y) = x^2 + y^2$ is at its smallest when both $x = 0$ and $y = 0$, i.e. at the bottom left corner of the square. And $f(x, y)$ is at its largest when $x = 1$ and $y = 1$ both hold. This happens at the top right corner of the square.

Note that the boundary of the rectangle Q has two different kinds of points: it has four corner points, and then all the other points which lie on the edges.

If you change the rectangle Q then the minimum can appear at a corner point, a point on an edge, or in an interior point. See Figure 1.

2.3. A fishy example. Consider the function $f(x, y) = x^2 - x^3 - y^2$. Its zero set is the curve $y^2 = x^2 - x^3$, which is shaped like the letter α, or like a fish – see Figure 2. The function is positive on the tail (D_1) and also on the body (D_2) of the fish, it vanishes on the curve which traces out the fish, and f is negative elsewhere.

Theorem 2.1 does not apply to the region D_1 because D_1 is not bounded (it contains the whole negative x-axis). But the region D_2 is bounded, and our function f is continuous, so Theorem 2.1 does apply to D_2. The theorem tells us that the function f has a maximal value and a minimal value in D_2. In the interior of D_2 the function is strictly positive, and on the boundary of D_2 we have $f = 0$. Therefore each boundary point is a minimum point of f on D_2. The point(s) in D_2 where f attains its highest value must be somewhere in the interior of D_2. In the next section we will see how to find it (and how to check that in this case there really only is ***one*** such point.)

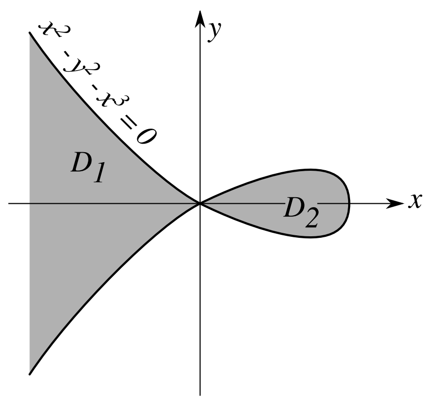


Figure 2: The region where $f(x, y) = x^2 - x^3 - y^2$ is positive consists of two parts, one bounded (D_2), and the other unbounded (D_1). Theorem 2.1 does not apply to the unbounded region, but it does apply to the bounded region D_2. In that region f must attain a maximum and a minimum. Since $f = 0$ on the boundary of the region D_2, and $f > 0$ in the interior, f achieves its lowest value in D_2 everywhere on the boundary of D_2 and its highest value somewhere in the interior. Theorem 2.1 does not tell you how to find that interior point, and allows for the possibility that there might be more interior maxima, as well as a few interior local minima.

3. Problems

80. Suppose you want to find the **maximal value** of $f(x, y) = x^2 - x^3 - y^2$ over all possible (x, y) with $x \geq 0$ (and no restriction on y – this region is called the *right half plane*).

(i) Explain why you should always choose $y = 0$ in order to maximize this particular function $f(x, y)$.

(ii) Use your answer to part **(i)** to find the point (x, y) which maximizes $f(x, y)$ over the right half plane.

(iii) Does our function $f(x, y)$ have a maximal value if (x, y) can be any point in the plane? (hint: what is $f(-1000, 0)$?)

81. Draw the region $R = \{(x, y) : y^2 \leq 4(x^3 - x^4)\}$. Find the largest and smallest values which the function $f(x, y) = x$ can have on this region. (Hint: where is $4(x^3 - x^4) = 4x^3(1 - x)$ positive? The region looks like an Onion).

4. Critical points

For functions $y = f(x), a \le x \le b$, of one variable the standard way of finding minima (and maxima) is to look for them in two different places: either the minimum is attained at one of the end points $x = a$ or $x = b$ of the interval, or else the minimum is attained at an interior point. At an interior minimum one has $f'(x) = 0$, so they can be found by solving the equation $f'(x) = 0$. The same approach works for functions of two or more variables. The basic fact that tells us this is so, is the following theorem.

4.1. Theorem. Local extrema are critical points. *If a function $z = f(x, y)$ defined on a domain D has a local minimum or local maximum at an interior point (a, b) then one has*

$$\frac{\partial f}{\partial x}(a, b) = 0, \text{ and } \frac{\partial f}{\partial y}(a, b) = 0.$$

Picture proof. If f has a local maximum at an interior point (a, b) then $f(x, y) \le f(a, b)$ for all (x, y) close to (a, b). This means that a small piece of the graph of f near its local maximum at $(a, b, f(a, b))$ lies below the plane $z = f(a, b)$. This plane must therefore be the tangent plane to the graph of f. Being horizontal, its slopes are zero, and these slopes are exactly the partial derivatives of f at (a, b). □

Frozen variable proof. Suppose f has a local maximum at an interior point (a, b) of the domain D. Then we can freeze the y-variable at the value $y = b$ and consider the function of one variable $g(x) = f(x, b)$. This function has a maximum at $x = a$, so by first semester calculus we know that $g'(x) = 0$. By definition $g'(a) = f_x(x, b)$, so we conclude that $f_x(a, b) = 0$.

By freezing x instead of y you show that $f_y(a, b) = 0$ also must hold.

The same arguments apply to a local minimum. □

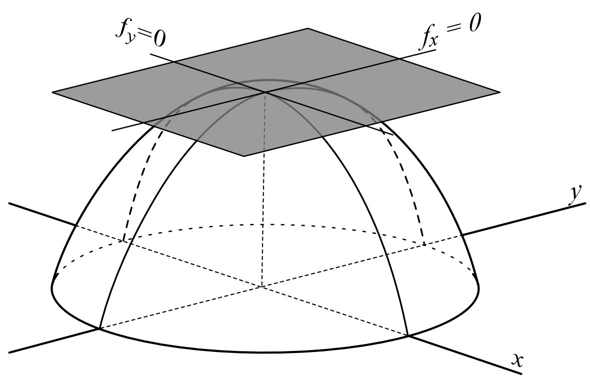


Figure 3: Theorem 4.1. At a local maximum the tangent plane to the graph is horizontal. The partial derivatives w.r.t. both x and y vanish, and in fact, the derivative along **any** path through (a, b) vanishes. To see a picture of a local minimum turn the page upside down.

4.2. Three typical critical points. Let's find the critical points of the following three functions:

$$f(x,y) = x^2 + y^2, \quad g(x,y) = x^2 - y^2, \quad h(x,y) = -x^2 - y^2.$$

► $f(x,y) = x^2 + y^2$. Computing the partial derivatives we find for the first function

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

If (x,y) is a critical point of f then x and y must satisfy the equations $f_x(x,y) = 0$ and $f_y(x,y) = 0$, in this case, $2x = 0$ and $2y = 0$. So we see that f has exactly one critical point, namely the origin $(x,y) = (0,0)$.

Is this critical point perhaps a minimum or a maximum? Since squares can never be negative, $f(x,y) = x^2 + y^2$ is always non-negative, and it is at its smallest when both terms x^2 and y^2 vanish, i.e. $f(x,y)$ has a global minimum at the origin.

► $h(x,y) = -x^2 - y^2$. This function is just $-f(x,y)$, and without looking at its derivatives you can tell that it has a global maximum at the origin. If you differentiate it you find

$$\frac{\partial h}{\partial x} = -2x, \quad \frac{\partial h}{\partial y} = -2y$$

so that the origin is the only critical point of this function too.

► $g(x,y) = x^2 - y^2$. The derivatives of g are

$$\frac{\partial g}{\partial x} = 2x, \quad \frac{\partial g}{\partial y} = -2y,$$

so, once again, the origin is the only critical point. But, unlike the previous two functions, g has neither a maximum nor a minimum at the origin. You can see this by first looking at what g does on the x-axis, and then what g does on the y-axis:

On the x-axis you have $g(x,0) = +x^2$, so g has a ***minimum*** at the origin.

On the y-axis you have $g(0,y) = -y^2$, so g has a ***maximum*** at the origin.

So arbitrarily close to the origin you can find points (x,y) where $g(x,y)$ is larger than $g(0,0)$, and you can find other points where $g(x,y)$ is smaller than $g(0,0)$. Therefore g does not have a local maximum or a local minimum at the origin. See Figure 4.

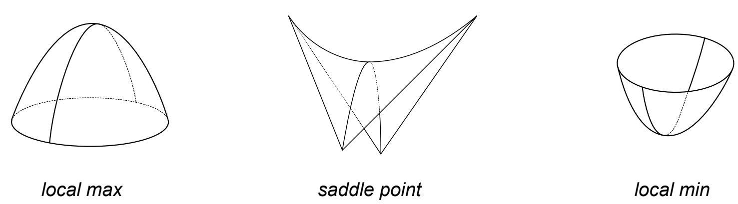
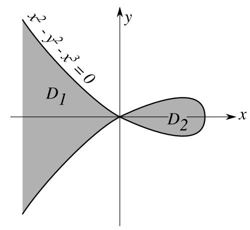


Figure 4: The three most common kinds of critical point. See the examples in §4.2 and also the second derivative test in §9.

4.3. Critical points in the fishy example. *What are the critical points of the function $f(x,y) = x^2 - x^3 - y^2$ from §2.3?* We compute the partial derivatives of the function

$$\frac{\partial f}{\partial x} = 2x - 3x^2 = (2-3x)x, \qquad \frac{\partial f}{\partial y} = -2y.$$

The equation $f_y = 0$ implies that $y = 0$, while $f_x = 0$ implies $x = 0$ or $x = \frac{2}{3}$. Therefore f has two critical points: one at the origin $(0,0)$, and the other at $(\frac{2}{3}, 0)$.



In this example we could have already predicted from the shape of the zero set of f that f has at least two critical points – we don't need to compute the derivatives of f for that. Namely, the zero set of f is a curve which crosses itself at the origin, so the Implicit Function Theorem 6.1 (chapter 2) can't hold at the origin, and hence $f_x = f_y = 0$ there. And in § 2.3 we argued that the function f must have a local maximum in the region D_2 (Figure 2), so f must have at least two critical points. On the other hand, by computing the critical points we have found that there is only one local maximum in the region D_2.

4.4. Another example: Find the critical points of $f(x, y) = x - x^3 - xy^2$. *Solution:* The derivatives of our function are

$$\frac{\partial f}{\partial x} = 1 - 3x^2 - y^2, \qquad \frac{\partial f}{\partial y} = -2xy.$$

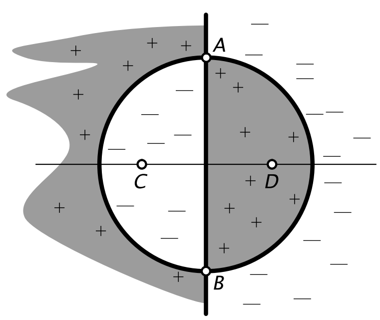
The critical points are therefore the solutions of the equations

$$1 - 3x^2 - y^2 = 0, \qquad -2xy = 0.$$

This is a system of two equations, with two unknowns (that always happens when you look for critical points, since you're looking for solutions of $f_x(x, y) = 0$, $f_y(x, y) = 0$.) The second equation, $-2xy = 0$, implies that either $x = 0$ or $y = 0$ (or both). We have to treat these two cases separately:

The case $x = 0$. If $x = 0$ then we only have the first equation left, which tells us $1 - y^2 = 0$, i.e. $y = \pm 1$. We find two critical points with $x = 0$, namely, $(0, 1)$ and $(0, -1)$.

The other case, $x \neq 0$. If $x \neq 0$, then the second equation ($-2xy = 0$) implies $y = 0$. Substitute this in the first equation and you find $1 - 3x^2 = 0$, i.e. $x = \pm\frac{1}{3}\sqrt{3}$, so that we have two critical points with $x \neq 0$, namely, $(-\frac{1}{3}\sqrt{3}, 0)$ and $(\frac{1}{3}\sqrt{3}, 0)$.



The conclusion is that this function has four critical points, two on the x-axis, and two on the y-axis. Without looking into this in any further detail we don't know if any of these points are local maxima or minima. In general the second derivative test (to be explained in § 9) will provide this information. For this example a look at the zero set of f also helps us figure out what kind of critical points we have found. Since f factors as $f(x, y) = x(1 - x^2 - y^2)$, you see that its zero set consists of the line $x = 0$ (a.k.a. the y-axis), and the unit circle $x^2 + y^2 = 1$. In the above picture $f > 0$ in the grey region, and $f < 0$ in the white area. Consider the right half of the unit disc. The function is positive in the interior, and zero on the boundary of this region. Just as in the "fishy example" of § 2.3, we have another case where the maximum of the function must be attained at one or more interior points of the right half of the unit disc. According to our computation f only has *one* critical point in the right half circle, and therefore this point must be a local maximum of the function. Conclusion: $D = (\frac{1}{3}\sqrt{3}, 0)$ is a local maximum.

In the same spirit you can argue that f has a local minimum at C.

The other two points A, B are neither local maxima nor minima, since *arbitrarily close to A or B* there are both points (x, y) with $f(x, y)$ positive, and points with $f(x, y)$

negative. The points A and B turn out to be "saddle points" (see §9 on the second derivative test.)

5. When you have more than two variables

The whole discussion so far has been about functions of two variables. Fortunately, not much changes when you have more variables. The concepts *local minimum* and *local maximum* are defined in the same way, and it turns out that any *interior local maximum or minimum must be a critical point of the function.* Here, by definition, a critical point of a function $w = f(x_1, \ldots, x_n)$ of n variables is a solution of the equations

$$\frac{\partial f}{\partial x_1}(x_1, \cdots, x_n) = 0$$
$$\vdots$$
$$\frac{\partial f}{\partial x_n}(x_1, \cdots, x_n) = 0.$$

Note that there are n equations, and there are also n unknowns $(x_1, \ldots, x_n)$ so that you should *in principle* be able to solve these equations. In practice the system of equations you get can be very easy, or simply impossible to solve.

6. Problems

82. Find all critical points of the following functions. Try to classify them into local/global maxima/minima, saddles, or other kind of critical points. (Write clear solutions. You will need your solutions later in problem 96.)

(i) $f(x, y) = x^2 + 4y^2 - 2x + 8y - 1$
(ii) $f(x, y) = x^2 - y^2 + 6x - 10y + 2$
(iii) $f(x, y) = x^2 + 4xy + y^2 - 6y + 1$
(iv) $f(x, y) = x^2 - xy + 2y^2 - 5x + 6y - 9$
(v) $f(x, y) = y^2 - 18x^2 + x^4$
(vi) $f(x, y) = y^4 - 4y^2 - 18x^2 + x^4$
(vii) $f(x, y) = 9 + 4x - y - 2x^2 - 3y^2$
(viii) $f(x, y) = xy(4 - x - 2y)$
(ix) $f(x, y) = (x - y^2)(x - 1)$
(x) $f(x, y) = (x - y)(xy - 4)$
(xi) $f(x, y) = y^2 + \cos x$
(xii) $f(x, y) = x^2y - \frac{1}{3}y^3$
(xiii) $f(x, y) = (x - y^2)(x - 1)$
(xiv) $f(x, y) = (x - y)(xy - 4)$
(xv) $f(x, y) = x^2$
(xvi) $f(x, y) = x^2y$
(xvii) $f(x, y) = \left(1 - x^2 - y^2\right)^2$
(xviii) $f(x, y) = x^2y$

83. **(i)** Draw the zero set of the function $f(x, y) = \sin(x)\sin(y)$.

(ii) Where is the function f positive? Find as many critical points as you can without computing f_x or f_y.

(iii) Find all critical points of $f(x, y)$. Which are local minima or local maxima?

84. Find the critical points of the function

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

85. Draw the zero set and find the critical points of the functions

$$f(x, y, z) = x^2 + y^2 - z^2 \text{ and } g(x, y, z) = x^2 - y^2 - z^2$$

86. Given the three points $(1, 4)$, $(5, 2)$, and $(3, -2)$,

$$f(x, y, z) = (x - 1)^2 + (y - 4)^2 + (x - 5)^2 + (y - 2)^2 + (x - 3)^2 + (y + 2)^2$$

is the sum of the squares of the distances from point (x, y) to the three points. Find x and y so that this quantity is minimized.

87. Given the three points (a, b), (c, d), and (e, f), let $f(x, y, z)$ be the sum of the squares of the distances from point (x, y) to the three points. Find x and y so that this quantity is minimized.

88. Suppose a function $f(x, y)$ factors, i.e. you can write it as the product of two other differentiable functions, $f(x, y) = g(x, y)h(x, y)$. *Prove: if a point (a, b) lies in the zero set of g and also in the zero set of h, then (a, b) is a critical point of f.*

89. Find the critical points of the functions

(i) $f(x, y, z) = x^2 + y^2 + z^2 - 2x + 4y - 2$ **(ii)** $f(x, y, z) = x^4 + y^2 + z^2 - 2xz + 4y$
(iii) $f(x, y, z) = xyze^{-x-y-z}$ **(iv)** $f(x, y, z) = x^2 + y^2 + z^2 - 2xyz$

7. A Minimization Problem: Linear Regression

Suppose you are measuring two quantities x and y in some experiment, and suppose that you expect that there is a linear relation of the form $y = ax + b$ between x and y. If you have a set of data points (x_k, y_k) from your experiment, then what do they tell you about a and b? ***Which choice of coefficients a and b bests fits your data?*** Because of experimental errors you would not expect your data points to lie on a straight line, but instead, you expect them to be clustered around a straight line. You could plot the data points, get a ruler, and draw a straight line by hand that looks like the best match – then you could measure a, b from your drawing. A more systematic approach is to first define what you mean by "best match" and then find the line that best matches according to your chosen criterion.

A very common criterion is the least-mean-square-fit. To describe it, imagine you have N data points, (x_1, y_1), $\ldots (x_N, y_N)$, and consider the line with coefficients a and b. Most data points (x_k, y_k) will then probably not lie on the line $y = ax + b$, and one uses

$$E_k = \tfrac{1}{2}\bigl(ax_k + b - y_k\bigr)^2$$

as a measure for the mismatch between the data point (x_k, y_k) and the line $y = ax+b$ (the factor $\frac{1}{2}$ makes formulas later on nicer). Adding all these errors we get the total "mean square" error

$$E = E_1 + \cdots + E_N.$$

If we think of all the numbers $x_1, \ldots, x_N, y_1, \ldots, y_N$ as given constants (after all, you measured them, so you shouldn't change them any more), then the total error only depends on the coefficients a and b. It is a measure for how well the line $y = ax + b$ fits our data points, and the common method of ***linear regression*** consists in choosing the coefficients a and b so as to minimize this error E.

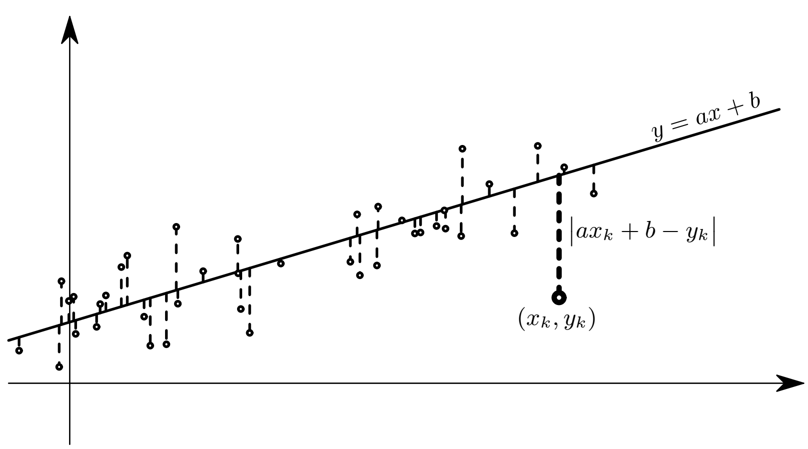


Figure 5: Which line best fits a set of data points?

This leads us to the problem of finding the critical points of the total error E as a function of a and b. We have to solve

$$\frac{\partial E}{\partial a} = 0 \qquad \frac{\partial E}{\partial b} = 0.$$

The total error is the sum of the individual errors $E_k(a, b)$ so we get

$$\frac{\partial E}{\partial a} = \frac{\partial E_1}{\partial a} + \cdots + \frac{\partial E_N}{\partial a}, \qquad \frac{\partial E}{\partial b} = \frac{\partial E_1}{\partial b} + \cdots + \frac{\partial E_N}{\partial b}.$$

The individual errors have the following derivatives:

$$\frac{\partial E_k}{\partial a} = x_k\bigl(ax_k + b - y_k\bigr), \qquad \frac{\partial E_k}{\partial b} = ax_k + b - y_k.$$

Adding all these derivatives then leads to

$$\begin{aligned}\frac{\partial E}{\partial a} &= \sum x_k\bigl(ax_k + b - y_k\bigr) \\ &= (\textstyle\sum x_k^2)a + (\sum x_k)b - \sum x_k y_k\end{aligned}$$

and

$$\begin{aligned}\frac{\partial E}{\partial b} &= \sum\bigl\{ax_k + b - y_k\bigr\} \\ &= (\textstyle\sum x_k)a + Nb - \sum y_k\end{aligned}$$

Here "$\sum$" represents summation over $k = 1, \cdots, N$, i.e. $\sum x_k y_k = x_1 y_1 + \cdots + x_N y_N$, etc.

If (a, b) is a critical point then a and b must satisfy

$$\begin{aligned}(\textstyle\sum x_k^2)a + (\sum x_k)b &= \textstyle\sum x_k y_k \\ (\textstyle\sum x_k)a + Nb &= \textstyle\sum y_k\end{aligned}$$

These are two linear equations for the two unknowns a and b. Solving them leads to

$$a = \frac{N\sum x_k y_k \;-\; \sum x_k \sum y_x}{N\sum x_k^2 - \bigl(\sum x_k\bigr)^2}\;; \qquad b = \frac{-\sum x_k \sum x_k y_k + \sum x_k^2 \sum y_k}{N\sum x_k^2 - \bigl(\sum x_k\bigr)^2}\,.$$

These are the standard formulas for the coefficients a and b provided by the method of linear regression. Most calculators, and certainly all spreadsheets (like Excel) have these formulas preprogrammed, so you only have to enter the data points (x_k, y_k) and "push the right button" to get a and b.

