1 Multivariable Functions

1.1 Functions, twospace, and threespace

Recall that a *function* is a rule that takes an input and assigns a specific output. Note that a function always gives exactly one output, and always gives the same output for a given input.

Importantly, we haven't specified what types of outputs functions have—in the abstract, a function can take any type of input and give any type of output. We say that the set of possible inputs for a function is the *domain*.

In your previous calculus courses you have seen "single–variable" functions, which take in a single real number and output a single real number. In this course we will study functions of "multiple variables". That is, either their input, or their output, or both will consist of more than one real number.

Example 1.1.
- Altitude as a function of latitude and longitude
- Gas pressure as a function of volume and temperature
- Temperature as a function of latitude and longitude and altitude
- Volume of a box as a function of length and width and height
- Position as a functionof time
- Direction and distance to your destination as a function of where you are

Definition 1.2. We use $\mathbb{R}$ to denote the set of all real numbers.

We use $\mathbb{R}^2 = \{(x, y) : x, y \in \mathbb{R}\}$ to denote the set of ordered pairs of real numbers. This is sometimes called the *real plane* or the *cartesian plane*.

We use $\mathbb{R}^3 = \{(x, y, z) : x, y, z \in \mathbb{R}\}$ to denote the set of ordered triples of real numbers. This is sometimes called *real threespace*.

Remark 1.3. We can use $\mathbb{R}^n$ to refer to the set of ordered sets of n real numbers. This is an n-dimensional space. Everything we do in this course can be generalized to work in $\mathbb{R}^n$ rather than $\mathbb{R}^2$ or $\mathbb{R}^3$, but in this course we're going to confine ourselves to the two- and three-dimensional cases.

This is mainly because those cases are easier to draw. But it's also harder to talk about the higher-dimensional cases without linear algebra.

Geometrically, we think of $\mathbb{R}^2$ as representing a plane; the first real number gives the x-coordinate, or horizontal distance from the origin; the second gives the y-coordinate or vertical distance.

Similarly, $\mathbb{R}^3$ represents three-dimensional space, with the three numbers giving the distances along three perpendicular axes. We put these axes together according to the *right-hand rule*, which says that if we look **from** the positive side of the z-axis we see the xy-plane in the usual orientation.

Example 1.4.

We can compute the *distance* between two points using the pythagorean theorem. If (x, y) and (a, b) are two points in $\mathbb{R}^2$, then the distance between them is given by $\sqrt{(x-a)^2 + (y-b)^2}$. Similarly, if (x, y, z) and (a, b, c) are two points in $\mathbb{R}^3$, then the distance between them is given by $\sqrt{(x-a)^2 + (y-b)^2 + (z-c)^2}$.

Example 1.5. The distance between $(3, 5)$ and $(1, 7)$ is

$$\sqrt{(3-1)^2 + (5-7^2)} = \sqrt{4+4} = 2\sqrt{2}.$$

The distance betweent $(1, -1, 0)$ and $(3, 2, 1)$ is

$$\sqrt{(1-3)^2 + (-1-2)^2 + (0-1)^2} = \sqrt{4+9+1} = \sqrt{14}.$$

Example 1.6. Suppose we want to describe a sphere in threespace. The unit sphere is the set of all points of distance one from the origin. That is, it's the set of all points (x, y, z) such that $\sqrt{(x-0)^2 + (y-0)^2 + (z-0)^2} = 1$. We can simplify this equation and see that the unit sphere is precisely the set of points that satisfy the equation $x^2 + y^2 + z^2 = 1$.

More generally, if we have a sphere of radius r centered at the point (a, b, c), then it is the set of all points of distance r from (a, b, c). So it's the set of all points such that $\sqrt{(x-a)^2 + (y-b)^2 + (z-c)^2} = r$, or in other words $(x-a)^2 + (y-b)^2 + (z-c)^2 = r^2$.

In this course we'll think about four or five basic categories of functions. This taxonomy is a bit artificial but it will still be useful to think about:

Type of function	Name	Example of use
$\mathbb{R} \to \mathbb{R}$	Single-variable functions	
$\mathbb{R}^2$ or $\mathbb{R}^3 \to \mathbb{R}$	Multivariable function	Temperature as a function of position
$\mathbb{R} \to \mathbb{R}^2$ or $\mathbb{R}^3$	Parametrization of a path	Position as a function of time
$\mathbb{R}^2 \to \mathbb{R}^3$	Parametrized surface	
$\mathbb{R}^2 \to \mathbb{R}^2$ or $\mathbb{R}^3 \to \mathbb{R}^3$	Vector field	Direction of a force as a function of position

You are already familiar with single-variable functions from your previous calculus courses. For the first part of this course we're going to focus on multivariable functions, that take in multiple variables but output a single variable. In the second half of the course we will look at functions that output more than one variable as well.

1.2 Graphing multivariable functions

To describe and understand single-variable functions, we would draw a graph, with one dimension representing the input and one dimension representing the output. We would like to do the same thing for multivariable functions, but the situation is a bit more difficult because it's much harder to draw three-dimensional pictures. (And all but impossible to draw four- or six-dimensional pictures).

1.2.1 Graphing functions of two variables as surfaces

Recall that when we graphed a single-variable function f, we took all the points (x, y) such that $y = f(x)$. Similarly, we can define:

Definition 1.7. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be a function of two variables. Then the *graph* of f is the set $\{(x, y, z) : z = f(x, y)\}$ of all points with $z = f(x, y)$.

The graph of a two-variable function will generally look like a curved two-dimensional surface in three-dimensional space.

A graph of a two-variable function will still have to pass the vertical line test: a vertical line given by $x = a, y = b$ will intersect the surface in at most one point. This is because a given (x, y) input has only one output.

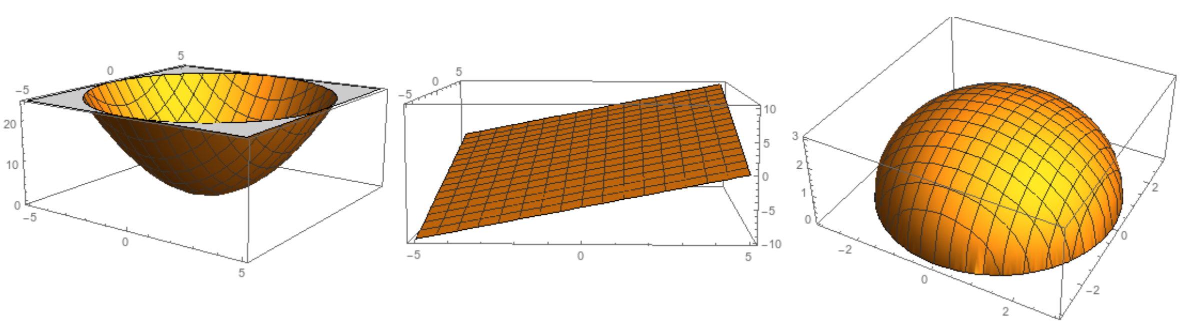

Figure 1.1: Graphs of the functions $x^2 + y^2$, $x + y$, and $\sqrt{9 - x^2 - y^2}$

1.2.2 Transformations of two-variable functions

If you already know the graph of one function, you can often figure out what the graphs of related functions must look like.

- The graph of $f(x, y)+c$ is the graph of $f(x, y)$ shifted up (along the z-axis) by c units.
- The graph of $f(x-a, y-b)$ is the same as the graph of $f(x, y)