8. Problems

90. You are given N measurements $x_1, \ldots, x_N$ from some experiment, and, inspired by the Linear Regression example, you decide to see which number a "best fits the data." You define the error (or "measure of misfit") for each measurement to be

$$E_k(a) = \tfrac{1}{2}(a - x_k)^2$$

and you look for the number a which minimizes the total error

$$E(a) = E_1(a) + \cdots + E_N(a).$$

(i) Is this a problem about several variable calculus, or about one variable calculus?

(ii) Which number a do you find?

91. You have a series of data points (x_k, y_k), and when you plot them you think you see a convex curve rather than a straight line. In fact it looks like a parabola to you, and so you set out to find a quadratic function $y = ax^2 + bx + c$ which minimizes the error

$$E(a, b, c) = E_1 + \cdots + E_N, \text{ with } E_k(a, b, c) = \tfrac{1}{2}\bigl(ax_k^2 + bx_k + c - y_k\bigr)^2.$$

(i) How many variables are there in this problem?

(ii) If (a, b, c) is a critical point of $E(a, b, c)$ then a, b, and c satisfy three linear equations. Find these equations (don't solve them).

92. A measurement in a certain experiment results in three numbers (x, y, z). The point of the experiment is to see if there is a linear relation of the form $z = ax + by + c$ between the three measured quantities, and to estimate the coefficients a, b, c.

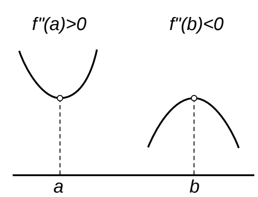
After repeating the experiment N times you have N data points (x_k, y_k, z_k) $(k = 1, \ldots, N)$. You decide to choose a, b, c so as to minimize the mean square error

$$E = E_1 + \cdots + E_N, \text{ with } E_k(a, b, c) = \tfrac{1}{2}\big(ax_k + by_k + c - z_k\big)^2.$$

Which (linear) equations do you get for a, b, and c?

9. The Second Derivative Test

9.1. The one-variable second derivative test. For a function $y = f(x)$ of one variable you can tell if a critical point a is a local maximum or minimum by looking at the sign of the second derivative $f''(a)$ of the function at that point. If $f''(a) > 0$ then the graph of f is curved upwards and f has a local minimum at a, if $f''(a) < 0$ then f has a local max. This section is about the analogous test for critical points of functions of two variables.



One way to understand the second derivative test is to look at the Taylor expansion of the function $y = f(x)$. If $x = a$ is a critical point for f, then

$$f(a + \Delta x) = f(a) + f'(a)\Delta x + \tfrac{1}{2} f''(a)(\Delta x)^2 + \cdots$$

Since a is a critical point of f we have $f'(a) = 0$, so that the Taylor expansion reduces to

$$f(a + \Delta x) = f(a) + \tfrac{1}{2} f''(a)(\Delta x)^2 + \cdots \tag{44}$$

If we ignore the remainder term (the dots), then we find that

$$f(a + \Delta x) \approx f(a) + \tfrac{1}{2} f''(a)(\Delta x)^2.$$

Near the critical point the graph of $y = f(x)$ is a approximately a parabola. It is curved upwards if $f''(a) > 0$, and downwards if $f''(a) < 0$.

To apply the same reasoning to a function of two (or more) variables we need to know the Taylor expansion of such a function.

9.2. Taylor's formula for a function of several variables. The Taylor expansion of a function $z = f(x, y)$ should give us an approximation of $f(a + \Delta x, b + \Delta y)$ in terms involving powers of Δx and Δy. There is a general formula, but here we only need the second order terms, so we'll derive those and stop there.

The trick to finding the Taylor expansion is to consider the function

$$g(t) = f(a + t\Delta x, b + t\Delta y). \tag{45}$$

By definition

$$g(1) = f(a + \Delta x, b + \Delta y)$$

is the quantity we want to approximate, and $g(0) = f(a, b)$. Since $g(t)$ is a function of one variable, we can apply Taylor's formula from math 222 to it. You get:

$$g(t) = g(0) + g'(0)t + g''(0)\frac{t^2}{2!} + \cdots \tag{46}$$

The dots contain the remainder term, which we will ignore in this course. Now we set $t = 1$, and we get

$$g(1) = g(0) + g'(0) + \frac{1}{2}g''(0) + \cdots$$

The derivatives of g can be computed with the chain rule:

$$\begin{aligned} g'(t) &= \frac{df(a+t\Delta x, b+t\Delta y)}{dt} \\ &= f_x(a+t\Delta x, b+t\Delta y)\frac{d(a+t\Delta x)}{dt} + f(a+t\Delta x, b+t\Delta y)\frac{d(b+t\Delta y)}{dt} \\ &= f_x(a+t\Delta x, b+t\Delta y)\Delta x + f_y(a+t\Delta x, b+t\Delta y)\Delta y. \end{aligned} \tag{47}$$

The second derivative is

$$\begin{aligned} g''(t) = f_{xx}(a+t\Delta x, b+t\Delta y)(\Delta x)^2 \\ + 2f_{xy}(a+t\Delta x, b+t\Delta y)\Delta x\Delta y \\ + f_{yy}(a+t\Delta x, b+t\Delta y)(\Delta y)^2. \end{aligned} \tag{48}$$

In computing $g''(t)$ you run into terms involving f_{xy} and terms with f_{yx}. Because of Clairaut's theorem these are the same, and combining them leads to the coefficient "2" in front of f_{xy} above.

Setting $t = 0$ in (47) and in (48) gives you expressions for $g'(0)$ and $g''(0)$, and by substituting these in (46) you get ***the second order Taylor expansion of a function of two variables:***

$$\begin{aligned} f(a+\Delta x, b+\Delta y) = f(a,b) + f_x(a,b)\Delta x + f_y(a,b)\Delta y \\ + \frac{1}{2}\Big\{f_{xx}(a,b)(\Delta x)^2 + 2f_{xy}(a,b)\Delta x\Delta y + f_{yy}(a,b)(\Delta y)^2\Big\} + \cdots \end{aligned} \tag{49}$$

The first three terms are exactly the linear approximation of the function from Chapter 2, §10. As always, the dots contain the remainder term. By carefully including the one-variable Lagrange remainder in the derivation you can get a formula for the remainder in (49). We will not do that, but it can be shown that the remainder is $o\big((\Delta x)^2 + (\Delta y)^2\big)$, i.e. that it is small compared to the other terms in the expansion, at least when Δx and Δy are small.

9.3. Example: Compute the Taylor expansion of $f(x,y) = \sin 2x \cos y$ at the point $(\frac{1}{6}\pi, \frac{1}{6}\pi)$. To find the expansion we need to compute $f, f_x, f_y, f_{xx}, f_{xy}$, and f_{yy} at $(\frac{1}{6}\pi, \frac{1}{6}\pi)$. Here goes:

$$\begin{aligned} f &= \sin 2x\cos y = \tfrac{3}{4} & f_{xx} &= -4\sin 2x\cos y = -3 \\ f_x &= 2\cos 2x\cos y = \tfrac{1}{2}\sqrt{3} & f_{xy} &= -2\cos 2x\sin y = -\tfrac{1}{2} \\ f_y &= -\sin 2x\sin y = -\tfrac{1}{4}\sqrt{3} & f_{yy} &= -\sin 2x\cos y = -\tfrac{3}{4}. \end{aligned}$$

Substituting in the Taylor expansion we get

$$\begin{aligned} &f\big(\tfrac{1}{6}\pi + \Delta x, \tfrac{1}{6}\pi + \Delta y\big) \\ &\qquad = \tfrac{3}{4} + \tfrac{1}{2}\sqrt{3}\Delta x - \tfrac{1}{4}\sqrt{3}\Delta y + \frac{1}{2}\Big\{-3(\Delta x)^2 - 2\cdot\tfrac{1}{2}\Delta x\Delta y - \tfrac{3}{4}(\Delta y)^2\Big\} + \cdots \\ &\qquad = \tfrac{3}{4} + \tfrac{1}{2}\sqrt{3}\Delta x - \tfrac{1}{4}\sqrt{3}\Delta y - \tfrac{3}{2}(\Delta x)^2 - \tfrac{1}{2}\Delta x\Delta y - \tfrac{3}{8}(\Delta y)^2 + \cdots \end{aligned}$$

Note that the first three terms in the expansion are the linear approximation of the function:

$$f\big(\tfrac{1}{6}\pi + \Delta x, \tfrac{1}{6}\pi + \Delta y\big) = \tfrac{3}{4} + \tfrac{1}{2}\sqrt{3}\Delta x - \tfrac{1}{4}\sqrt{3}\Delta y + \cdots$$

9.4. Another example: the Taylor expansion of $f(x,y) = x^3 + y^3 - 3xy$ at the point $(1,1)$. The function $f(x,y) = x^3 + y^3 - 3xy$ has the following derivatives at $(1,1)$

$$
\begin{aligned}
f &= x^3 + y^3 - 3xy &&= 1 \qquad & f_{xx} &= 6x &&= 6\\
f_x &= 3x^2 - 3y &&= 0 & f_{xy} &= -3 &&= -3\\
f_y &= 3y^2 - 3x &&= 0 & f_{yy} &= 6y &&= 6
\end{aligned}
$$

The first derivatives vanish, so $(1,1)$ is a critical point of f. The second order Taylor expansion of f at $(1,1)$ is

$$f(1+\Delta x, 1+\Delta y) = 1 + 3(\Delta x)^2 - 3\Delta x\Delta y + 3(\Delta y)^2 + \cdots \tag{50}$$

Quadratic forms

An expression of the type

$$Q(x,y) = Ax^2 + Bxy + Cy^2$$

where A, B, and C are constants, is called a *quadratic form* in the variables x and y. They show up whenever you compute second order Taylor expansions of a function of two variables. In this context you want to know for which (x,y) the form $Q(x,y)$ is positive or negative. Since

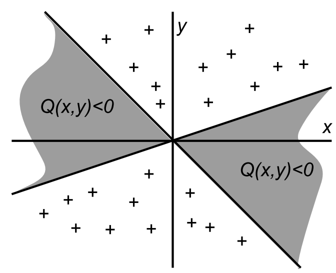
$$
\begin{aligned}
Q(x,y) &= y^2\left\{A\left(\tfrac{x}{y}\right)^2 + B\tfrac{x}{y} + C\right\}\\
&= y^2(At^2 + Bt + C)
\end{aligned}
$$

the question of when $Q(x,y)$ is positive or negative is closely related to the question of when the quadratic function $At^2 + Bt + C$ is positive or negative. There is a familiar method for distinguishing between these cases.

Completing the square. For instance, if the form is $Q(x,y) = -3x^2 + 9xy + 6y^2$ then you rewrite this as follows:

$$
\begin{aligned}
Q &= -3x^2 + 6xy + 9y^2\\
&= -3\left(x^2 - 2xy - 3y^2\right)\\
&= -3\big[\underbrace{x^2 - 2xy + y^2}_{\text{complete square}} - 4y^2\big]\\
&= -3\left[(x-y)^2 - 4y^2\right]\\
&= -3(x - y - 2y)(x - y + 2y)\\
&= -3(x - 3y)(x + y).
\end{aligned}
$$

This shows that $Q(x,y) > 0$ when $y > \frac{1}{3}x$ or $y < -x$, and $Q(x,y) < 0$ when $-x < y < \frac{1}{3}x$.



Classification of forms. When you apply this method to a quadratic form you always end up with one of the following results:

Q **is an indefinite form** when Q is the product of two distinct factors,

$$Q(x,y) = (ax + by)(cx + dy),$$

as in the example above.

Q **is a positive-definite form** when $Q(x,y)$ is the sum of two squares,

$$Q(x,y) = (ax + by)^2 + cy^2 \quad (c > 0)$$

Q **is a negative-definite form** when $-Q(x,y)$ is the sum of two squares,

$$Q(x,y) = -(ax + by)^2 - cy^2 \quad (c > 0)$$

Q **is a degenerate form** when $Q(x,y)$ is a perfect square,

$$Q(x,y) = (ax + by)^2.$$

□

Figure 6: Finding the signs of a quadratic form

To see what kind of critical point $(1,1)$ is, we have to analyze the second order, quadratic, terms

$$3(\Delta x)^2 - 3\Delta x\Delta y + 3(\Delta y)^2 \tag{51}$$

This expression is a quadratic form in Δx and Δy, and by completing the square (see Figure 6) you find that

$$3(\Delta x)^2 - 3\Delta x\Delta y + 3(\Delta y)^2 = 3\Big[\big(\Delta x - \tfrac{1}{2}\Delta y\big)^2 + \tfrac{3}{4}(\Delta y)^2\Big].$$

In particular, the quadratic terms in the Taylor expansion of f at the critical point are always positive, no matter what Δx and Δy we choose (as long as they are not both zero). If we are allowed to ignore the remainder term (the "$\cdots$"), then this implies that the function has a local minimum: after all, the Taylor expansion (50) says that for small Δx and Δy the function value $f(1+\Delta x, 1+\Delta y)$ is

$$f(1+\Delta x, 1+\Delta y) \approx f(1,1) + 3\big(\Delta x - \tfrac{1}{2}\Delta y\big)^2 + \tfrac{9}{4}(\Delta y)^2$$

which is more than $f(1,1)$.

9.5. Example of a saddle point. The same function $f(x,y) = x^3 + y^3 - 3xy$ has another critical point, namely, the origin. By calculating the derivatives at $(0,0)$ you find that the Taylor expansion at the origin is

$$f(\Delta x, \Delta y) = -3\Delta x\Delta y + \cdots \tag{52}$$

Ignoring the remainder terms we see that near the origin $f(\Delta x, \Delta y) \approx -3\Delta x\Delta y$, which suggests that f is positive when Δx and Δy are both positive or both negative, while f is negative when Δx and Δy have opposite signs.

Arbitrarily close to the origin the function f therefore has both positive and negative values, and therefore f has neither a local maximum nor a local minimum at the origin. In fact the Taylor expansion (52) suggests that the graph of f should look like that of the "saddle function" $z = xy$.

9.6. The two-variable second derivative test. The last two examples essentially show you how the second derivative test for functions of two variables works. To explain how it works in general, let's suppose a function f has a critical point at (a,b). Then the first partial derivatives of f vanish at (a,b) and hence the Taylor expansion simplifies a bit. You get

$$\begin{aligned} f(a+\Delta x, b+\Delta y) = f(a,b) + {}& \\ \frac{1}{2}\Big\{f_{xx}(a,b)(\Delta x)^2 + 2f_{xy}(a,b)\Delta x\Delta y &+ f_{yy}(a,b)(\Delta y)^2\Big\} + \cdots \end{aligned} \tag{53}$$

This is the two-variable analog of equation (44). To see if (a,b) is a local maximum or minimum (or neither), we have to see if the quadratic terms in (53) are always negative, positive, or if they can have any sign, depending on the choice of Δx, Δy.

The precise statement of the second derivative test uses the terminology introduced in Figure 6.

Theorem (second derivative test). *If (a,b) is a critical point of $f(x,y)$, and if*

$$Q(\Delta x, \Delta y) = \frac{1}{2}\Big\{f_{xx}(a,b)(\Delta x)^2 + 2f_{xy}(a,b)\Delta x\Delta y + f_{yy}(a,b)(\Delta y)^2\Big\}$$

is the quadratic part of the Taylor expansion of f at the critical point, then

- *If Q is* ***positive definite*** *then (a,b) is a* ***local minimum*** *of f,*
- *If Q is* ***negative definite*** *then (a,b) is a* ***local maximum*** *of f,*
- *If Q is* ***indefinite*** *then (a,b) is a* ***saddle point*** *of f*
- *If Q is* ***degenerate*** *the second derivative test is inconclusive.*

When the form Q is indefinite, so that it can be factored as

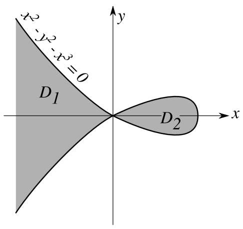
$$Q(\Delta x, \Delta y) = (k\Delta x + l\Delta y)(m\Delta x + n\Delta y)$$

then the level set of the function f containing the critical point (a, b) consists of two curves. One of these curves is tangent to the line

$$k\Delta x + l\Delta y = 0, \text{ i.e. } k(x-a) + l(y-b) = 0$$

while the other is tangent to

$$m\Delta x + n\Delta y = 0, \text{ i.e. } m(x-a) + l(y-b) = 0.$$



9.7. Example: Apply the second derivative test to the fishy example. In § 2.3 and § 4.3 we had found that the function $f(x, y) = x^2 - x^3 - y^2$ has two critical points, one at the origin, and one at the point $(\frac{2}{3}, 0)$. By carefully looking at the zero set of the function we discovered that the origin is neither a local maximum nor a local minimum, and that the point $(\frac{2}{3}, 0)$ is a local maximum. The second derivative test provides a more systematic way of reaching these conclusions. To apply the test we need to know the second derivatives of f at the critical points. They are:

(x, y)	$f_{xx}(x, y)$	$f_{xy}(x, y)$	$f_{yy}(x, y)$
(x, y)	$2 - 6x$	0	-2
$(0, 0)$	2	0	-2
$(\frac{2}{3}, 0)$	-2	0	-2

Therefore the second order Taylor expansion of f at the origin is

$$\begin{aligned} f(\Delta x, \Delta y) &= f(0,0) + \tfrac{1}{2}\left\{2 \cdot (\Delta x)^2 + 2 \cdot 0 \cdot \Delta x \Delta y + (-2)(\Delta y)^2\right\} + \cdots \\ &= (\Delta x)^2 - (\Delta y)^2 + \cdots \\ &= (\Delta x - \Delta y)(\Delta x + \Delta y) + \cdots \end{aligned}$$

The quadratic part of the Taylor expansion factors (we have called this "indefinite"). It can be both positive and negative, depending on your choice of Δx and Δy. The second derivative test implies that the origin is a saddle point. It also says that the zero set of f near the origin consists of two curves, whose tangents at the origin are given by the two equations

$$\Delta x - \Delta y = 0 \text{ and } \Delta x + \Delta y = 0.$$

In this case the point (a, b) is the origin, so $\Delta x = x - a = x$ and $\Delta y = y - b = y$, and the two tangents are the lines $y = \pm x$.

The second origin Taylor expansion at the other critical point $(\frac{2}{3}, 0)$ is given by

$$f(\tfrac{2}{3} + \Delta x, \Delta y) = -(\Delta x)^2 - (\Delta y)^2 + \cdots$$

This time you see that the second order terms of the Taylor expansion are negative definite. The second derivative test therefore says that we have a local maximum at $(\frac{2}{3}, 0)$.

10. Problems

93. Compute the second order Taylor expansion of the following functions at the indicated points:

(i) $f(x, y) = (1 - x + xy)^2$ at $(0, 0)$ **(ii)** $f(x, y) = (1 - x + xy)^2$ at $(1, 1)$
(iii) $f(x, y) = e^{x - y^2}$ at $(0, 0)$ **(iv)** $f(x, y) = e^{x - y^2}$ at $(1, 1)$
(v) $f(x, y) = \dfrac{x}{1 - y}$ at $(0, 0)$ **(vi)** $f(x, y) = \dfrac{x}{1 + y}$ at $(1, 0)$

94. Factor, or complete the square in the following quadratic forms, draw their zero sets, and determine where they are positive definite, negative definite, indefinite or degenerate.

(i) $Q(x,y) = x^2 + 3xy + y^2$
(ii) $Q(x,y) = x^2 + xy + y^2$
(iii) $Q(x,y) = 2x^2 + 3xy - 4y^2$
(iv) $Q(x,y) = 2x^2 + 3xy - 5y^2$
(v) $Q(\Delta x, \Delta y) = (\Delta x)^2 + (\Delta y)^2$
(vi) $Q(\Delta x, \Delta y) = (\Delta x)^2 - 3(\Delta y)^2$
(vii) $Q(\Delta x, \Delta y) = \Delta x \Delta y$
(viii) $Q(\Delta x, \Delta y) = \Delta x \Delta y - 2(\Delta y)^2$

95. If a is a constant, then for which values of a is the form $Q(x,y) = x^2 + 2axy + y^2$ positive/negative definite, indefinite, or degenerate?

96. Find all critical points of the following functions (you did many of these in problem 82). Apply the second derivative test to all critical points you find.

(i) $f(x,y) = x^2 + 4y^2 - 2x + 8y - 1$
(ii) $f(x,y) = x^2 - y^2 + 6x - 10y + 2$
(iii) $f(x,y) = x^2 + 4xy + y^2 - 6y + 1$
(iv) $f(x,y) = x^2 - xy + 2y^2 - 5x + 6y - 9$
(v) $f(x,y) = y^2 - 18x^2 + x^4$
(vi) $f(x,y) = y^4 - 4y^2 - 18x^2 + x^4$
(vii) $f(x,y) = 9 + 4x - y - 2x^2 - 3y^2$
(viii) $f(x,y) = xy(4 - x - 2y)$
(ix) $f(x,y) = (x - y^2)(x - 1)$
(x) $f(x,y) = (x - y)(xy - 4)$
(xi) $f(x,y) = y^2 + \cos x$
(xii) $f(x,y) = x^2 y - \frac{1}{3}y^3$
(xiii) $f(x,y) = (x - y^2)(x - 1)$
(xiv) $f(x,y) = (x - y)(xy - 4)$
(xv) $f(x,y) = x^2$
(xvi) $f(x,y) = x^2 - y^4$
(xvii) $f(x,y) = x^2 + y^4$
(xviii) $f(x,y) = x^2 y$

97. **(i)** Draw the zero set of the function $f(x,y) = \sin(x)\sin(y)$. **(ii)** Where is the function f positive? Find as many critical points as you can without computing f_x or f_y.

(iii) Find all critical points of $f(x,y)$. Which are local minima or local maxima?

98. Find all critical points of the following functions, and apply the second derivative test to the points you find.

(i) $f(x,y) = x^2 + y^2 - \frac{1}{2}xy^2$
(ii) $f(x,y) = x^2 + y^2 - x^2y^2$
(iii) $f(x,y) = x + 2y - xy^2$
(iv) $f(x,y) = 8x^4 + y^4 - xy^2$

99. Suppose that $f(x,y) = x^2 + y^2 + kxy$. Find and classify the critical points, and discuss how they change when k takes on different values.

100. Consider the function $f(x,y) = x^3 - 3x^2y$.

(i) Show that $(0,0)$ is the only critical point of f.

(ii) Show that the second derivative test is inconclusive for f.

(iii) Draw the zero set of f, and indicate where $f \gtrless 0$.

(iv) What kind of critical point is $(0,0)$?

11. Second derivative test for more than two variables

The ideas that lead to the second derivative test for functions of two variables also work if you have a function with more variables. However, in math 234 you won't be asked to use the test in any problems involving more than two variables, and this short section tries to explain why.

11.1. The second order Taylor expansion. If $z = f(x_1, x_2, \cdots, x_n)$ is a function of n variables, then its Taylor expansion of order two at some point $(a_1, a_2, \cdots, a_n)$ turns out to be

$$\begin{aligned} f(a_1 + \Delta x_1, \cdots, a_n + \Delta x_n) = f(a_1, \cdots, a_n) &+ f_{x_1}\Delta x_1 + \cdots + f_{x_n}\Delta x_n \\ &+ \tfrac{1}{2}\Big\{ f_{x_1x_1}(\Delta x_1)^2 + \cdots + f_{x_1x_n}\Delta x_1 \Delta x_n \\ &\quad + f_{x_2x_1}\Delta x_2\Delta x_1 + \cdots + f_{x_2x_n}\Delta x_2\Delta x_n \\ &\qquad \vdots \\ &\quad + f_{x_nx_1}\Delta x_n\Delta x_1 + \cdots + f_{x_nx_n}(\Delta x_n)^2 \Big\} + \cdots \end{aligned}$$

where the partial derivatives f_{x_i} and $f_{x_i x_j}$ are to be evaluated at the point $(a_1, \cdots, a_n)$. The same trick involving the function "$g(t)$" that was used in §9.2 to derive the two-variable Taylor expansion works without modification.

If $(a_1, \cdots, a_n)$ is a critical point then $f_{x_1} = f_{x_2} = \cdots = f_{x_n} = 0$, so the linear terms are absent, and the function is described by the quadratic terms of the Taylor expansion

$$\begin{aligned} f(a_1 + \Delta x_1, \cdots, a_n + \Delta x_n) = \\ f(a_1, \cdots, a_n) + \tfrac{1}{2}\Big\{ & f_{x_1x_1}(\Delta x_1)^2 + \cdots + f_{x_1x_n}\Delta x_1 \Delta x_n \\ & + f_{x_2x_1}\Delta x_2 \Delta x_1 + \cdots + f_{x_2x_n}\Delta x_2\Delta x_n \\ & \qquad \vdots \\ & + f_{x_nx_1}\Delta x_n \Delta x_1 + \cdots + f_{x_nx_n}(\Delta x_n)^2 \Big\} + \cdots \end{aligned}$$

Just as in the two-variable case you could now try to see if the quadratic terms are positive definite or negative definite by completing squares. The procedure is however much more complicated, and best understood in terms of *"eigenvalues of matrices"*, a subject which is explained in courses on linear algebra or matrix algebra (math 340 or 320). Therefore, we will only use the second derivative test for functions of two variables in this course.

12. Optimization with constraints

In many optimization problems you want to find the maximal or minimal value of a function $f(x, y)$ where (x, y) can be any point satisfying a certain ***constraint***

$$g(x, y) = C. \tag{54}$$

Thus the domain D of the function you want to minimize consists of all points (x, y) which satisfy the equation $g(x, y) = C$: it is a level set of g.

12.1. Solution by elimination or parametrization. One approach to minimization problems with a constraint is to "eliminate one variable." If you are asked to find the minimal value that $f(x, y)$ can have if (x, y) must satisfy the constraint $g(x, y) = C$, then you first try to solve the constraint equation:

$$g(x, y) = C \iff y = h(x).$$

Now the only (x, y) that you have to consider are points of the form $(x, h(x))$, so the old minimization problem is equivalent to a new problem: find the minimal value of $F(x) = f(x, h(x))$, where there are no constraints on x. This new problem is a one variable problem of the kind we learned to solve in math 221.

12.2. Example. *Which rectangle with perimeter 1 has the largest area?*

If the sides of the rectangle are x and y, then its area is xy and its perimeter is $2(x + y)$. Hence the function we want to maximize is $f(x, y) = xy$ and the constraint is $g(x, y) = 2(x + y) = 1$.

Solving the constraint for y tells you that $y = \frac{1}{2} - x$, so we want to maximize the function $F(x) = f(x, \frac{1}{2} - x) = x(\frac{1}{2} - x)$. There is no constraint on x... well, this is not quite true: x cannot be negative, and neither can $y = \frac{1}{2} - x$, so we want to maximize $F(x) = x(\frac{1}{2} - x)$ over all x in the interval $0 \le x \le \frac{1}{2}$.

12.3. Example. *Maximize $x + 2y$ over the unit circle.*

We are asked to find the maximal value of $f(x, y) = x + 2y$ where (x, y) is allowed to be any point which satisfies the constraint $g(x, y) = x^2 + y^2 = 1$. If we try to solve for y we find that there are two solutions, $y = \pm\sqrt{1 - x^2}$, and so the "function" $F(x) = x + 2y = x \pm 2\sqrt{1 - x^2}$ is note really a function at all. In this case we can still solve the problem by noting that any point on the unit circle can be written as $(x, y) = (\cos\theta, \sin\theta)$ for some angle θ, and thus we have to maximize the function

$$F(\theta) = f(\cos\theta, \sin\theta) = \cos\theta + 2\sin\theta.$$

Here there are no constraints on θ, and we again have a first semester calculus problem.

12.4. Solution by Lagrange multipliers. In both examples above we were lucky because we could either solve the constraint equation or we could parametrize all possible points that satisfy the constraint. There is a method due to Joseph-Louis Lagrange (known from the remainder term) that does not require this kind of luck. His method is based on the following observation (see Figure 7).

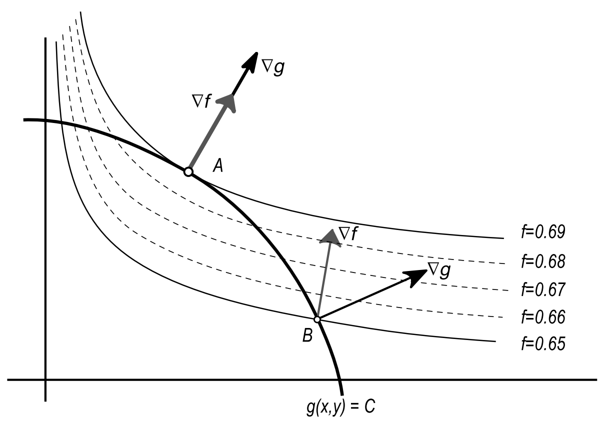


Figure 7: Lagrange multipliers

Let $B = (x, y)$ be a point on the constraint set as in the figure. Assume that $\vec{\nabla} g \neq \vec{0}$ at B, then near B the Implicit Function Theorem says that the constraint set $g(x, y) = C$ is a curve, and that its tangent is perpendicular to $\vec{\nabla} g(B)$.

If $\vec{\nabla} f(B)$ is not perpendicular to the constraint set at B, then it provides us a direction along the constraint set in which f will increase (see Figure 7). Therefore f does not have a maximum at B. It follows that at a maximum of f on the constraint set $g(x, y) = C$ the gradient $\vec{\nabla} f(B)$ must be perpendicular to the constraint set, and hence it must be parallel to $\vec{\nabla} g(B)$. Since one vector is parallel to another if it is a multiple of the other vector, we have found the following fact.

12.5. Theorem (Lagrange multipliers). *If the function $z = f(x, y)$ attains its largest value among all points which satisfy the constraint $g(x, y) = C$ at the point (a, b), and if*

$$\vec{\nabla} g(a, b) \neq 0, \tag{55}$$

then the point (a, b) *satisfies the Lagrange Multiplier equations,*

$$\vec{\nabla} f(a,b) = \lambda \vec{\nabla} g(a,b) \tag{56}$$

The number λ is called the *Lagrange multiplier*, and it is one of the unknowns in the equations you must solve when you use Lagrange's method.

12.6. Example. We again try to find the largest rectangle with perimeter 1, as in example 12.2.

The problem is to maximize $f(x, y) = xy$ with constraint $g(x, y) = 2x + 2y = 1$. We compute the gradients

$$\vec{\nabla} f = \begin{pmatrix} y \\ x \end{pmatrix}, \qquad \vec{\nabla} g = \begin{pmatrix} 2 \\ 2 \end{pmatrix},$$

The gradient of g never vanishes, i.e. $\vec{\nabla} g(x, y) \neq \vec{0}$ for all (x, y), so Lagrange tells us that any minimum or maximum satisfies the following equations

$$f_x = \lambda g_x, \text{ i.e. } y = 2\lambda \tag{57}$$

$$g_y = \lambda g_y, \text{ i.e. } x = 2\lambda \tag{58}$$

$$g(x, y) = C, \text{ i.e. } 2x + 2y = 1. \tag{59}$$

The first two equations come from $\vec{\nabla} f = \lambda \vec{\nabla} g$, and the last equation is the constraint. We have three equations, and we also have three unknowns: x, y and the Lagrange multiplier λ.

In this case it is easy to solve the equations: the first two say that both y and x equal 2λ, so in particular, they equal each other: $y = x$. This already tells us that the solution is a square! To complete the problem we must still solve for x, y, λ. Since $x = y$ the constraint implies $4x = 1$, so $x = y = \frac{1}{4}$. Finally, either of the first two equations provides $\lambda = \frac{1}{2}x = \frac{1}{2}y = \frac{1}{8}$.

What is the meaning of λ*?* In this example you see that we first found the solution (x, y), and then computed λ. The multiplier λ is the ratio between the lengths of the gradients of f and g at the maximum, and is usually of no interest. Nonetheless, when using Lagrange's method, *you must always also find* λ*,* or at least make sure that a λ exists for the x and y you have found.

Did we find a maximum or a minimum? Lagrange's method does not tell us if we have a maximum or a minimum, and we will have to use different methods to figure this out. There does exist a second derivative test for constrained minimization problems, but it falls outside the scope of this course.

12.7. A three variable example. *Find the largest value of* $x+y+z$ *on the sphere with equation* $x^2 + y^2 + z^2 = 1$.

Solution: We must maximize $f(x, y, z) = x + y + z$ with constraint $g(x, y, z) = x^2 + y^2 + z^2 = 1$.

Lagrange's method says that the minimum and maximum either occur at a point (x_0, y_0, z_0) with $\vec{\nabla} g(x_0, y_0, z_0) = \vec{0}$, or else at a point which satisfies Lagrange's equations. The gradient of g is

$$\vec{\nabla} g(x, y, z) = \begin{pmatrix} 2x \\ 2y \\ 2z \end{pmatrix},$$

and the only point where $\vec{\nabla} g = \vec{0}$ is at the origin. The origin does not satisfy the constraint $g(x, y, z) = 1$, so we can rule out the possibility of the maximum or minimum occurring at a point with $\vec{\nabla} g = \vec{0}$.

This leads us to consider the Lagrange multiplier equations, which are

$$\begin{aligned} 1 &= \lambda \cdot 2x && (f_x = \lambda g_x) \\ 1 &= \lambda \cdot 2y && (f_y = \lambda g_y) \\ 1 &= \lambda \cdot 2z && (f_z = \lambda g_z) \\ x^2 + y^2 + z^2 &= 1 && (g(x, y, z) = C) \end{aligned}$$

Solve the first three equations for x, y, z and substitute the result in the constraint, and you find

$$\frac{1}{4\lambda^2} + \frac{1}{4\lambda^2} + \frac{1}{4\lambda^2} = 1 \implies \frac{3}{4\lambda^2} = 1 \implies \lambda = \pm\tfrac{1}{2}\sqrt{3}.$$

We therefore find two points on the ellipsoid,

$$(x, y, z) = \left(\tfrac{1}{3}\sqrt{3}, \tfrac{1}{3}\sqrt{3}, \tfrac{1}{3}\sqrt{3}\right) \text{ and } (x, y, z) = \left(-\tfrac{1}{3}\sqrt{3}, -\tfrac{1}{3}\sqrt{3}, -\tfrac{1}{3}\sqrt{3}\right)$$

By computing the function values you find that the first point maximizes $x + y + z$, and the second minimizes $x + y + z$.

13. Problems

101. Minimize xy subject to the constraint

$$x^2 + \tfrac{1}{4}y^2 = 1.$$

Draw the constraint set.

102. A six-sided rectangular box is to hold $1/2$ cubic meter. Which shape should the box be to minimize surface area?

(i) Find the solution without using Lagrange's method.

(ii) Use Lagrange multipliers to solve this problem.

103. Using the methods of this section, find the shortest distance from the origin to the plane $x + y + z = 10$. (suggestion: instead of minimizing the distance, you can also minimize the square of the distance.)

104. Use Lagrange multipliers to find the largest and smallest values of $f(x, y) = x$ under the constraint $g(x, y) = y^2 - x^3 + x^4 = 0$.

105. **(i)** Using Lagrange multipliers, find the shortest distance from the point $(2, 1, 4)$ to the plane $2x - y + 3z = 1$.

(ii) Using Lagrange multipliers, find the shortest distance from the point (x_0, y_0, z_0) to the plane $ax + by + cz = d$.

106. **(i)** Find the shortest distance from the point $(0, b)$ to the parabola $y = x^2$, using Lagrange multipliers.

(ii) Find the shortest distance from the point $(0, 0, b)$ to the paraboloid $z = x^2 + y^2$.

(iii) Find the shortest distance from the point $(0, 0, b)$ to the paraboloid $z = x^2 + \frac{1}{4}y^2$.

107. Find the volume of the largest rectangular box with edges parallel to the axes that can be inscribed in the ellipsoid

$$2x^2 + 72y^2 + 18z^2 = 288.$$

108. A six-sided rectangular box is to hold $1/2$ cubic meter; what shape should the box be to minimize surface area?

109. A circular cone has height H, and its base has radius R. If the volume of the cone is fixed, then which ratio of radius to height $(R : H)$ minimizes the surface area of the cone? (The area of the cone is $A = \pi R\sqrt{R^2 + H^2}$, its volume is $V = \frac{1}{3}\pi R^2 H$, and instead of minimizing the area you could also minimize the square of the area.)

110. The post office will accept packages whose combined length and girth are at most 130 inches (girth is the maximum distance around the package perpendicular to the length). What is the largest volume that can be sent in a rectangular box?

111. The bottom of a rectangular box costs twice as much per unit area as the sides and top. Find the shape for a given volume that will minimize cost.

112. Find all points on the surface

$$xy - z^2 + 1 = 0$$

that are closest to the origin.

113. The material for the bottom of an aquarium costs half as much as the high strength glass for the four sides. Find the shape of the cheapest aquarium that hold a given volume V.

114. The plane $x-y+z=2$ intersects the cylinder $x^2+y^2=4$ in an ellipse. Find the points on the ellipse closest to and farthest from the origin. (Hint: on the plane you always have $z=2-x+y$, so you can eliminate z and make this a problem about functions of (x,y) only.)

CHAPTER 4

Integrals

1. Overview

1.1. The one variable integral. To begin, let's quickly recall how the integral of a function of one variable is defined. Given a function $y = f(x)$ and an interval $[a, b]$, we choose a ***partition*** of the interval $[a, b]$, meaning we split the interval $[a, b]$ into shorter intervals $[x_0, x_1]$, $[x_1, x_2]$, $\ldots$, $[x_{N-1}, x_N]$, where $a = x_0 < x_1 < \cdots < x_N = b$, and we choose one sample number ξ_k from each interval $[x_{k-1}, x_k]$. From these ingredients we compute the ***Riemann sum***

$$R = f(\xi_1)\Delta x_1 + \cdots + f(\xi_N)\Delta x_N = \sum_{k=1}^{N} f(\xi_k)\Delta x_k$$

where $\Delta x_k = x_k - x_{k-1}$ is the length of the k^{th} interval.

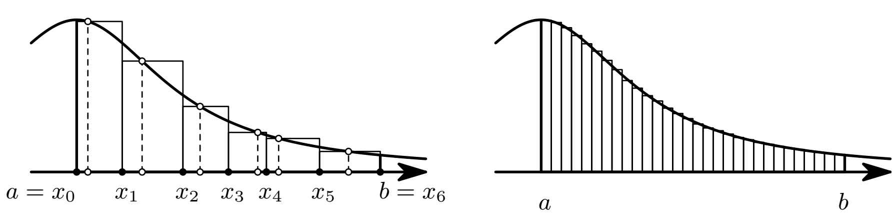


Figure 1: Riemann sums for $\int_a^b f(x)dx$ with one partition on the left, and a finer partition on the right. The dashed lines in the figure on the left indicate where the intermediate points ξ_k were chosen.

Upon making the intervals $[x_{k-1}, x_k]$ shorter (and hence choosing more partition intervals), the resulting Riemann sums get closer to one particular value, which we call the integral of the function $f(x)$ over the interval $[a, b]$:

$$\int_a^b f(x)dx = \lim_{\substack{\text{“as the partition}\\ \text{gets finer”}}} f(\xi_1)\Delta x_1 + \cdots + f(\xi_N)\Delta x_N.$$

The individual terms in the Riemann sum are areas of narrow rectangles in the figure, and following this lead one finds that the integral is the area between the graph of $y = f(x)$ and the x-axis (at least in the case that f is a positive function, so that its graph lies above the x-axis.)

1.2. Generalizing the one variable integral. While there is essentially only one kind of integral in single variable calculus, there are many different ways of integrating a function of several variables. All these different notions of "integral" are bound together by one idea, namely that they all satisfy the following rough description.

In each of the notions of integral you have these ingredients:

- a ***domain***. Depending on the kind of integral, this can be a region in the plane, in space, a plane curve, a space curve, or even some surface in three dimensional space.
- a ***function*** which is defined on the domain
- a way of measuring the "***size***" of pieces of the domain

To define the integral you "partition" the region, i.e. you divide it into lots of little pieces. Given any such partition of the region into smaller pieces, you then form the following "Riemann sum"

$$\sum_{\substack{\text{pieces in the}\\ \text{partition}}} \begin{pmatrix} f \text{ at sample point} \\ \text{in piece \#k} \end{pmatrix} \times \left\{\text{Size of piece \#k}\right\}$$

This gives you a number for each way of partitioning the region. As you make the partition finer, i.e. as you choose more, smaller, pieces, the Riemann sums tend to get closer to one particular number, which is called the integral of the function. In short, the integral is the limit of the Riemann sums you find as you take finer and finer partitions:

$$\int_{\text{some region}} f(x)\, dx = \lim_{\substack{\text{as the}\\ \text{partition}\\ \text{gets finer}}} \sum_{\substack{\text{pieces in the}\\ \text{partition}}} \begin{pmatrix} f \text{ at sample point} \\ \text{in piece \#k} \end{pmatrix} \times \left\{\text{Size of piece \#k}\right\}$$

Depending on what kind of function we have, and what kind of region the function is defined on, and also how we decide to measure the size of the small pieces in the partition, this process can led to many different kinds of integrals. The integrals we will meet in this chapter are ***double integrals***, ***triple integrals***, ***line integrals***, and ***surface integrals***. See Table 1.

2. Double Integrals

Let $z = f(x, y)$ be a function of two variables defined on some region D in the plane. The ***double integral of*** f ***over*** D is defined in terms of Riemann sums, following the general scheme described in the previous section. To form a Riemann sum you first need a partition of the region D into smaller regions $D_1, \ldots, D_N$, and you need to choose a sample point (x_k, y_k) from each region D_k. If ΔA_k is the area of region D_k, then the Riemann sum corresponding to the partition $D_1, \cdots, D_N$ and the choice of sample points $(x_1, y_1), \ldots, (x_N, y_N)$ is

$$R = f(x_1, y_1)\Delta A_1 + \cdots + f(x_N, y_N)\Delta A_N = \sum_{k=1}^{N} f(x_k, y_k)\, \Delta A_k. \tag{60}$$

If the partition is "sufficiently fine" then this Riemann sum will in many cases be close to one particular number, which we will call the integral of the function f over the region D. Thus

$$\iint_D f(x, y)\, dA = \lim_{\substack{\text{as the partition}\\ \text{"gets finer\&finer"}}} \sum_{k=1}^{N} f(x_k, y_k)\, \Delta A_k. \tag{61}$$

To make this more precise one has to resort to εs and δs, which results in the following definition

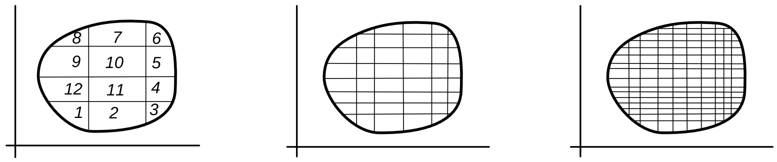


Figure 2: On the left: a region in the plane with some partition. Many pieces of the partition are rectangles. This is a common choice, but the pieces don't have to be rectangles: here the pieces that touch the boundary of the domain have at least one curved edge. **On the right:** the same region with two finer partitions.

2.1. Definition. *If for every $\varepsilon > 0$ there is a $\delta > 0$ such that the Riemann sum corresponding to any partition of the region D into smaller pieces $D_1, \ldots, D_N$ whose pieces have diameter no more than δ satisfies*

$$\left| I - \sum_{k=1}^{N} f(x_k, y_k)\ \Delta A_k \right| < \varepsilon$$

then we say that

$$\iint_D f(x,y)\ dA = I.$$

On one hand it can be shown in many cases that that the integral of a function exists according to the above definition. On the other hand the ε-δ definition is neither a practical method of computing such integrals, nor does it provide an easy intuitive understanding of the properties of the integral. Therefore, we will stick to the less precise definition (61) in this course.

Kind of integral	**Domain**	**Typical piece of partition**	**Size of piece**
"Good old 221 Integral" $\int_a^b f(x)\ dx$	interval $a \le x \le b$	small subinterval (x_{k-1}, x_k)	length of subinterval $\Delta x_k = x_k - x_{k-1}$
Multiple integral $\iint_D f(x,y) dA$	region in the plane	tiny sub domain	area ΔA of tiny sub domain
Multiple integral $\iiint_D f(x,y,z) dV$	region in space	tiny sub domain	volume ΔV of tiny sub domain
Line integral $\int_{\mathcal{C}} f(x,y)\ ds$	curve in the plane	short sub arc of curve	length Δs of the sub arc
Line integral $\int_{\mathcal{C}} f(x,y,z)\ ds$	curve in space	short sub arc of curve	length Δs of the sub arc
Surface integral $\iint_{\mathcal{S}} f(x,y,z)\ dA$	surface in space	small patch on the surface	area ΔA of the patch

Table 1: A list of the different kinds of integrals that we will encounter in math 234.

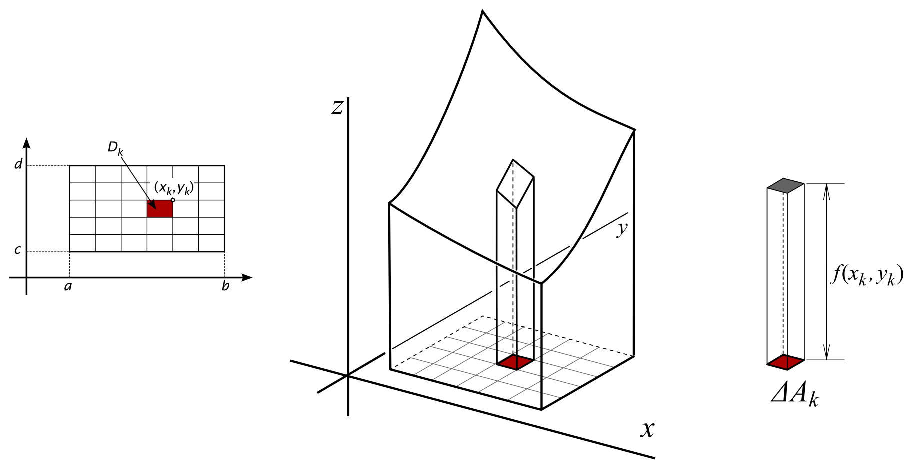


Figure 3: On the left: the domain of the function f partitioned into 6×5 pieces, each with the same width Δx and height Δy. To form a Riemann sum we have to choose one sample point (x_k, y_k) in each piece D_k of the partition. Below we will always choose the upper-right-hand corner of the rectangle to be the sample point. **On the right:** Any piece in the partition corresponds to a term in the Riemann sum of the form $f(x_k, y_k)\Delta A_k$. This is the volume of a block of height $f(x_k, y_k)$, and base D_k, which is approximately the volume of the region under the graph of f and above the piece D_k. Adding all these volumes together you see that a Riemann sum approximates the total volume between the graph and the region D.

2.2. The integral is the volume under the graph, when $f \geq 0$. If the function f is positive, then its graph lies above the xy-plane, and there is a simple interpretation of the integral, namely

$$\iint_D f(x, y)\, dA = \text{Volume of } \mathcal{R},$$

where $\mathcal{R}$ is "the region under the graph of f above the domain D" – in symbols,

$$\mathcal{R} = \big\{(x, y, z) : (x, y) \text{ lies in } D, \text{ and } 0 \leq z \leq f(x, y)\big\}. \tag{62}$$

To see why this is so, let's imagine that we have a positive function $z = f(x, y)$ defined on some region D in the xy-plane, and let's try to compute the integral $\iint_D f(x, y)dA$ "geometrically." To compute the integral we begin by finely partitioning the region D into smaller regions D_1, D_2, $\ldots$, D_N (see Figure 3 on the left where the small pieces were themselves chosen to be rectangles). We also choose one "sample point" (x_k, y_k) in each region D_k. The Riemann sum we get this way is

$$R = f(x_1, y_1)\Delta A_1 + \cdots + f(x_N, y_N)\Delta A_N$$

where ΔA_k is the area of D_k. The k^{th} term, $f(x_k, y_k)\Delta A_k$, is the volume of a block whose base is D_k and whose top is some point on the graph of the function above the region D_k. This volume is almost, but usually not exactly the same as the volume of the region between the graph of the function and the small region D_k in the xy-plane. The volume $f(x_k, y_k)\Delta A_k$ of the block above D_k is not exactly the same as the volume of the region under the graph because the top of the block is a piece of a horizontal plane while the graph of f will usually have a slope.

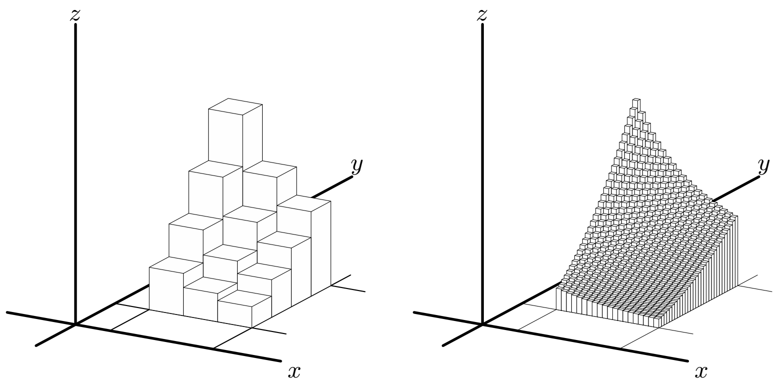


Figure 4: Approximating the region under the graph of $z = f(x, y)$ from Figure 3 by vertical blocks. The base of each block is a rectangle in a partition of the domain of f. As you choose finer and finer partitions, the region occupied by the vertical blocks gets closer to the region under the graph of f.

The total Riemann sum is therefore the sum of the volumes of such blocks, (see Figure 4) and this will approximate the volume between the graph of f and the domain of integration D. The finer the partition, the better the approximation so we can conclude[1] that the limit of the Riemann sums is the volume under the graph, to wit, the volume of the region $\mathcal{R}$ defined in (62).

2.3. How to compute a double integral. So far, we have a definition for the double integral $\iint_D f(x,y)dA$, and an interpretation of the integral as "volume under the graph of f." What is missing is a method of actually computing the integral. In this section we'll see how you can compute a double integral by doing two one-variable integrals.

Let's take another look at the integral of the function f over the rectangle

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

from the previous section.

We again partition D into smaller rectangles, as in Figure 3, but instead of just counting them and numbering the pieces 1, 2, ..., N, we can use the fact that the smaller rectangles appear in rows and columns. If we take N rectangles in the x direction, and M in the y direction, then the smaller rectangles will measure Δx by Δy, where

$$\Delta x = \frac{b-a}{N}, \qquad \Delta y = \frac{d-c}{M}.$$

We let (x_k, y_l) be the upper-right-hand corner of the rectangle in the k^{th} column from the left, and the l^{th} row from below. Then

$$x_k = a + k\Delta x, \qquad y_l = c + l\Delta y. \tag{63}$$

---

[1]This is not a very precise "proof," but to prove that the limit of Riemann sums exists you would have to use $\varepsilon\&\delta$ arguments.

The Riemann sum corresponding to this partition and choice of sample points (x_k, y_l) is

$$(64)\qquad R = \sum f(x_k, y_l)\Delta x\Delta y = \begin{array}{c} f(x_1,y_1)\Delta x\Delta y + \cdots + f(x_N,y_1)\Delta x\Delta y \\ + f(x_1,y_2)\Delta x\Delta y + \cdots + f(x_N,y_2)\Delta x\Delta y \\ \vdots \\ + f(x_1,y_M)\Delta x\Delta y + \cdots + f(x_N,y_M)\Delta x\Delta y \end{array}$$

Since we are choosing the upper-right-hand corner of each rectangle as sample point in that rectangle, the sample point for the rectangle at the top-right is (x_N, y_M). (See Figure 3 on the left.) Therefore, in this summation k can have any value with $1 \le k \le N$ and l can be any integer with $1 \le l \le M$.

The term corresponding to rectangle (k, l) represents the volume of a block whose height is $f(x_k, y_l)$ and whose base is a $\Delta x \times \Delta y$ rectangle. Together these blocks approximate the region between the graph of the function and the xy-plane.

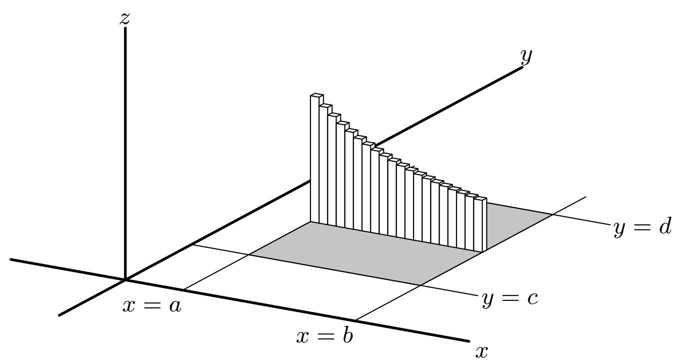


Figure 5: This picture shows the blocks corresponding to all those terms in the Riemann sum R from equation (64) in which $y = y_k$. These terms $\{f(x_1, y_k)\Delta x + \cdots + f(x_N, y_k)\Delta x\}\Delta y$ give you the total volume of one row of "matchsticks" from Figure 4. In this sum y is frozen at the value $y = y_k$, so you can think of $f(x_1, y_k)\Delta x + \cdots + f(x_N, y_k)\Delta x$ as a Riemann sum for the integral $\int_a^b f(x, y_k)\,dx$.

Consider the terms on the k^{th} row in equation (64); after factoring out Δy you get

$$\text{row } \#k \text{ of } (64) = \Delta y \Big\{f(x_1, y_k)\Delta x + f(x_2, y_k)\Delta x + \cdots + f(x_N, y_k)\Delta x\Big\}.$$

Note that in this sum the function is always evaluated at the same value of y, namely y_k. The sum between braces $\{\cdots\}$ is actually a Riemann sum for the one-variable integral

$$I = \int_a^b f(x, y_k)dx$$

in which we treat $f(x, y_k)$ as a function of x only and consider the variable y to be frozen at $y = y_k$. The value of this integral will of course depend the value at which y is frozen, so it is better to write

$$I(y) = \int_a^b f(x, y)dx.$$

With this notation we find that

$$\text{row } \#k \text{ of } (64) \approx \Delta y \times \big\{I(y_k)\big\} = I(y_k)\Delta y.$$

To find the value of the Riemann sum which approximates the double integral $\iint_D f(x, y)dA$ we add the rows in (64) and find

$$R \approx I(y_1)\Delta y + I(y_2)\Delta y + \cdots + I(y_M)\Delta y.$$

The sum on the right is again a Riemann sum for a one variable integral, namely, $\int_c^d I(y)dy$. Therefore we find that

$$R \approx \int_c^d I(y)dy$$

If we now take the limit in which we let the size of the pieces in the partition go to zero, then it can be shown (with quite a bit of effort) that the approximation above gets better, and that one has

$$\iint_D f(x,y)dA = \int_c^d I(y)dy,$$

and hence, remembering the definition of $I(y)$, we have found the following method of computing a double integral.

2.4. Theorem. *If $f(x,y)$ is a function defined on a rectangle*

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

then the double integral of f over D is given by

$$\iint_D f(x,y)dA = \int_c^d \Big\{ \int_a^b f(x,y)dx \Big\} dy. \tag{65}$$

One can also first integrate with respect to y and then x, so that

$$\iint_D f(x,y)dA = \int_a^b \Big\{ \int_c^d f(x,y)dy \Big\} dx. \tag{66}$$

The second way of computing the double integral $\iint_D f(x,y)\, dA$, i.e. equation (66), follows by the same reasoning that led us to (65), except in (64) one groups the terms by columns rather than rows.

To compute the right hand side in this equation we have to compute two one-variable integrals. The expression

$$\int_c^d \Big\{ \int_a^b f(x,y)dx \Big\} dy = \int_c^d \int_a^b f(x,y)\, dx\, dy$$

is called an ***iterated integral***.

2.5. Example: the volume under the graph of the paraboloid $z = x^2 + y^2$ above the square $Q = \{(x,y) : 0 \le x \le 1, 0 \le y \le 1\}$. The double integral we have to compute is

$$\text{Volume} = \iint_Q (x^2+y^2)dA$$

and to compute it we write it as an iterated integral

$$\iint_Q (x^2+y^2)dA = \int_0^1 \Big\{ \int_0^1 (x^2+y^2)dx \Big\} dy$$

In the inner integral the variable y is frozen, so to compute the inner integral, you simply treat y as a constant, and integrate with respect to x. You get

$$\int_0^1 (x^2+y^2)dx = \left[\tfrac{1}{3}x^3 + y^2x\right]_{x=0}^1 = \tfrac{1}{3} + y^2.$$

(This is $I(y)$ in the notation of the previous section.)

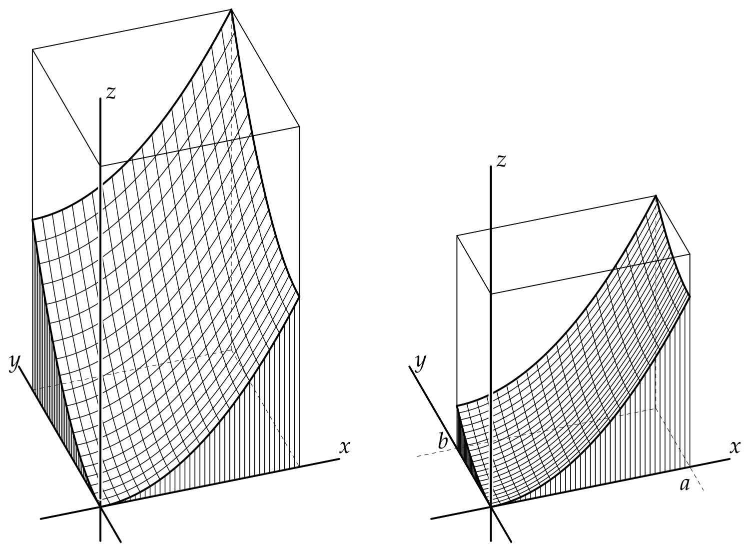


Figure 6: The graph of $z = x^2 + y^2$ above the unit square Q on the left, and rectangle $\{(x, y) : 0 \le x \le a$ and $0 \le y \le b\}$, on the right, together with the surrounding block. *What fraction of the volume of the block lies below the graph?*

To get the double integral we must still do the outer integral:

$$\begin{aligned}\iint_Q (x^2 + y^2)dA &= \int_0^1 \Big\{ \int_0^1 (x^2 + y^2)dx \Big\} dy \\ &= \int_0^1 \left(\tfrac{1}{3} + y^2\right) dy \\ &= \left[\tfrac{1}{3}y + \tfrac{1}{3}y^3\right]_0^1 \\ &= \tfrac{1}{3} + \tfrac{1}{3} = \tfrac{2}{3}.\end{aligned}$$

Since the surrounding block (Figure 6) is a $1 \times 1 \times 2$ block, its volume is 2, and the region under the graph occupies exactly one third of the whole block.

To compute the volume of the region under the graph of the same function above the rectangle $\{(x, y) : 0 \le x \le a, 0 \le y \le b\}$ one can compute either of the iterated integrals

$$\int_0^a \int_0^b \left(x^2 + y^2\right) dy\, dx \text{ or } \int_0^b \int_0^a \left(x^2 + y^2\right) dx\, dy.$$

2.6. Double integrals when the domain is not a rectangle. We have seen how to compute a double integral when the domain is a rectangle. The reasoning which led us from a double integral to an iterated integral also works for non rectangular domains, provided they are not too complicated. Suppose you want to compute $\iint_D f(x, y)dA$ where the domain D is the region caught between the graphs of two functions:

$$D = \big\{(x, y) : a \le x \le b,\ f(x) \le y \le g(x)\big\}.$$

We again partition the region by cutting it along many vertical lines $x = x_1$, $x = x_2$, $\ldots$, $x = x_N$, and many horizontal lines $y = y_1$, $\ldots$, $y = y_M$. Most of the pieces of the partition will be rectangles, but those which overlap with the boundary of the region D may have curved edges. See Figures 7 and 8.

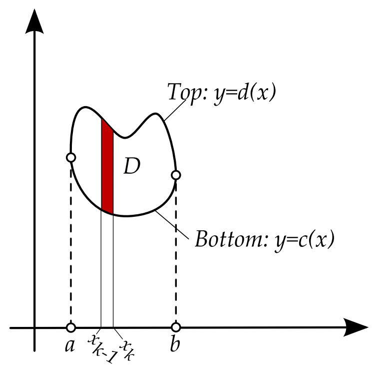


Figure 7: The region between the graphs of $y = f(x)$ and $y = g(x)$.

This time, all the terms in a Riemann sum corresponding to one particular strip $x_{k-1} \leq x \leq x_k$ add up to a Riemann sum for an integral over the y variable,

$$\int_{c(x)}^{d(x)} f(x_k, y)\, dy \;\times\; \Delta x,$$

and adding all these we get the iterated integral

$$\iint_D f(x,y)\, dA = \int_a^b \left\{ \int_{c(x)}^{d(x)} f(x,y)\, dy \right\} dx. \tag{67}$$

2.7. An example–the parabolic office building. Consider the region under the graph of $f(x, y) = x + y$, above the domain

$$D = \left\{(x,y) : 0 \leq x \leq 1, (1-x)^2 \leq y \leq 1\right\}.$$

The volume of this region is given by

$$V = \iint_D (x+y) dA.$$

We can compute this volume by finding the following iterated integral

$$V = \int_{x=0}^{1} \int_{(1-x)^2}^{1} (x+y)\, dy\, dx. \tag{68}$$

Alternatively, the region D can also be described as

$$D = \{(x,y) : 0 \leq y \leq 1,\ 1 - \sqrt{y} \leq x \leq 1\}.$$

This leads to the following iterated integral for the volume

$$V = \int_{y=0}^{1} \int_{1-\sqrt{y}}^{1} (x+y)\, dx\, dy. \tag{69}$$

Both iterated integrals should give the same answer. Let's compute the first one:

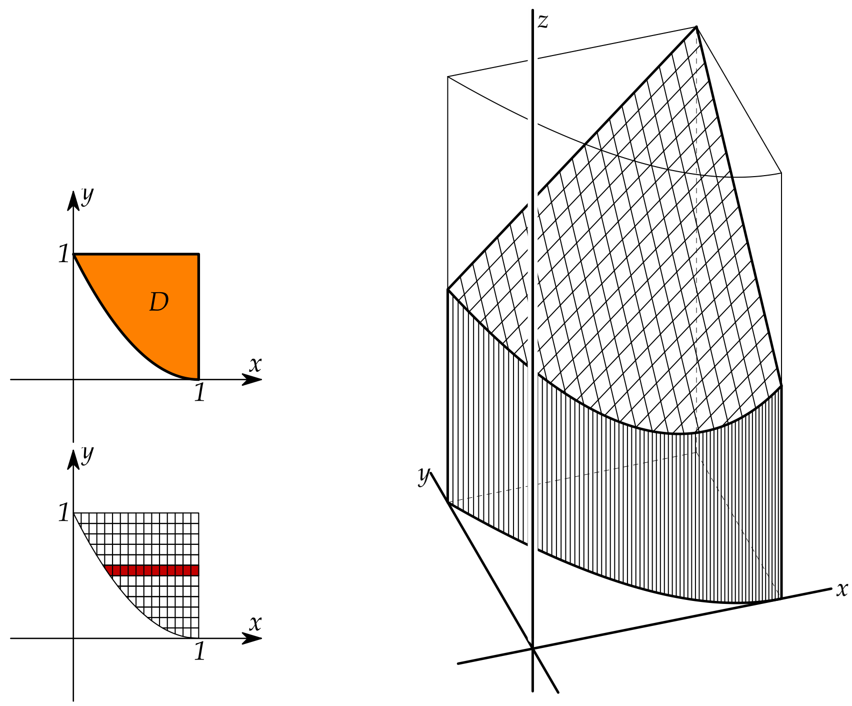
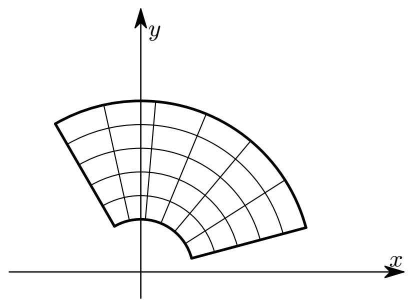


Figure 8: On the left: the domain of integration, a partition, and all pieces in the partition corresponding to one value of y. **On the right:** The "parabolic office building," being the region whose volume is computed in example 2.7

$$
\begin{aligned}
V &= \int_0^1 \int_{(1-x)^2}^1 (x+y)\,dy\,dx \\
&= \int_0^1 \left[\tfrac{1}{2}xy + \tfrac{1}{2}y^2\right]_{(1-x)^2}^1 \,dx \\
&= \int_0^1 \left[x\bigl(1-(1-x)^2\bigr) + \tfrac{1}{2}\bigl(1^2-(1-x)^4\bigr)\right] \,dx \\
&= \int_0^1 \left[2x^2 - x^3 + \tfrac{1}{2}\bigl(4x - 6x^2 + 4x^3 - x^4\bigr)\right] \,dx \\
&= \int_0^1 \left[2x^2 - x^3 + 2x - 3x^2 + 2x^3 - \tfrac{1}{2}x^4\right] \,dx \\
&= \tfrac{2}{3} - \tfrac{1}{4} + 1 - 1 + 2\times\tfrac{1}{4} - \tfrac{1}{2}\times\tfrac{1}{5} \\
&= \tfrac{16}{15}.
\end{aligned}
$$

Note that even though the function we are integrated is very simple (it's just $x+y$) the integral can still become complicated because of the shape of the domain D over which we are integrating.

2.8. Double integrals in Polar Coordinates. Sometimes Cartesian coordinates are just not the best choice. For instance, a disc or radius R, centered at the origin, is very easy to describe in polar coordinates as "all points with $r \le R$." In Cartesian coordinates you need Pythagoras, and you have to say "all points with $x^2 + y^2 \le R^2$." In the same



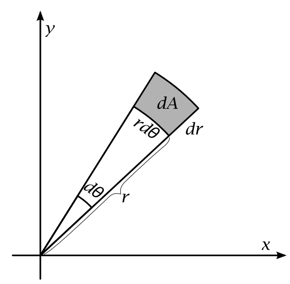


Figure 9: Left: A "polar rectangle" and a partition by lines θ =constant (the spokes) and r =constant (the arcs). **Right:** The area of a small piece of such a partition is approximately $\Delta A \approx \Delta r \times r\Delta\theta$.

spirit a "polar rectangle" is a domain of the form

$$R = \{\text{all points with } \theta_0 \le \theta \le \theta_1, r_0 \le r \le r_1\}.$$

See Figure 9 (on the left). There is a very natural way of partitioning such a region into many smaller regions, by cutting the region along curves of constant r (arcs centered at the origin) or constant θ (rays emanating from the origin). If the partition is sufficiently fine, then the pieces in the partition will almost be real Cartesian rectangles, with sides $r\Delta\theta$ and Δr ($\Delta\theta$ being the angle between adjacent rays, and Δr being the difference in radius between two consecutive arcs). The area of such a small partition piece is therefore $\Delta A \approx r\Delta\theta \times \Delta r$, and one arrives at the following formula for the integral of a function of a polar rectangle

$$\iint\limits_R f(x, y)\, dA = \int_{r_0}^{r_1} \int_{\theta_0}^{\theta_1} F(r, \theta)\, r d\theta\, dr = \int_{\theta_0}^{\theta_1} \int_{r_0}^{r_1} F(r, \theta)\, r dr\, d\theta. \tag{70}$$

Here $F(r, \theta) = f(r \cos\theta, r \sin\theta)$ is the function $f(x, y)$ written in polar coordinates.[2]

There is a similar formula for more complicated domains. If a domain can be described in polar coordinates by

$$D = \{\text{all points with } \alpha \le \theta \le \beta, a(x) \le r \le b(x)\}$$

and if you want to integrate a function $z = f(x, y)$ of this domain, then you can again partition the domain D into many small pieces which are bounded by circular arcs centered at the origin, and lines emanating from the origin. The area of a small piece in the partition is once again given by $\Delta A \approx \Delta r \times r\Delta\theta$, and therefore the integral of f over D is

$$\iint\limits_D f(x, y)\, dA = \int_\alpha^\beta \int_{a(\theta)}^{b(\theta)} F(r, \theta)\, r\, dr\, d\theta. \tag{71}$$

[2]It is very common to use the same letter f for both functions, i.e. to write $f(x, y)$ for f as a function of Cartesian coordinates, and also $f(r, \theta)$ for the same function but written in Polar coordinates. This begs the question of what $f(0.3, 1.24)$ means – are $(0.3, 1.24)$ the Polar or Cartesian coordinates of the point at which f is to be evaluated? To avoid this kind of ambiguity we will try to use different letters for the same quantity regarded as a function of Cartesian coordinates, and of Polar coordinates.

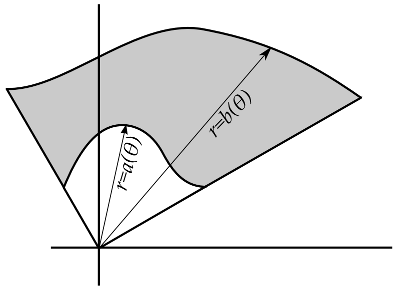


Figure 10: The gray region is the region between the polar graphs $r = a(\theta)$ and $r = b(\theta)$.

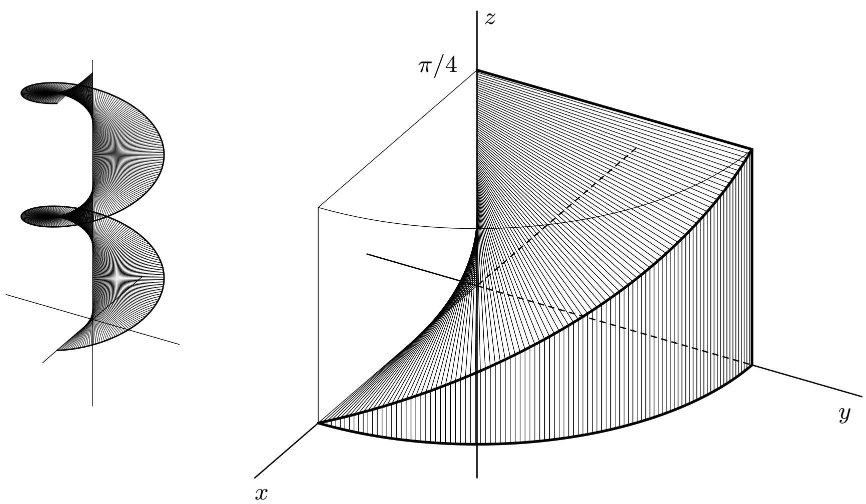


Figure 11: The graph of the function $z = a\theta$ in polar coordinates is called the *helicoid*. **Left:** two turns of a helicoid. **Right:** one quarter turn of a helicoid with $a = \frac{1}{2}$ is show. The volume under the helicoid is given by a double integral which is best computed using polar coordinates. Which fraction of the volume in the surrounding quarter cylinder lies beneath the helicoid?

2.9. Example: the volume under a quarter turn of a helicoid. A ***helicoid*** is the surface which in polar coordinates is given by

$$z = a\theta$$

where $a > 0$ is some constant. The polar angle is multivalued, because at any point in the plane the polar angle is only determined up to multiple of 2π, except at the origin, where the polar angle θ isn't defined at all. If you take all the possible different values the polar angle of a point (x, y) can have the graph of $z = \theta$ looks like the picture on the left in Figure 11.

If we choose the constant $a = \frac{1}{2}$, and take the first quarter turn of this surface, on which $0 \le \theta \le \frac{1}{2}\pi$, then we get the picture on the right in Figure 11. In that drawing we have only included the part with $0 \le r \le 1$. To compute the volume of the region under the quarter helicoid using Cartesian coordinates, you would have to compute this integral

$$V = \int_0^1 \int_0^{\sqrt{1-x^2}} \tfrac{1}{2} \arctan \frac{y}{x}\, dy\, dx.$$

(Try to set up this integral yourself!)

In Polar coordinates things are easier. The domain is a polar rectangle,

$$0 \le r \le 1, \quad 0 \le \theta \le \tfrac{1}{2}\pi,$$

and the function is very simple,

$$F(r, \theta) = \tfrac{1}{2}\theta.$$

The double integral that represents the volume is therefore

$$\begin{aligned} V &= \iint\limits_D \tfrac{1}{2}\theta\, dA \\ &= \int_0^1 \int_0^{\pi/2} \tfrac{1}{2}\theta r\, d\theta\, dr \\ &= \frac{\pi^2}{32}. \end{aligned}$$

3. Problems

115. Compute these iterated integrals:

(i) $\displaystyle\int_0^1 \int_0^4 x\, dy\, dx$ **(ii)** $\displaystyle\int_0^1 \int_0^4 x\, dx\, dy$ **(iii)** $\displaystyle\int_{-1}^1 \int_0^{x^2} dy\, dx$

(iv) $\displaystyle\int_0^1 \int_0^y \frac{\sin y}{y}\, dx\, dy$ **(v)** $\displaystyle\int_0^1 \int_0^\theta \frac{\sin\theta}{\theta}\, dr\, d\theta$ **(vi)** $\displaystyle\int_0^1 \int_0^{\sqrt{1-x^2}} dy\, dx$

116. What is wrong with the iterated integral $\displaystyle\int_x^1 \Big\{ \int_0^1 \sin(\pi x) dx \Big\} dy$?

Is the answer a number – does it depend on x or y?

117. Is the following true or false?

For any two functions $f(x)$ and $g(y)$ one has

$$\int_0^1 \int_0^2 f(x)g(y)\, dx\, dy = \int_0^1 f(x)\, dx \;\times\; \int_0^2 g(y)\, dy\,.$$

Explain your answer (if you claim "true" give a proof, if you claim "false" give a counterexample.)

118. Answer the question posed in Figure 6.

119. Compute the following double integrals. In each case sketch the domain of integration and show which iterated integral you must compute to find the given double integral.

(**i**) $\iint_D (1+x)\,dA$ where $D = \{(x,y) : 0 \le x \le 2, 0 \le y \le 4\}$.

(**ii**) Compute $\iint_D (x+y)\,dA$ where $D = \{(x,y) : |x| \le 1, 0 \le y \le 4\}$

(**iii**) Compute $\iint_D xy\,dA$ where $D = \{(x,y) : 0 \le x \le y, 1 \le y \le 2\}$.

(**iv**) Compute $\iint_D dA$ where $D = \left\{(x,y) : \frac{1}{2}y^2 \le x \le \sqrt{y}, 0 \le y \le 1\right\}$.

(**v**) Compute $\iint_D \frac{x^2}{y^2}\,dA$ where $D = \{(x,y) : 1 \le x \le 2, 1 \le y \le x\}$.

(**vi**) Compute $\iint_D \frac{y}{e^x}\,dA$ where $D = \left\{(x,y) : 0 \le x \le 1, 0 \le y \le x^2\right\}$.

(**vii**) Compute $\iint_D x\cos y\,dA$ where $D = \left\{(x,y) : 0 \le x \le \sqrt{\pi/2}, 0 \le y \le x^2\right\}$.

(**viii**) Compute: $\iint_D \sqrt{x^3+1}\,dA$ where $D = \{(x,y) : 0 \le y \le 1, \sqrt{y} \le x \le 1\}$.

(**ix**) Compute: $\iint_D y\sin(x^2)\,dA$ where $D = \left\{(x,y) : 0 \le y \le 1, y^2 \le x \le 1\right\}$.

(**x**) Compute: $\iint_D x\sqrt{1+y^2}\,dA$ where $D = \left\{(x,y) : 0 \le x \le 1, x^2 \le y \le 1\right\}$.

(**xi**) Compute: $\iint_D \frac{2}{\sqrt{1-x^2}}\,dA$ where D is the triangle bounded by the y axis, the line $y=1$ and the line $y=x$.

120. Find the volumes of the following regions by computing a double integral.

(**i**) the region bounded by $z = x^2+y^2$ and $z=4$.

(**ii**) the region in the first octant bounded by $y^2 = 4-x$ and $y=2z$.

(**iii**) the region in the first octant bounded by $y^2 = 4x$, $2x+y=4$, $z=y$, and $y=0$.

(**iv**) the region in the first octant bounded by $x+y+z=9$, $2x+3y=18$, and $x+3y=9$.

(**v**) the region in the first octant bounded by $x^2+y^2=a^2$ and $z=x+y$.

(**vi**) the region bounded by $4x^2+y^2=4z$ and $z=2$.

(**vii**) the region bounded by $z=x^2+y^2$ and $z=y$.

121. The average value of a function $f(x,y)$over a domain D is by definition

$$\text{average } f \text{ over } D = \frac{\iint_D f(x,y)\,dA}{\text{area of } D}$$

Find the average value of $f(x,y) = e^y\sqrt{x+e^y}$ on the rectangle with vertices $(0,0)$, $(4,0)$, $(4,1)$ and $(0,1)$.

122. Suppose $f(x)$ is a positive function defined on an interval $a \le x \le b$. Let A be the area under the graph of $y=f(x)$, $(a \le c \le b)$, and let B be the area under the graph of $y=f(x)^2$ $(a \le c \le b)$

(**i**) Compute $\int_a^b \int_0^{f(x)} dydx$.

(**ii**) Compute $\int_a^b \int_0^{f(x)} ydydx$.

123. Let V be the volume under the graph of the function $z = \frac{2xy}{x^2+y^2}$, above the region

$$D = \left\{(x,y) : x \ge 0,\ y \ge 0,\ x^2+y^2 \le 1\right\}.$$

(See chapter 1 for a picture of the graph of this function.)

(**i**) Write an iterated integral for the volume V, using Cartesian coordinates. (You don't have to compute the integral you get.)

(**ii**) Compute V using polar coordinates.

124. Let V be the volume under the graph of $z = xy$ above the domain

$$D = \left\{(x, y) : x \geq 0,\ y \geq 0,\ x^2 + y^2 \leq 4\right\}.$$

Try to draw the region D, and the graph of $z = xy$ above D.

(**i**) Use Cartesian coordinates to compute V. (Hint: this is similar to part (i) of the previous problem, but the integral in this problem isn't as bad.)

(**ii**) Use Polar Coordinates to compute V.

Answers and Hints

(2) You should use a graphing program to produce pictures of the graphs in these problems.

(2i) $z - x^2 = 0$. Domain $\mathbb{R}^2$. Graph is a **parabolic cylinder** and consists of horizontal lines perpendicular to the xz-plane, going through the parabola $y = x^2$ in that plane.

Level sets: parallel straight lines $x = \pm\sqrt{z}$ if $z > 0$, the x axis if $z = 0$, the empty set if $z < 0$.

(2ii) $z^2 - x = 0$. Implicit function. At least two functions are defined, namely $z = \pm\sqrt{x}$. Domain: all points (x, y) with $x \geq 0$. Graph is **half a parabolic cylinder** and consists of horizontal lines perpendicular to the xz-plane, going through the parabola $z = \sqrt{x}$ (or $z = -\sqrt{x}$, depending on which function you choose) in that plane.

Level sets (assuming we choose the function $z = +\sqrt{x}$): the line $x = z^2$ if $z \geq 0$, empty set otherwise.

(2iii) $z - x^2 - y^2 = 0$. Domain is the whole plane. Graph is a paraboloid of revolution, obtained by rotating the parabola $z = x^2$ in the xz-plane around the z axis.

Level sets: circle with radius $\sqrt{z}$ for $z > 0$, the origin for $z = 0$ (note: this level set is a point rather than a curve), empty for $z < 0$.

(2iv) $z^2 - x^2 - y^2 = 0$. Implicit function. Domain all of $\mathbb{R}^2$. Possible functions are $z = \pm\sqrt{x^2 + y^2}$. Graph is the cone obtained by rotating the half line $z = x, x \geq 0$ in the xz-plane around the z axis (or the half line $z = -x, x \geq 0$, if you chose $z = -\sqrt{x^2 + y^2}$.)

Level sets (assuming we choose $z = +\sqrt{x^2 + y^2}$): circle with radius z when $z > 0$, origin when $z = 0$, empty when $z < 0$.

(2v) $xyz = 1$. Domain the whole plain with the x and y-axes removed, i.e. all points (x, y) with $xy \neq 0$. Function is $f(x, y) = \frac{1}{xy}$. For each y the graph is the hyperbola $z = 1/(yx)$ which is just the standard hyperbola $z = 1/x$ stretched vertically by a factor $1/y$. As $y \to 0$ this factor goes to ∞.

(2vi) $xy/z^2 = 1$. Implicit function. Domain first and third quadrants (all points with $xy > 0$). Functions $z = \pm\sqrt{xy}$. Cross sections with planes $y =$constant are half parabolas.

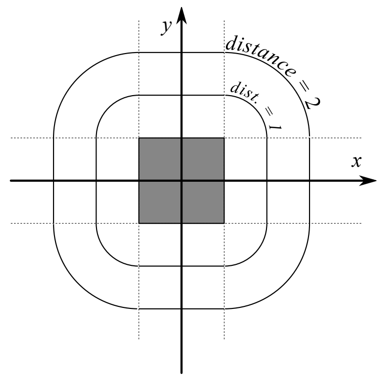
Note: Harder to see, but the surface with equation $xy = z^2$ is in fact the cone obtained by rotating the x-axis around the line $x = y$ in the xy-plane.

(4i) $(0, \frac{1}{2})$ is in the square Q, so it is the point closest to $(0, \frac{1}{2})$.
The point $(0, 1)$ on the top edge of the square is closest to $(0, 2)$.
The corner point $(1, 1)$ is closest to $(3, 4)$.

(4ii) $f(0, \frac{1}{2}) = 0$; $f(0,2) = 1$ and $f(3,4)) = \sqrt{2^2 + 3^2} = \sqrt{13}$.

(4iii) The zero set of f is the square Q.

(4iv) The level set at level -1 is empty. The others are "rounded rectangles," see this drawing, in which the square is grey, the dashed lines are given by $x = \pm 1$ or $y = \pm 1$.



(4v) The lines $x = \pm 1$ and $y = \pm 1$ divide the plane into nine regions. On each region the function is given by a different formula. Here they are:

$f(x,y)$	**if** ...
0	(x,y) in Q
$x - 1$	$x \geq 1, \|y\| \leq 1$
$y - 1$	$\|x\| \leq 1, y \geq 1$
$-x - 1$	$x \leq -1, \|y\| \leq 1$
$-y - 1$	$\|x\| \leq 1, y \leq -1$
$\sqrt{(x-1)^2 + (y-1)^2}$	$x \geq 1$ and $y \geq 1$
$\sqrt{(x-1)^2 + (y+1)^2}$	$x \geq 1$ and $y \leq -1$
$\sqrt{(x+1)^2 + (y-1)^2}$	$x \leq -1$ and $y \geq 1$
$\sqrt{(x+1)^2 + (y+1)^2}$	$x \leq -1$ & $y \leq -1$

(6) See answers to problem 2.

(7) The graph of f is obtained by taking the part of the graph of $z = g(x)$ with $x \geq 0$ and rotating it around the z-axis.

Each level set of f are circles, the origin, or is the empty set.

(8i) The two rectangular strips $-3 \leq x \leq 3, 2 \leq y < \infty$ and $-3 \leq x \leq 3, -\infty < y \leq -2$.

(8ii) By definition $\arcsin(x)$ is only defined if $-1 \leq x \leq 1$. For $\arcsin(x^2 + y^2 - 2)$ to be defined, we must therefore have $-1 \leq x^2 + y^2 - 2 \leq 1$, i.e. $1 \leq x^2 + y^2 \leq 3$.

The domain of this function is the ring-shaped region between the circles with radii 1 and $\sqrt{3}$, both centered at the origin. Circles are included in the domain.

(8iii) The way this function is written both $\sqrt{x}$ and $\sqrt{y}$ must be defined, so the domain consists off all (x,y) with $x \geq 0$ and $y \geq 0$.

(8iv) $\sqrt{xy}$ must exist, which happens for all (x, y) in the first and third quadrants (axes included.)

(8vi) The region in the plane given by $x^2+4y^2 \leq 16$, which is the region enclosed by an ellipse with major axis of length 4, along the x axis, and minor axis of length 2 along the y-axis. The ellipse is included.

(9) The level sets of the function whose graph is a cone are equally spaced circles (the level set at level c is a circle with radius c). Hence the one on the right corresponds to the cone, and the one on the left corresponds to the paraboloid.

(10i) At time t we have a line through the origin with slope $\sin t$. As time progresses this lines turns up and down, and up and down, etc.

(10ii) Same as previous problem, but twice as fast.

(10iii) At all times one sees the graph of $y = \sin x$ stretched vertically by a factor t.

(10iv) Same as previous problem, but twice as fast.

(10v) The graph of $y = \sin 2x$ stretched vertically by a factor t.

(10vi) Parabola with its minimum on the x-axis at $x = t$. So we see the parabola $y = x^2$ translating from the left to the right with constant speed 1.

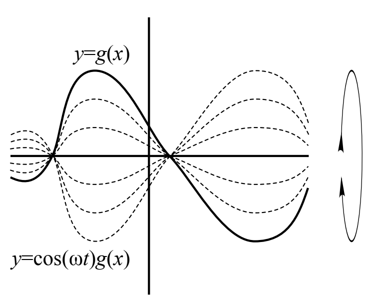
(10vii) Parabola with its minimum on the x-axis at $x = \sin t$. So we see the parabola $y = x^2$ translating back and forth horizontally every 2π time units.

(10x) At time t we see Agnesi's witch, i.e. the graph $y = a/(1 + x^2)$ with amplitude $a = 1/(1 + t^2)$. Thus we see a bump whcich starts out small at $t = -\infty$, grows to its maximal size at time $t = 0$, and then decays again, until it vanishes at $t = +\infty$.

(11) The graph of $y = g(x - a)$ is obtained from the graph of $y = g(x)$ by translating the graph of $y = g(x)$ by a units to the right.

Hence the graph of $g(x - ct)$ is the graph of $g(x)$ translated by ct units to the right. As time changes the graph of $g(x - ct)$ therefore moves with velocity c to the right.

(12) If you know the graph of a function $y = g(x)$, then you get the graph of $y = cg(x)$ by stretching the graph of g vertically by a factor c (here c is a constant.) If you allow this constant to depend on time, e.g. as in this problem by setting $c = \cos(\omega t)$, then the "movie" you get is of a version of the graph of g which is growing and shrinking vertically.



(16) The level set at level $z = c$ is the set of points which satisfy the equation

$$\frac{x^2 - y^2}{x^2 + y^2} = c.$$

You can simplify this equation by rewriting it as follows:

$$\frac{x^2 - y^2}{x^2 + y^2} = c \iff$$

$$x^2 - y^2 = cx^2 + cy^2 \iff$$

$$(1-c)x^2 = (1+c)y^2 \iff$$

$$\frac{y}{x} = \pm\sqrt{\frac{1-c}{1+c}}.$$

So we see that if $\frac{1-c}{1+c} \geq 0$ the level set consists of two **straight lines**, with the indicated slopes. This happens exactly when $-1 < c < 1$.

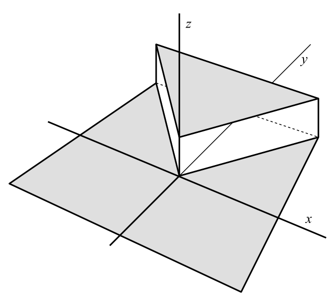
When $c = \pm 1$ we get either the equation $x^2 = 0$ or $y^2 = 0$, so that the corresponding level sets consist of either the y-axis or the x-axis.

(17) For f we have

$$\lim_{x \to 0} f(x, mx) = \lim_{x \to 0} \frac{x^2 - m^2x^2}{x^2 + m^2x^2} = \frac{1 - m^2}{1 + m^2}.$$

In fact this computation shows that f is constant on lines of the form $y = mx$, which we already found in the previous problem.

(18i) The function has been defined for all (x, y), so its domain is the whole plane. The graph looks like this, roughly:



Note that the drawing doesn't tell you what the function values are on the jump curve, $y = |x|$.

(18ii) To compute A we must take the limit as $x \to 0$ of $\lim_{y\to 0} f(x, y)$, so we must know this limit when $x \neq 0$.

For all $x \neq 0$ one has $\lim_{y\to 0} f(x, y) = 0$. Hence $A = 0$.

To find B, compute for any $y \neq 0$

$$\lim_{x\to 0} f(x, y) = \begin{cases} 1 & \text{if } y > 0 \\ 0 & \text{if } y < 0. \end{cases}$$

Hence the iterated limit

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

does not exist (limits for $y \nearrow 0$ and $y \searrow 0$ are different).

(18iii) The limit doesn't exist because one of the iterated limits A and B does not exist.

(18iv) The function is continuous at all points (x, y) except those with $y = |x|$, where the function has a jump discontinuity.

(20i)

$$\begin{aligned}&\lim_{\substack{(x,y)\to(0,0)\\ y=mx}} h(x,y)\\ &\qquad = \lim_{x\to 0} h(x, mx)\\ &\qquad = \lim_{x\to 0} \frac{x^4 - m^2x^2}{x^4 + m^2x^2}\\ &\qquad = \lim_{x\to 0} \frac{x^2 - m^2}{x^2 + m^2}.\end{aligned}$$

If $m \neq 0$ then this limit is

$$\frac{-m^2}{m^2} = -1$$

But when $m = 0$ you get

$$\lim_{x\to 0} \frac{x^2 - m^2}{x^2 + m^2} = \lim_{x\to 0} \frac{x^2}{x^2} = 1.$$

(20ii) The two iterated limits

$$\lim_{x\to 0} \lim_{y\to 0} h(x, y) = 1$$

and

$$\lim_{y\to 0} \lim_{x\to 0} h(x, y) = -1$$

are different, so the limit $\lim_{(x,y)\to(0,0)} h(x, y)$ does not exist.

(20iv)

$$\begin{aligned}&\lim_{\substack{(x,y)\to(0,0)\\ y=mx^2}} h(x,y)\\ &\qquad = \lim_{x\to 0} h(x, mx^2)\\ &\qquad = \lim_{x\to 0} \frac{x^4 - m^2x^4}{x^4 + m^2x^4}\\ &\qquad = \lim_{x\to 0} \frac{1 - m^2}{1 + m^2}\\ &\qquad = \frac{1 - m^2}{1 + m^2}.\end{aligned}$$

So the answer does depend on m, i.e. depending on which parabola $y = mx^2$ you follow to get to the origin, you get a different limit.

(22ii) $-2xy \sin(x^2y)$, $-x^2 \sin(x^2y) + 3y^2$

(22iii) $(y^2 - x^2y)/(x^2 + y)^2$, $x^3/(x^2 + y)^2$

(22vii) $2xe^{x^2+y^2}$, $2ye^{x^2+y^2}$

(22viii) $y \ln(xy) + y$, $x \ln(xy) + x$

(22ix) $-x/\sqrt{1-x^2-y^2}$, $-y/\sqrt{1-x^2-y^2}$

(22xii) $\tan y$, $x/\cos^2 y$

(22xiii) $-1/(x^2y)$, $-1/(xy^2)$

(26i) The linear approximation formula is equation (10), in which $x_0 = a = 3$, $y_0 = b = 1$, and $\Delta x = x - a = x - 3$, $\Delta y = y - b = y - 1$. So for this problem the linear approximation of $f(x, y) = xy^2$ at $(3, 1)$ is

$$f(x, y) \approx 3 + (x - 3) + 6(y - 1) = x + 6y - 6.$$

This approximation is only expected to be good when (x, y) is close to $(3, 1)$. The approximation contains an error which is small compared to $|x - 3|$ and $|y - 1|$.

FAQ: What is the relation between the linear approximation and the tangent plane?

Answer: They are very closely related: the tangent plane is the graph of the linear approximation. The linear approximation is the equation for the tangent plane. To compute either you have to do the same thing.

(26ii) $x/y^2 \approx 3 + (x - 3) - 6(y - 1) = x - 6y + 6$ when x is close to 3 and y is close to 1.

(26iii) $\sin x + \cos y \approx -1 + (-1)(x - \pi) + (0)(y - \pi) = \pi - 1 - x$ when x is close to π and y is close to π.

(26iv) $\frac{xy}{x+y} \approx \frac{3}{4} + \frac{1}{16}(x - 3) + \frac{9}{16}(y - 1)$ when x is close to 3 and y is close to 1.

(27) $z = 1$

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

(29) $z = (x - 2) + 4(y - 1/2)$

(30) The graph has equation $z = x^2 - 2xy$, The tangent plane has equation $z = 2x - 4y$.

The part of the tangent plane which lies under the graph is given by

$$2x - 4y < x^2 - 2xy,$$

i.e. by

$$x^2 - 2xy - 2x + 4y > 0$$

With some luck you see that the LHS can be factored as

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

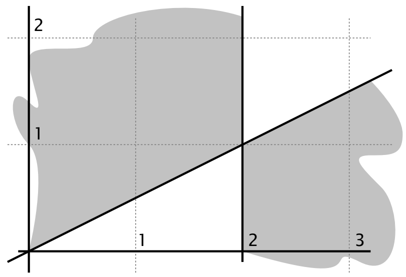
so that the part of the tangent plane which is under the graph consists of those points (x, y) for which either

$$x > 2, \text{ and } x > 2y$$

or

$$x < 2, \text{ and } x < 2y$$

holds.



(31i) The tangent plane has equation $z = ab + b(x-a) + a(y-b) = bx + ay - ab$.

(31ii) The point (x, y, z) lies on the intersection if $z = xy$ and $z = bx + ay - ab$. Therefore x and y must satisfy $xy - bx - ay + ab = 0$. This equation factors as follows:

$$xy - bx - ay + ab = (x-a)(y-b) = 0,$$

so that the intersection contains the line $x = a$, $z = ay$, and also the line $y = b$, $z = bx$.

(32i) Solve for z: $z = \pm\sqrt{2x^2 + 3y^2 - 4}$. In this problem we are looking at the point $(1, 1, -1)$ so we have the graph of $z = f(x, y) = -\sqrt{2x^2 + 3y^2 - 4}$. The partials are

$$\frac{\partial f}{\partial x} = \frac{-2x}{\sqrt{2x^2 + 3y^2 - 4}}, \qquad \frac{\partial f}{\partial y} = \frac{-3y}{\sqrt{2x^2 + 3y^2 - 4}}$$

so that, at $(1, 1, -1)$ you get $f_x = -2$, $f_y = -3$. There for the equation for the tangent plane is $z = -2(x-1) - 3(y-1) - 1$

(33i) The tangent plane has equation $z = z_0 + A(x-x_0) + B(y-y_0)$. By putting the variables x, y, z on one side, and all the constants on the other, you can write this as

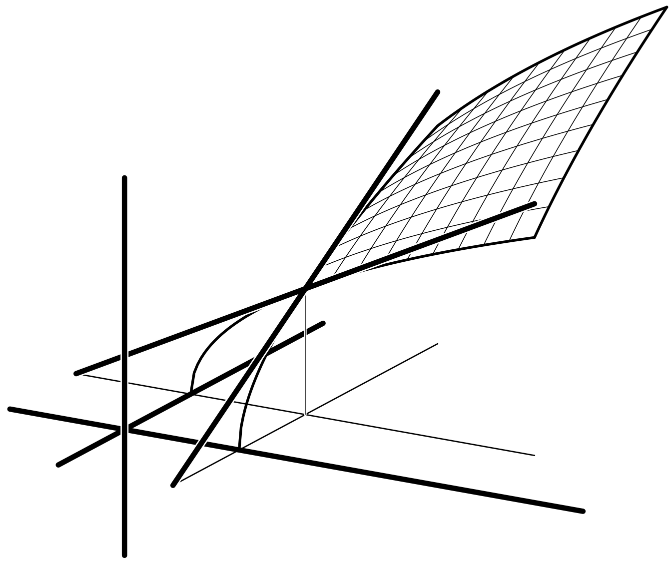
$$Ax + By - z = Ax_0 + By_0 - z_0.$$

This is the equation for a plane whose normal is $\vec{n} = \begin{pmatrix} A \\ B \\ -1 \end{pmatrix}$. Any other multiple of this vector is also a valid normal to the plane, in particular, $\begin{pmatrix} -A \\ -B \\ +1 \end{pmatrix}$ is OK.

(33ii) We want a normal to the graph of $z = f(x, y) = \frac{1}{2}x^2 + 2y^2$ at the point P. By the previous problem a normal is given by $\vec{n} = \begin{pmatrix} f_x(2,1) \\ f_y(2,1) \\ -1 \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \\ -1 \end{pmatrix}$.

A line through P in the direction of $\vec{n}$ is given by $\vec{r}(t) = \begin{pmatrix} 2 \\ 1 \\ 4 \end{pmatrix} + t\begin{pmatrix} 2 \\ 4 \\ -1 \end{pmatrix}$

(34) Below you see the graph of a function and two (solid) lines which are tangent to the graph. On one line you have $x = a$ (hence constant), and its slope is $f_x(a, b)$; on the other you have $y = b$, and it has slope $f_y(a, b)$.



The tangent plane to the graph (not drawn here, but see Figure 2 in the notes) is the plane containing the two lines in the drawing.

(35i) At $(2,1)$ the gradient is $\vec{\nabla} T = \begin{pmatrix} -2x \\ -9y^2 \end{pmatrix} = \begin{pmatrix} -4 \\ -9 \end{pmatrix}$. To cool off as fast as possible the bug should go in the opposite direction, i.e. in the direction of $\begin{pmatrix} 4 \\ 9 \end{pmatrix}$, or any positive multiple of this vector.

(35ii) At $(1,3)$ the gradient is $\vec{\nabla} T = \begin{pmatrix} -2 \\ -81 \end{pmatrix}$. To keep its temperature constant the bug should walk in any direction perpendicular to the gradient. The vector $\begin{pmatrix} 81 \\ -2 \end{pmatrix}$ is perpendicular to the gradient, so the bug should go in the direction of $\begin{pmatrix} 81 \\ -2 \end{pmatrix}$ or the opposite direction, $\begin{pmatrix} -81 \\ 2 \end{pmatrix}$.

Any non-zero multiple of $\begin{pmatrix} -81 \\ 2 \end{pmatrix}$ is also a valid answer, since we can only give the *direction* and not the speed.

Remember: the vector $\begin{pmatrix} -b \\ a \end{pmatrix}$ is perpendicular to $\begin{pmatrix} a \\ b \end{pmatrix}$.

(36) The function is $f(x,y) = x\ln(xy)$. We have $f(2,\frac{1}{2}) = 2\ln(2\cdot\frac{1}{2}) = \ln 1 = 0$. The gradient of the function is $\vec{\nabla} f = \begin{pmatrix} \ln(xy)+1 \\ x/y \end{pmatrix}$. At the point $(2,\frac{1}{2})$ this is $\vec{\nabla} f = \begin{pmatrix} 1 \\ 4 \end{pmatrix}$, so the linear approximation is

$$f(x,y) \approx f(2,\frac{1}{2}) + 1\cdot(x-2) + 4\cdot(y-\frac{1}{2}),$$

i.e.

$$f(x,y) \approx 1(x-2) + 4(y-\frac{1}{2}).$$

(This is also the answer to problem 29.)

Here we don't want to describe the tangent plan, but we want to find the value of $f(x, y)$ for $(x, y) = (1.98, 0.4)$. Substituting these values of x and y in the linear approximation we get $f(1.98, 0.4) \approx (1.98 - 2) + 4(0.4 - 0.5) = -0.42$.

This is only an approximation, and you wonder how good it is. We have $\Delta x = 1.98 - 2 = -0.02$, and $\Delta y = 0.4 - \frac{1}{2} = -0.1$... are these numbers "small"? To find the error in the approximation you could use a Lagrange-type remainder term, but that's not part of math 234. Instead we grab a calculator and compute $f(1.98, 0.4) = 1.98 \cdot \ln(1.98 \cdot 0.4) = -0.46172\cdots$. So our linear approximation formula is off by $0.04\cdots$.

(38) $\frac{\partial(f+g)}{\partial x} = f_x + g_x$, and $\frac{\partial(f+g)}{\partial y} = f_y + g_y$, so

$$\begin{pmatrix} \frac{\partial(f+g)}{\partial x} \\ \\ \frac{\partial(f+g)}{\partial y} \end{pmatrix} = \begin{pmatrix} f_x + g_x \\ f_y + g_y \end{pmatrix} = \begin{pmatrix} f_x \\ f_y \end{pmatrix} + \begin{pmatrix} g_x \\ g_y \end{pmatrix}$$

Hence $\vec{\nabla}(f + g) = \vec{\nabla} f + \vec{\nabla} g$.

(39ii) The gradient is $\vec{\nabla} f = \left(\begin{smallmatrix} 2x \\ 8y \end{smallmatrix}\right)$. This vector is parallel to $\left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right)$ if there is a number s such that $\vec{\nabla} f = s\left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right)$, i.e. $\begin{pmatrix} f_x \\ f_y \end{pmatrix} = \left(\begin{smallmatrix} s \\ s \end{smallmatrix}\right)$. This happens if $f_x(x, y) = f_y(x, y)$. From our computation of the partial derivatives of f we find that $\vec{\nabla} f$ is parallel to $\left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right)$ when $2x = 8y$. This happens at every point on the line $y = \frac{1}{4}x$.

We are asked which points on the level set $f = 4$ satisfy this condition, so we must find where the line $y = \frac{1}{4}x$ intersects the level set $x^2 + 4y^2 = 4$. Solving the two equations gives two points $(\frac{4}{5}\sqrt{5}, \frac{1}{5}\sqrt{5})$ and $(-\frac{4}{5}\sqrt{5}, -\frac{1}{5}\sqrt{5})$.

(39iii) $\vec{\nabla} g = \begin{pmatrix} 4y^2 \\ 8xy \end{pmatrix}$. This is parallel to $\left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right)$ when $y = 2x$. This line intersects the level set $g = 4$ in the point $(\frac{1}{2}\sqrt[3]{2}, \sqrt[3]{2})$.

Note: when you solve the equations $\vec{\nabla} g = \left(\begin{smallmatrix} s \\ s \end{smallmatrix}\right)$, you find $y = 2x$, but also the line $y = 0$ (x-axis). On this line the gradient actually vanishes, i.e. $\vec{\nabla} g = \vec{\mathbf{0}}$ and has no direction, so you can't really say it is parallel to $\left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right)$.

(40i) It's a paraboloid of revolution.

(40ii) $\vec{\nabla} f = \begin{pmatrix} 2x \\ 2y \\ -2 \end{pmatrix} = s\begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}$ if $-2 = 2s$, i.e. $s = -1$. This then implies $2x = -1$, $2y = -1$, so that $x = y = -\frac{1}{2}$. Since the point has to lie on the zero set of f, we find $z = \frac{1}{2}(x^2 + y^2) = \frac{1}{4}$.

(41) The zero set doesn't have to be a curve. For example the zero set of the function $f(x, y)$ =distance from (x, y) to the square Q (Problems 4 and 25) is the whole square Q.

(42) $\|\vec{\nabla} f\|$ is larger at the top right, because there the function f changes faster.

(47ii) The result of a rather long calculation is that $\|\vec{\nabla} f\| = 1$ everywhere outside the square, and $\|\vec{\nabla} f\| = 0$ inside the square (because f is constant in the square.)

(49) $ax + by + cz = R^2$.

(50) $4xt\cos(x^2 + y^2) + 6yt^2\cos(x^2 + y^2)$

(51) $2xy\cos t + 2x^2 t$

(52) $2xyt\cos(st) + 2x^2 s$, $2xys\cos(st) + 2x^2 t$

(53) $2xy^2 t - 4yx^2 s$, $2xy^2 s + 4yx^2 t$

(55i) $\frac{\partial T_B}{\partial Y} = -\sin\alpha \frac{\partial T_A}{\partial x} + \cos\alpha \frac{\partial T_A}{\partial y}$.

(55ii) Take the formulas for $\frac{\partial T_B}{\partial X}$ and $\frac{\partial T_B}{\partial Y}$ and work out the right hand side in this problem.

(58i) $\vec{\mathbf{E}} = -\vec{\boldsymbol{\nabla}} \ln r = \frac{1}{r^2}\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right)$.

(58ii) $\|\vec{\mathbf{E}}\| = 1/r = \frac{1}{\sqrt{x^2+y^2}}$.

(62i) Height $= -(x^2 - y^2)/(x^2 + y^2)$

(62ii) Height $= \sin 2\theta$.

(62iii) Height $= \cos 2\varphi$.

(64) $f_x = 3x^2y^2$, $f_y = 2x^3y + 5y^4$, $f_{xx} = 6xy^2$, $f_{yy} = 2x^3 + 20y^3$, $f_{xy} = 6x^2y$

(65) $f_x = 12x^2 + y^2$, $f_y = 2xy$, $f_{xx} = 24x$, $f_{yy} = 2x$, $f_{xy} = 2y$

(66) $f_x = \sin y$, $f_y = x\cos y$, $f_{xx} = 0$, $f_{yy} = -x\sin y$, $f_{xy} = \cos y$

(72) A function of two variables has

$$f_{xx}, \qquad f_{xy} = f_{yx}, \qquad f_{yy},$$

so it has **three** different partial derivatives of second order.

A function of three variables has these partial derivatives:

$$\begin{matrix} f_{xx} & f_{xy} & f_{xz} \\ f_{yx} & f_{yy} & f_{yz} \\ f_{zx} & f_{zy} & f_{zz} \end{matrix}$$

The ones "below the diagonal" are the same as corresponding derivatives above the diagonal, so there are only six different partial derivatives of second order, namely these:

$$\begin{matrix} f_{xx} & f_{xy} & f_{xz} \\ & f_{yy} & f_{yz} \\ & & f_{zz} \end{matrix}$$

A function of two variables has

$$\begin{gathered} f_{xxx}, \\ f_{xxy} = f_{xyx} = f_{yxx}, \\ f_{xyy} = f_{yxy} = f_{yyx}, \\ \text{and } f_{yyy} \end{gathered}$$

so **four** different partial derivatives of third order.

(78) We have $g(u,v) = f(u+v, u-v)$, so

$$\frac{\partial g}{\partial u} = \frac{\partial f}{\partial x}\frac{\partial(u+v)}{\partial u} + \frac{\partial f}{\partial y}\frac{\partial(u-v)}{\partial u} = f_x(u+v,u-v) + f_y(u+v,u-v).$$

Similarly,

$$\frac{\partial g}{\partial v} = f_x(u+v,u-v) - f_y(u+v,u-v).$$

Differentiate again to get

$$\frac{\partial^2 g}{\partial u^2} = f_{xx}(u+v,u-v) + 2f_{xy}(u+v,u-v) + f_{yy}(u+v,u-v)$$
$$\frac{\partial^2 g}{\partial v^2} = f_{xx}(u+v,u-v) - 2f_{xy}(u+v,u-v) + f_{yy}(u+v,u-v)$$
$$\frac{\partial^2 g}{\partial u \partial v} = f_{xx}(u+v,u-v) - f_{yy}(u+v,u-v)$$

(80i) If $y \neq 0$ then you can increase $x^2 - x^3 - y^2$ by setting $y = 0$. To put it differently, no matter what you choose for y, you always have

$$f(x,y) = x^2 - x^3 - y^2 \leq x^2 - x^3 = f(x,0).$$

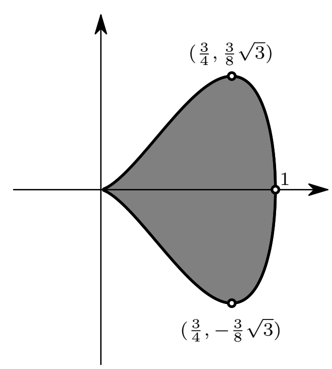
(80ii) The maximum has to appear on the x axis, so the question is *which $x \geq 0$ maximizes $f(x,0) = x^2 - x^3$?*

This is a Math 221 question. The answer is at $x = 2/3$.

(80iii) No, $\lim_{x\to-\infty} f(x,y) = +\infty$, so f has no largest value.

(81)

The quantity $4(x^3 - x^4) = 4x^3(1-x)$ is negative when $x < 0$ or $x > 1$, so the region is confined to the vertical strip $0 \leq x \leq 1$. Within this strip R is comprised of those points which satisfy $-\sqrt{4(x^3-x^4)} \leq y \leq +\sqrt{4(x^3-x^4)}$. The largest x value is attained at the point with $x = 1$, where $y = 0$, so, at the point $(1,0)$. The smallest x value is attained at the point $(0,0)$. The largest y value is attained at the point where $y^2 = 4x^3 - 4x^4$ is maximal. This happens when $x = \frac{3}{4}$, and the largest y value is therefore $\sqrt{4[(3/4)^3 - (3/4)^4]} = \frac{3}{8}\sqrt{3}$. The smallest y value also occurs at $x = \frac{3}{4}$ and is given by $y = -\frac{3}{8}\sqrt{3}$.



(82i) $f_x = 2x-2$, $f_y = 8y+8$, $f_{xx} = 2$, $f_{xy} = 0$, $f_{yy} = 8$.

There is exactly one critical point, at $(x,y) = (1,-1)$.

The 2nd order Taylor expansion at this point is

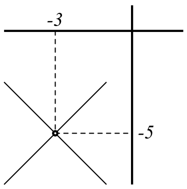
$$f(1+\Delta x, -1+\Delta y) = f(1,-1) + (\Delta x)^2 + 4(\Delta y)^2 + \cdots$$

The quadratic part is positive definite, therefore f has a **local minimum** at $(1,-1)$.

(82ii) $f_x = 2x + 6$, $f_y = -2y - 10$, $f_{xx} = 2$, $f_{xy} = 0$, $f_{yy} = -2$.

There is exactly one critical point, at $(x, y) = (-3, -5)$.

The 2nd order Taylor expansion at this point is

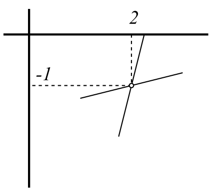


Critical point and level set near the critical point.

$$f(-3+\Delta x, -5+\Delta y) = f(-3,-5)+(\Delta x)^2-(\Delta y)^2+\cdots = f(-3,-5)+\big(\Delta x-\Delta y\big)\big(\Delta x+\Delta y\big)+\cdots$$

The quadratic part factors, therefore f has a **saddle point** at $(-3, -5)$. The level set near the critical point consists of two crossing curves whose tangents are given by the equations $\Delta x = \Delta y$ and $\Delta x = -\Delta y$. Since $\Delta x = x - a = x + 3$ and $\Delta y = y - b = y + 5$, the two tangent lines have equations $x + 3 = y + 5$ and $x + 3 = -(y + 5)$.

(82iii) $f_x = 2x + 4y$, $f_y = 4x + 2y$, $f_{xx} = 2$, $f_{xy} = 4$, $f_{yy} = 2$. There is one critical point: $(x, y) = (2, -1)$.



Critical point and level set near the critical point.

The 2nd order Taylor expansion at this point is

$$\begin{aligned} f(2+\Delta x, -1+\Delta y) &= f(2,-1) + (\Delta x)^2 + 4\Delta x\Delta x + (\Delta y)^2 + \cdots \\ &= f(2,-1) + \big(\Delta x + 2\Delta y\big)^2 - 3(\Delta y)^2 + \cdots \\ &= f(2,-1) + \big(\Delta x + (2+\sqrt{}3)\Delta y\big)\big(\Delta x + (2-\sqrt{}3)\Delta y\big) + \cdots \end{aligned}$$

The quadratic part factors, therefore f has a **saddle point** at $(2, -1)$. The level set near the critical point consists of two crossing curves whose tangents are given by the equations $\Delta x = (2+\sqrt{}3)\Delta y$ and $\Delta x = (2-\sqrt{}3)\Delta y$. Since $\Delta x = x-a = x-2$ and $\Delta y = y-b = y+1$, the two tangent lines have equations $x - 2 = (2+\sqrt{}3)(y+1)$ and $x - 2 = (2-\sqrt{}3)(y+1)$.

(82iv) $f_x = 2x - y - 5$, $f_y = -x + 4y + 6$, $f_{xx} = 2$, $f_{xy} = -1$, $f_{yy} = 4$.

There is again one critical point: $x = 2$, $y = -1$.

The 2nd order Taylor expansion at this point is

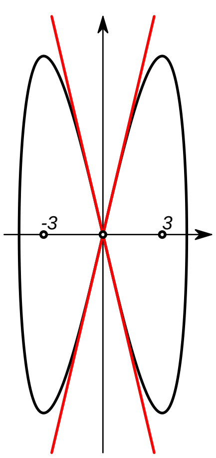
$$\begin{aligned} f(2+\Delta x, -1+\Delta y) &= f(2,-1) + (\Delta x)^2 - \Delta x\Delta x + 2(\Delta y)^2 + \cdots \\ &= f(2,-1) + \big(\Delta x - \tfrac{1}{2}\Delta y\big)^2 + \tfrac{7}{4}(\Delta y)^2 + \cdots \end{aligned}$$

The second order part of the Taylor expansion is positive, so $(2, -1)$ is a **local minimum**.

(82v) $f_x = -36x + 4x^3$, $f_y = 2y$, $f_{xx} = -36 + 12x^2$, $f_{xy} = 0$, $f_{yy} = 2$.

The equation $f_x = 0$ has three solutions, $x = 0$ and $x = \pm 3$. The equation $f_y = 0$ has only one solution $y = 0$. Therefore there are three critical points, the origin and the points $(\pm 3, 0)$.

The taylor expansions at these points are



Critical points and zero set.

$$\begin{aligned} f(\Delta x, \Delta y) &= f(0,0) - 18(\Delta x)^2 + (\Delta y)^2 + \cdots \\ &= f(0,0) + \big(\Delta y - \sqrt{18}x\big)\big(\Delta y + \sqrt{18}x\big) + \cdots \\ f(3+\Delta x, \Delta y) &= f(3,0) + 36(\Delta x)^2 + (\Delta y)^2 + \cdots \\ f(-3+\Delta x, \Delta y) &= f(-3,0) + 36(\Delta x)^2 + (\Delta y)^2 + \cdots \end{aligned}$$

The second order terms in the Taylor expansions at $(3, 0)$ and at $(-3, 0)$ are both positive for all Δx and Δy, so both points $(\pm 3, 0)$ are local minima. The second order part of the expansion at the origin factors and hence the origin is a saddle point. The tangents to the zeroset at the origin are the lines $\Delta y = \pm\sqrt{18}\Delta x = \pm 3\sqrt{2}\Delta x$. Since here $\Delta x =$ "$x - a$" $= x$, and $\Delta y = y$, the tangents are the lines through the origin given by $y = \pm 3\sqrt{2}x$.

You can try to draw the zeroset of this function and analyze it in the same way as the "fishy example" in 4.3. The zeroset of f consists of the graphs of $y = \pm\sqrt{18x^2 - x^4} = \pm|x|\sqrt{18 - x^2}$. It looks like a squashed "∞" or a butterfly (you decide.)

(82vi) There are nine critical points. Four global minima at $(\pm 3, \pm\sqrt{3})$, four saddle points at $(0, \pm\sqrt{3})$ and $(\pm 3, 0)$ respectively, and finally, a local but not global maximum at the origin.

(82vii) critical point at $(1, -1/6)$ $f_x = 4 - 4x$, $f_y = 1 - 6y$, $f_{xx} = -4$, $f_{xy} = 0$, $f_{yy} = -6$.

Second order Taylor expansion at the critical point:

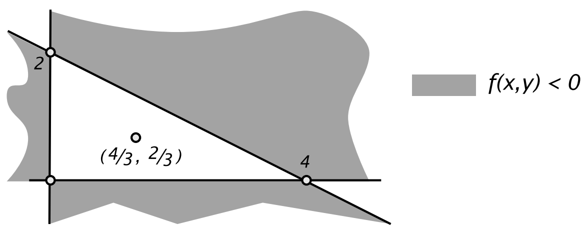
$$f(-1 + \Delta x, -\tfrac{1}{6} + \Delta y) = f(1, -\tfrac{1}{6}) - 2(\Delta x)^2 - 3(\Delta y)^2 + \cdots$$

The second order terms are always negative so $(1, -\frac{1}{6})$ is a local maximum.

(82viii) The derivatives are:

$$f_x = 4y - 2xy - 2y^2, \quad f_y = 4x - x^2 - 4xy, \quad f_{xx} = -2y, \quad f_{xy} = 4 - 2x - 4y, \quad f_{yy} = -4x.$$

This function is given in factored form, so without solving the equations $f_x = 0$, $f_y = 0$ you can say the following about this problem. The zero set consists of the three lines: the y-axis ($x = 0$), the x-axis ($y = 0$) and the line with equation $4 - x - 2y = 0$. It follows that the intersection points $(0, 0)$, $(4, 0)$, and $(0, 2)$ of these lines are saddle points. Since $f > 0$ in the triangle formed by the three lines this triangle must contain at least one local maximum.



To find all critical points solve these equations:

$$f_x = 4y - 2xy - 2y^2 = 0 \iff \boxed{y = 0 \text{ or } 4 - 2x - 2y = 0}$$

and

$$f_y = 4x - x^2 - 4xy = 0 \iff \boxed{x = 0 \text{ or } 4 - x - 4y = 0}$$

Since both equations $f_x = 0$ and $f_y = 0$ lead to two possibilities, we have to consider $2 \times 2 = 4$ cases:

- $y = 0$ **&** $x = 0$**:** This tells us the origin is a critical point
- $y = 0$ **&** $4 - x - 4y = 0$**:** Solving these equations leads to $x = 4, y = 0$, so $(4, 0)$ is a critical point.
- $4 - 2x - 2y = 0$ **&** $x = 0$**:** Solve and you find that $(0, 2)$ is a critical point.
- $4 - 2x - 2y = 0$ **&** $4 - x - 4y = 0$**:** Solve these equations and you get $(x, y) = (\frac{4}{3}, \frac{2}{3})$.

The first three critical points are the saddle points we predicted. The fourth critical point must be a local maximum, since there has to be one in the triangle, and of all the critical points we have found the others are all saddle points.

(82x) Two saddle points: $(2, 2)$ and $(-2, -2)$

(82xii) The origin. Neither a local max, min, nor saddle. The graph of this function is called the "Monkey Saddle" as it accommodates two legs and a tail too. Draw it in your graphing program to see this.

(82xiii) Zero set is the parabola with equation $x = y^2$, and the line $x = 1$. They intersect at $(1, \pm 1)$, so the function has two saddle points $(1, 1)$ and $(1, -1)$. The region between the line $x = 1$ and the parabola must contain local minimum. It is located at $(\frac{1}{2}, 0)$.

(82xiv) Two saddle points : $(2, 2)$ and $(-2, -2)$. Yes, this problem appeared twice.

(82xv) All points on the y-axis are critical points. They are all global minima, but the second derivative test doesn't tell you so.

(82xvi) All points on the y-axis are again critical points. Those with $y > 0$ are local minima, those with $y < 0$ are local maxima, and the origin is neither. The second derivative test applies to none of these points.

(82xvii) All points on the unit circle are global minima, because the function vanishes there, and is positive everywhere else. The origin is a local maximum. The 2nd derivative test applies to the origin, but not to any of the other critical points.

(82xviii) All points on the y-axis are again critical points. Those with $y > 0$ are local minima, those with $y < 0$ are local maxima, and the origin is neither. The second derivative test applies to none of these points.

(86) $(3, 4/3)$

(87) $x = (a + c + e)/3$, $y = (b + d + f)/3$.

(88) You have to show that $f_x(a, b) = f_y(a, b) = 0$. By the product rule $f_x(a, b) = g_x(a, b)h(a, b) + g(a, b)h_x(a, b)$. Since both $g(a, b) = 0$ and $h(a, b) = 0$, it follows that $f_x(a, b) = 0$. The same reasoning applies to $f_y(a, b)$.

(90i) One variable calculus. There's only one variable, a, and you must solve $E'(a) = 0$.

(90ii) $a = (x_1 + \cdots + x_N)/N$, i.e. the average provides "the best fit."

(91i) Three: a, b, and c.

(91ii) The equations for (a, b, c) are:

$$\begin{array}{ccccccccc}
(\sum x_k^4) & a & + & (\sum x_k^3) & b & + & (\sum x_k^2) & c & = & \sum x_k^2 y_k \\
(\sum x_k^3) & a & + & (\sum x_k^2) & b & + & (\sum x_k) & c & = & \sum x_k y_k \\
(\sum x_k^2) & a & + & (\sum x_k) & b & + & N & c & = & \sum y_k
\end{array}$$

(92) The equations are

$$\begin{array}{ccccccccc}
(\sum x_k^2) & a & + & (\sum x_k y_k) & b & + & (\sum x_k) & c & = & \sum x_k z_k \\
(\sum x_k y_k) & a & + & (\sum y_k^2) & b & + & (\sum y_k) & c & = & \sum y_k z_k \\
(\sum x_k) & a & + & (\sum y_k) & b & + & N & c & = & \sum z_k
\end{array}$$

(93) In this problem you are asked to find Taylor expansions of functions at various points. Since these points are not critical points, the expansions you find will generally have first and second oder terms. In the expansions you will compute when you use the second derivative test later on, there will be no first order terms.

(93i) $f(\Delta x, \Delta y) = \Big(1 - \Delta x + \Delta x \Delta y\Big)^2 = 1 - 2\Delta x + \Delta x^2 + 2\Delta x \Delta y + \cdots$

(93ii) $f(1+\Delta x, 1+\Delta y) = \Big(1 - (1+\Delta x) + (1+\Delta x)(1+\Delta y)\Big)^2 = 1 + 2\Delta y + 2\Delta x \Delta y + 2(\Delta y)^2 + \cdots$

(93iii) $f(\Delta x, \Delta y) = e^{\Delta x - (\Delta y)^2} = 1 + \Delta x + \frac{1}{2}(\Delta x)^2 - (\Delta y)^2 + \cdots$

(93iv) $f(1+\Delta x, 1+\Delta y) = e^{(1+\Delta x)-(1+\Delta y)^2} = 1 + \Delta x - 2\Delta y + \frac{1}{2}(\Delta x)^2 - 2\Delta x \Delta y + (\Delta y)^2 + \cdots$

(95) Complete the square and you get

$$Q(x, y) = \big(x - ay\big)^2 + \big(1 - a^2\big)y^2.$$

When $1 - a^2 > 0$, i.e. when $-1 < a < 1$ the form is positive definite. When $a = \pm 1$ the form is a perfect square, namely,

$$x^2 \pm 2xy + y^2 = \big(x \pm y\big)^2.$$

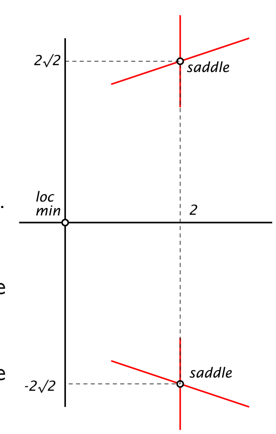
When $1 - a^2 < 0$, i.e. when $a > 1$ or $a < -1$, the form is indefinite:

$$x^2 + 2axy + y^2 = \Big(x - ay - \sqrt{a^2 - 1}y\Big)\Big(x - ay + \sqrt{a^2 - 1}y\Big) = (x - k_+ y)(x - k_- y),$$

where $k_\pm = -a \pm \sqrt{a^2 - 1}$.

(96) See the solutions to Problem 82 for the solutions to this problem.

(98i) $f_x = 2x - \frac{1}{2}y^2$, $f_y = 2y - xy$. The equation $f_y = y(2 - x) = 0$ leads to two possibilities: $x = 2$ or $y = 0$. If $y = 0$ then $f_x = 0$ implies $x = 0$, which gives us one critical point, the origin $(0, 0)$. If on the other hand $x = 2$, then $f_x = 0$ implies $y^2 = 8 \iff y = \pm 2\sqrt{2}$. We therefore get two more critical points $(2, \pm 2\sqrt{2})$.



The second derivatives are $f_{xx} = 2$, $f_{xy} = -y$, $f_{yy} = 2 - x$. Therefore we have the following Taylor expansions at the three critical points:

$$\begin{aligned}
f(\Delta x, \Delta y) &= f(0,0) + (\Delta x)^2 + (\Delta y)^2 + \cdots && \implies \text{loc.min.}\\
f(2+\Delta x, 2\sqrt{2}+\Delta y) &= f(2, 2\sqrt{2}) + (\Delta x)^2 - 2\sqrt{2}\Delta x \Delta y + 0(\Delta y)^2 + \cdots \\
&= f(2, 2\sqrt{2}) + \big(\Delta x - 2\sqrt{2}\Delta y\big)\Delta x + \cdots && \implies \text{saddle}\\
f(2+\Delta x, -2\sqrt{2}+\Delta y) &= f(2, -2\sqrt{2}) + (\Delta x)^2 + 2\sqrt{2}\Delta x \Delta y + 0(\Delta y)^2 + \cdots \\
&= f(2, -2\sqrt{2}) + \big(\Delta x + 2\sqrt{2}\Delta y\big)\Delta x + \cdots && \implies \text{saddle}
\end{aligned}$$

The origin is therefore a local minimum, and the points $(2, \pm 2\sqrt{2})$ are saddle points. At $(0, 2\sqrt{2})$ the level set consists of two crossing curves, whose tangents are given by $\Delta x = 0$ (a vertical line) and $\Delta x = 2\sqrt{2}\Delta y$ (a line with slope $1/2\sqrt{2} = \frac{1}{4}\sqrt{2}$).

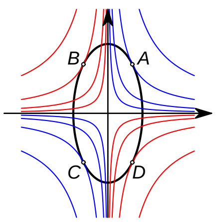
(98iii) $f_x = 1 - y^2$, $f_y = 2 - 2xy$. Critical points: $f_x = 0$ holds when $y = \pm 1$. If $y = +1$, then $f_y = 0$ implies $x = 1$, and if $y = -1$ then $f_y = 0$ implies $x = -1$. There are therefore two critical points, $(1, 1)$ and $(-1, -1)$.

(101) $f(x,y) = xy$, $g(x,y) = x^2 + \frac{1}{4}y^2$. $\vec{\nabla} f = \binom{y}{x}$, $\vec{\nabla} g = \binom{2x}{y/2}$.

First we check for possible max/minima which satisfy $\vec{\nabla} g = \vec{0}$. But the only point (x,y) satisfying $\vec{\nabla} g(x,y) = \binom{0}{0}$ is the origin $(x,y) = (0,0)$, and this point does not lie on the constraint set.

Therefore, **if** there is a minimum it is attained at a solution of Lagrange's equations

$$\begin{aligned} f_x = \lambda g_x &\iff y = 2\lambda x \\ f_y = \lambda g_y &\iff x = \lambda y/2 \\ g(x,y) = 1 &\iff x^2 + \tfrac{1}{4}y^2 = 1 \end{aligned}$$



Level sets of the function $f(x,y) = xy$ and the constraint set $x^2 + \frac{1}{4}y^2 = 1$

Multiply the first equation with y and the second with $4x$, then you get

$$y^2 = 2\lambda xy \text{ and } 4x^2 = 2\lambda xy$$

Hence $y^2 = 4x^2$. Put that in the constraint, and you find

$$1 = x^2 + \tfrac{1}{4}y^2 = 2x^2.$$

Thus $x = \pm\sqrt{1/2} = \pm\frac{1}{2}\sqrt{2}$ and $y = \pm\sqrt{2}$. In all we have found **four** possible solutions. Lagrange's method does not tell us which, if any, of these are minima.

By looking at the constraint set (it's an ellipse with horizontal axis of length 1 and vertical axis of length 2) and taking into account that $f(x,y) = xy$ is positive in the first and third quadrants, and negative in the second and fourth, you find out that the two points $(\frac{1}{2}\sqrt{2}, \sqrt{2})$ and $(-\frac{1}{2}\sqrt{2}, -\sqrt{2})$ (A and C in the figure) are maximum points, while $(-\frac{1}{2}\sqrt{2}, \sqrt{2})$ and $(\frac{1}{2}\sqrt{2}, -\sqrt{2})$ (B and D in the figure) are minimum points.

(102i) Let the sides of the box be x, y, z. We want to minimize the quantity $A = 2xy + 2yz + 2xz$, with the constraint $V = xyz = \frac{1}{2}$. The constraint implies that $x \neq 0$, $y \neq 0$ and $z \neq 0$ moreover, given x and y the only z which satisfies the constraint is $z = 1/(2xy)$. Thus we must minimize the following function of two variables

$$A(x,y) = xy + \frac{1}{2x} + \frac{1}{2y}$$

over all $x > 0$, $y > 0$.

A minimum must be an interior minimum (can't be on the x or y-axis since these are excluded), and thus must be a critical point.

$$\frac{\partial A}{\partial x} = y - \frac{1}{2x^2}, \qquad \frac{\partial A}{\partial y} = x - \frac{1}{2y^2}.$$

Solving $A_x = A_y = 0$ for (x,y) leads to $x = y = \sqrt[3]{2}$, so the solution is a cube $1/\sqrt[3]{2}$ on a side

(102ii) We wish to minimize $A(x,y,z) = 2yz + 2xz + 2xy$ with constraint $V(x,y,z) = xyz = \frac{1}{2}$, using Lagrange's method.

First we check for exceptional points on the constraint set, i.e. points (x,y,z) that satisfy both $V(x,y,z) = \frac{1}{2}$ and $\vec{\nabla} V(x,y,z) = \vec{0}$. Since

$$\vec{\nabla} V = \begin{pmatrix} yz \\ xz \\ xy \end{pmatrix}$$

the gradient $\vec{\nabla} V$ vanishes if at least two of the three coordinates x, y, z are zero. But such a point can never satisfy the constraint $xyz = \frac{1}{2}$. Therefore, if there is a box with least area, its sides x, y, z must satisfy Lagrange's equations.

Lagrange's equations are

$$\begin{aligned} A_x = \lambda V_x &\iff 2y + 2z = \lambda yz \\ A_y = \lambda V_y &\iff 2x + 2z = \lambda xz \\ A_z = \lambda V_z &\iff 2x + 2y = \lambda xy \end{aligned}$$

To get rid of λ multiply the first equation with x and the second with y to get

$$y(2x + 2z) = \lambda xyz = x(2y + 2z) \implies 2xy + 2yz = 2xy + 2xz \implies 2yz = 2xz.$$

Therefore we find that either $z = 0$ or $x = y$. But $z = 0$ is not possible, because (x, y, z) must satisfy the constraint $xyz = 0$. Therefore we get $x = y$.

If you multiply the second Lagrange equation with y and the third with z then the same reasoning as above tells you that $y = z$.

So, **if there is a minimum** then it happens when $x = y = z$, i.e. when the box is a cube. The only cube that satisfies the constraint has sides $x = y = z = 2^{-1/3}$.

As always, Lagrange's method does not rule out the possibility that the cube we have found actually maximizes the surface area, rather than minimizing it. That this is actually not the case is something you would have to prove by other means. We will not do that in this course.

(103) Answer: the shortest distance is $\sqrt{100/3}$.

Solution: If (x, y, z) is any point than its distance to the origin is $d(x, y, z) = \sqrt{x^2 + y^2 + z^2}$. We want to minimize $d(x, y, z)$ over all points (x, y, z) which satisfy the constraint $g(x, y, z) = x + y + z = 10$. Instead of minimizing $d(x, y, z)$ we will minimize $f(x, y, z) = d(x, y, z)^2 = x^2 + y^2 + z^2$. You can do this problem directly with the function $d(x, y, z)$ and you will get the same answer – the computations are just a little longer because f has easier derivatives than d.

We use Lagrange's method. First we check for exceptional points, i.e. points on the constraint set which satisfy $\vec{\nabla} g = \vec{0}$. Since $\vec{\nabla} g = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ the gradient of g can never be the zero vector, so there are no exceptional points. If there is a minimum of f on the constraint set, it must be a solution of Lagrange's equations.

The Lagrange equations are

$$\begin{aligned} f_x = \lambda g_x &\iff 2x = \lambda \\ f_y = \lambda g_y &\iff 2y = \lambda \\ f_z = \lambda g_z &\iff 2z = \lambda \end{aligned}$$

Therefore **if there is a nearest point to the origin on the plane** then it must satisfy $x = y = z = \lambda/2$ as well as the constraint. The only point satisfying these conditions is $(\frac{10}{3}, \frac{10}{3}, \frac{10}{3})$.

Lagrange's method does not tell us that this is the nearest point. As far as Lagrange is concerned it could also be the furthest point from the origin. (But because we know what a plane looks like we "know" that there has to be a nearest point to the origin.)

(105i) Minimize $f(x, y, z) = (x-2)^2 + (y-1)^2 + (z-4)^2$ subject to the constraint $g(x, y, z) = 2x - y + 3z = 1$.

First, since $\vec{\nabla} g - \begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix} \neq \vec{0}$, there are no exceptional points, so the nearest point (if it exists) is a solution of Lagrange's equations. These are

$$2(x - 2) = 2\lambda, \quad 2(y - 1) = -\lambda, \quad 2(z - 4) = 3\lambda.$$

Eliminate λ to get

$$x = -2y + 4, \qquad z = -3y + 7.$$

Combined with the constraint you then find

$$y = 2, \quad x = 0, \quad z = 1.$$

The Lagrange multiplier is $\lambda = x - 2 = -2$.

The distance from the point we found to the given point $(2, 1, 4)$ is

$$d = \sqrt{(x-2)^2 + (y-1)^2 + (z-4)^2} = \sqrt{14}$$

(105ii) $|ax_0 + by_0 + cz_0 - d|/\sqrt{a^2+b^2+c^2}$

(108) a cube

(110) $65/3 \times 65/3 \times 130/3$

(111) It has a square base, and is one and one half times as tall as wide. If the volume is V the dimensions are $\sqrt[3]{2V/3} \times \sqrt[3]{2V/3} \times \sqrt[3]{9V/4}$.

(112) $(0, 0, 1)$, $(0, 0, -1)$

(113) $\sqrt[3]{4V} \times \sqrt[3]{4V} \times \sqrt[3]{V/16}$

(114) Farthest: $(-\sqrt{2}, \sqrt{2}, 2 + 2\sqrt{2})$; closest: $(2, 0, 0)$, $(0, -2, 0)$

(115i) 2

(115ii) 8

(115iii) $2/3$

(115iv) $\displaystyle\int_0^\pi \int_0^y \frac{\sin y}{y}\, dx\, dy = \int_0^\pi \frac{\sin y}{y} \cdot y\, dy = \int_0^\pi \sin y\, dy = 2.$

(115v) Except for a change in notation ($y \to \theta$ and $x \to r$) this is the same integral as in the previous problem. The answer is again 2.

(115vi) Which function is being integrated? It's the function $f(x, y) = 1$.
$\int_0^1 \int_0^{\sqrt{1-x^2}} dy\, dx = \int_0^1 [y]_{y=0}^{y=\sqrt{1-x^2}}\, dx = \int_0^1 \sqrt{1-x^2}\, dx$. The last integral is the area of a quarter circle with radius 1, so the answer is $\pi/4$.

(116) Once you compute the inner integral

$$\int_0^1 \sin(\pi x) dx = [-\cos \pi x]_{x=0}^1 = -\cos \pi - (-\cos 0) = 2,$$

you get

$$\int_x^1 \Big\{\int_0^1 \sin(\pi x) dx\Big\} dy = \int_x^1 2dy = [2y]_{y=x}^1 = 2(1-x).$$

The result depends on x. The x in the answer and the two x-es in the inner integral refer to different quantities. This is at best confusing, and should really never be done.

(117) This is almost true, but in fact false. The correct statement which looks like the one in the problem is

For any two functions $f(x)$ and $g(y)$ one has

$$\int_0^1 \int_0^2 f(x)g(y)\, dx\, dy = \int_0^2 f(x)\, dx \;\times\; \int_0^1 g(y)\, dy\ .$$

(what's the difference? Look at the integration bounds!) To give a counterexample for the statement in the problem, almost any two functions f and g will do, as long as f is not a constant multiple of g. For instance, if you choose $f(x) = x$, $g(y) = 1$, then you get

$$\int_0^1 \int_0^2 f(x)g(y)\, dx\, dy = \int_0^1 \int_0^2 x dx dy = 2.$$

but

$$\int_0^1 f(x)\, dx \;\times\; \int_0^2 g(y)\, dy = \int_0^1 x\, dx \;\times\; \int_0^2 dy = \frac{1}{2} \times 2 = 1.$$

(118) The volume under the graph is $\frac{1}{3}ba^3 + \frac{1}{3}ab^3 = \frac{1}{3}ab(a^2 + b^2)$. The volume of the surrounding block is $a \times b \times (a^2 + b^2)$, so the region beneath the graph occupies one third of the surrounding block, no matter which a or b you choose.

(119i) 16

(119ii) 4

(119iii) $15/8$

(119iv) $1/2$

(119v) $5/6$

(119vi) $12 - 65/(2e)$.

(119vii) $1/2$

(119viii) $(2/9)2^{3/2} - (2/9)$

(119ix) $(1 - \cos(1))/4$

(119x) $(2\sqrt{2} - 1)/6$

(119xi) $\pi - 2$

(120i) 8π

(120ii) 2

(120iii) $5/3$

(120iv) $81/2$

(120v) $2a^3/3$

(120vi) 4π

(120vii) $\pi/32$

(122i) A

(122ii) $B/2$

(123i) $\displaystyle\int_0^1 \int_0^{\sqrt{1-x^2}} \frac{2xy}{x^2+y^2}\, dy\, dx.$

(123ii) In P.C. the function simplifies to $F(r,\theta) = 2\sin\theta\cos\theta$, so the volume is

$$V = \int_0^1 \int_0^{\pi/2} 2\sin\theta\cos\theta r\, d\theta\, dr = \int_0^1 \left[\sin^2\theta\right]_0^{\pi/2} r\, dr = \tfrac{1}{2}.$$

CHAPTER 5

GNU Free Documentation License

Preamble

The purpose of this License is to make a manual, textbook, or other functional and useful document "free" in the sense of freedom: to assure everyone the effective freedom to copy and redistribute it, with or without modifying it, either commercially or noncommercially. Secondarily, this License preserves for the author and publisher a way to get credit for their work, while not being considered responsible for modifications made by others.

This License is a kind of "copyleft", which means that derivative works of the document must themselves be free in the same sense. It complements the GNU General Public License, which is a copyleft license designed for free software.

We have designed this License in order to use it for manuals for free software, because free software needs free documentation: a free program should come with manuals providing the same freedoms that the software does. But this License is not limited to software manuals; it can be used for any textual work, regardless of subject matter or whether it is published as a printed book. We recommend this License principally for works whose purpose is instruction or reference.

1. APPLICABILITY AND DEFINITIONS

This License applies to any manual or other work, in any medium, that contains a notice placed by the copyright holder saying it can be distributed under the terms of this License. Such a notice grants a world-wide, royalty-free license, unlimited in duration, to use that work under the conditions stated herein. The "**Document**", below, refers to any such manual or work. Any member of the public is a licensee, and is addressed as "**you**". You accept the license if you copy, modify or distribute the work in a way requiring permission under copyright law.

A "**Modified Version**" of the Document means any work containing the Document or a portion of it, either copied verbatim, or with modifications and/or translated into another language.

A "**Secondary Section**" is a named appendix or a front-matter section of the Document that deals exclusively with the relationship of the publishers or authors of the Document to the Document's overall subject (or to related matters) and contains nothing that could fall directly within that overall subject. (Thus, if the Document is in part a textbook of mathematics, a Secondary Section may not explain any mathematics.) The relationship could be a matter of historical connection with the subject or with related matters, or of legal, commercial, philosophical, ethical or political position regarding them.

The "**Invariant Sections**" are certain Secondary Sections whose titles are designated, as being those of Invariant Sections, in the notice that says that the Document is released under this License. If a section does not fit the above definition of Secondary then it is not allowed to be designated as Invariant. The Document may contain zero Invariant Sections. If the Document does not identify any Invariant Sections then there are none.

The "**Cover Texts**" are certain short passages of text that are listed, as Front-Cover Texts or Back-Cover Texts, in the notice that says that the Document is released under this License. A Front-Cover Text may be at most 5 words, and a Back-Cover Text may be at most 25 words.

A "**Transparent**" copy of the Document means a machine-readable copy, represented in a format whose specification is available to the general public, that is suitable for revising the document straightforwardly with generic text editors or (for images composed of pixels) generic paint programs or (for drawings) some widely available drawing editor, and that is suitable for input to text formatters or for automatic translation to a variety of formats suitable for input to text formatters. A copy made in an otherwise Transparent file format whose markup, or absence of markup, has been arranged to thwart or discourage subsequent modification by readers is not Transparent. An image format is not Transparent if used for any substantial amount of text. A copy that is not "Transparent" is called "**Opaque**".

Examples of suitable formats for Transparent copies include plain ASCII without markup, Texinfo input format, LaTeX input format, SGML or XML using a publicly available DTD, and standard-conforming simple HTML, PostScript or PDF designed for human modification. Examples of transparent image formats include PNG, XCF and JPG. Opaque formats include proprietary formats that can be read and edited only by proprietary word processors, SGML or XML for which the DTD and/or processing tools are not generally available, and the machine-generated HTML, PostScript or PDF produced by some word processors for output purposes only.

The "**Title Page**" means, for a printed book, the title page itself, plus such following pages as are needed to hold, legibly, the material this License requires to appear in the title page. For works in formats which do not have any title page as such, "Title Page" means the text near the most prominent appearance of the work's title, preceding the beginning of the body of the text.

The "**publisher**" means any person or entity that distributes copies of the Document to the public.

A section "**Entitled XYZ**" means a named subunit of the Document whose title either is precisely XYZ or contains XYZ in parentheses following text that translates XYZ in another language. (Here XYZ stands for a specific section name mentioned below, such as "**Acknowledgements**", "**Dedications**", "**Endorsements**", or "**History**".) To "**Preserve the Title**" of such a section when you modify the Document means that it remains a section "Entitled XYZ" according to this definition.

The Document may include Warranty Disclaimers next to the notice which states that this License applies to the Document. These Warranty Disclaimers are considered to be included by reference in this License, but only as regards disclaiming warranties: any other implication that these Warranty Disclaimers may have is void and has no effect on the meaning of this License.

2. VERBATIM COPYING

You may copy and distribute the Document in any medium, either commercially or noncommercially, provided that this License, the copyright notices, and the license notice saying this License applies to the Document are reproduced in all copies, and that you add no other conditions whatsoever to those of this License. You may not use technical measures to obstruct or control the reading or further copying of the copies you make or distribute. However, you may accept compensation in exchange for copies. If you distribute a large enough number of copies you must also follow the conditions in section 3.

You may also lend copies, under the same conditions stated above, and you may publicly display copies.

3. COPYING IN QUANTITY

If you publish printed copies (or copies in media that commonly have printed covers) of the Document, numbering more than 100, and the Document's license notice requires Cover Texts, you must enclose the copies in covers that carry, clearly and legibly, all these Cover Texts: Front-Cover Texts on the front cover, and Back-Cover Texts on the back cover. Both covers must also clearly and legibly identify you as the publisher of these copies. The front cover must present the full title with all words of the title equally prominent and visible. You may add other material on the covers in addition. Copying with changes limited to the covers, as long as they preserve the title of the Document and satisfy these conditions, can be treated as verbatim copying in other respects.

If the required texts for either cover are too voluminous to fit legibly, you should put the first ones listed (as many as fit reasonably) on the actual cover, and continue the rest onto adjacent pages.

If you publish or distribute Opaque copies of the Document numbering more than 100, you must either include a machine-readable Transparent copy along with each Opaque copy, or state in or with each Opaque copy a computer-network location from which the general network-using public has access to download using public-standard network protocols a complete Transparent copy of the Document, free of added material. If you use the latter option, you must take reasonably prudent steps, when you begin distribution of Opaque copies in quantity, to ensure that this Transparent copy will remain thus accessible at the stated location until at least one year after the last time you distribute an Opaque copy (directly or through your agents or retailers) of that edition to the public.

It is requested, but not required, that you contact the authors of the Document well before redistributing any large number of copies, to give them a chance to provide you with an updated version of the Document.

4. MODIFICATIONS

You may copy and distribute a Modified Version of the Document under the conditions of sections 2 and 3 above, provided that you release the Modified Version under precisely this License, with the Modified Version filling the role of the Document, thus licensing distribution and modification of the Modified Version to whoever possesses a copy of it. In addition, you must do these things in the Modified Version:

A. Use in the Title Page (and on the covers, if any) a title distinct from that of the Document, and from those of previous versions (which should, if there were any, be listed in the History section of the Document). You may use the same title as a previous version if the original publisher of that version gives permission.
B. List on the Title Page, as authors, one or more persons or entities responsible for authorship of the modifications in the Modified Version, together with at least five of the principal authors of the Document (all of its principal authors, if it has fewer than five), unless they release you from this requirement.
C. State on the Title page the name of the publisher of the Modified Version, as the publisher.
D. Preserve all the copyright notices of the Document.
E. Add an appropriate copyright notice for your modifications adjacent to the other copyright notices.
F. Include, immediately after the copyright notices, a license notice giving the public permission to use the Modified Version under the terms of this License, in the form shown in the Addendum below.
G. Preserve in that license notice the full lists of Invariant Sections and required Cover Texts given in the Document's license notice.
H. Include an unaltered copy of this License.
I. Preserve the section Entitled "History", Preserve its Title, and add to it an item stating at least the title, year, new authors, and publisher of the Modified Version as given on the Title Page. If there is no section Entitled "History" in the Document, create one stating the title, year, authors, and publisher of the Document as given on its Title Page, then add an item describing the Modified Version as stated in the previous sentence.
J. Preserve the network location, if any, given in the Document for public access to a Transparent copy of the Document, and likewise the network locations given in the Document for previous versions it was based on. These may be placed in the "History" section. You may omit a network location for a work that was published at least four years before the Document itself, or if the original publisher of the version it refers to gives permission.
K. For any section Entitled "Acknowledgements" or "Dedications", Preserve the Title of the section, and preserve in the section all the substance and tone of each of the contributor acknowledgements and/or dedications given therein.
L. Preserve all the Invariant Sections of the Document, unaltered in their text and in their titles. Section numbers or the equivalent are not considered part of the section titles.
M. Delete any section Entitled "Endorsements". Such a section may not be included in the Modified Version.
N. Do not retitle any existing section to be Entitled "Endorsements" or to conflict in title with any Invariant Section.
O. Preserve any Warranty Disclaimers.

If the Modified Version includes new front-matter sections or appendices that qualify as Secondary Sections and contain no material copied from the Document, you may at your option designate some or all of these sections as invariant. To do this, add their titles to the list of Invariant Sections in the Modified Version's license notice. These titles must be distinct from any other section titles.

You may add a section Entitled "Endorsements", provided it contains nothing but endorsements of your Modified Version by various parties—for example, statements of peer review or that the text has been approved by an organization as the authoritative definition of a standard.

You may add a passage of up to five words as a Front-Cover Text, and a passage of up to 25 words as a Back-Cover Text, to the end of the list of Cover Texts in the Modified Version. Only one passage of Front-Cover Text and one of Back-Cover Text may be added by (or through arrangements made by) any one entity. If the Document already includes a cover text for the same cover, previously added by you or by arrangement made by the same entity you are acting on behalf of, you may not add another; but you may replace the old one, on explicit permission from the previous publisher that added the old one.

The author(s) and publisher(s) of the Document do not by this License give permission to use their names for publicity for or to assert or imply endorsement of any Modified Version.

5. COMBINING DOCUMENTS

You may combine the Document with other documents released under this License, under the terms defined in section 4 above for modified versions, provided that you include in the combination all of the Invariant Sections of all of the original documents, unmodified, and list them all as Invariant Sections of your combined work in its license notice, and that you preserve all their Warranty Disclaimers.

The combined work need only contain one copy of this License, and multiple identical Invariant Sections may be replaced with a single copy. If there are multiple Invariant Sections with the same name but different contents, make the title of each such section unique by adding at the end of it, in parentheses, the name of the original author or publisher of that section if known, or else a unique number. Make the same adjustment to the section titles in the list of Invariant Sections in the license notice of the combined work.

In the combination, you must combine any sections Entitled “History” in the various original documents, forming one section Entitled “History”; likewise combine any sections Entitled “Acknowledgements”, and any sections Entitled “Dedications”. You must delete all sections Entitled “Endorsements”.

6. COLLECTIONS OF DOCUMENTS

You may make a collection consisting of the Document and other documents released under this License, and replace the individual copies of this License in the various documents with a single copy that is included in the collection, provided that you follow the rules of this License for verbatim copying of each of the documents in all other respects.

You may extract a single document from such a collection, and distribute it individually under this License, provided you insert a copy of this License into the extracted document, and follow this License in all other respects regarding verbatim copying of that document.

7. AGGREGATION WITH INDEPENDENT WORKS

A compilation of the Document or its derivatives with other separate and independent documents or works, in or on a volume of a storage or distribution medium, is called an “aggregate” if the copyright resulting from the compilation is not used to limit the legal rights of the compilation’s users beyond what the individual works permit. When the Document is included in an aggregate, this License does not apply to the other works in the aggregate which are not themselves derivative works of the Document.

If the Cover Text requirement of section 3 is applicable to these copies of the Document, then if the Document is less than one half of the entire aggregate, the Document’s Cover Texts may be placed on covers that bracket the Document within the aggregate, or the electronic equivalent of covers if the Document is in electronic form. Otherwise they must appear on printed covers that bracket the whole aggregate.

8. TRANSLATION

Translation is considered a kind of modification, so you may distribute translations of the Document under the terms of section 4. Replacing Invariant Sections with translations requires special permission from their copyright holders, but you may include translations of some or all Invariant Sections in addition to the original versions of these Invariant Sections. You may include a translation of this License, and all the license notices in the Document, and any Warranty Disclaimers, provided that you also include the original English version of this License and the original versions of those notices and disclaimers. In case of a disagreement between the translation and the original version of this License or a notice or disclaimer, the original version will prevail.

If a section in the Document is Entitled “Acknowledgements”, “Dedications”, or “History”, the requirement (section 4) to Preserve its Title (section 1) will typically require changing the actual title.

9. TERMINATION

You may not copy, modify, sublicense, or distribute the Document except as expressly provided under this License. Any attempt otherwise to copy, modify, sublicense, or distribute it is void, and will automatically terminate your rights under this License.

However, if you cease all violation of this License, then your license from a particular copyright holder is reinstated (a) provisionally, unless and until the copyright holder explicitly and finally terminates your license, and (b) permanently, if the copyright holder fails to notify you of the violation by some reasonable means prior to 60 days after the cessation.

Moreover, your license from a particular copyright holder is reinstated permanently if the copyright holder notifies you of the violation by some reasonable means, this is the first time you have received notice of violation of this License (for any work) from that copyright holder, and you cure the violation prior to 30 days after your receipt of the notice.

Termination of your rights under this section does not terminate the licenses of parties who have received copies or rights from you under this License. If your rights have been terminated and not permanently reinstated, receipt of a copy of some or all of the same material does not give you any rights to use it.

10. FUTURE REVISIONS OF THIS LICENSE

The Free Software Foundation may publish new, revised versions of the GNU Free Documentation License from time to time. Such new versions will be similar in spirit to the present version, but may differ in detail to address new problems or concerns. See http://www.gnu.org/copyleft/.

Each version of the License is given a distinguishing version number. If the Document specifies that a particular numbered version of this License “or any later version” applies to it, you have the option of following the terms and conditions either of that specified version or of any later version that has been published (not as a draft) by the Free Software Foundation. If the Document does not specify a version number of this License, you may choose any version ever published (not as a draft) by the Free Software Foundation. If the Document specifies that a proxy can decide which future versions of this License can be used, that proxy’s public statement of acceptance of a version permanently authorizes you to choose that version for the Document.

11. RELICENSING

“Massive Multiauthor Collaboration Site” (or “MMC Site”) means any World Wide Web server that publishes copyrightable works and also provides prominent facilities for anybody to edit those works. A public wiki that anybody can edit is an example of such a server. A “Massive Multiauthor Collaboration” (or “MMC”) contained in the site means any set of copyrightable works thus published on the MMC site.

“CC-BY-SA” means the Creative Commons Attribution-Share Alike 3.0 license published by Creative Commons Corporation, a not-for-profit corporation with a principal place of business in San Francisco, California, as well as future copyleft versions of that license published by that same organization.

“Incorporate” means to publish or republish a Document, in whole or in part, as part of another Document.

An MMC is “eligible for relicensing” if it is licensed under this License, and if all works that were first published under this License somewhere other than this MMC, and subsequently incorporated in whole or in part into the MMC, (1) had no cover texts or invariant sections, and (2) were thus incorporated prior to November 1, 2008.

The operator of an MMC Site may republish an MMC contained in the site under CC-BY-SA on the same site at any time before August 1, 2009, provided the MMC is eligible for relicensing.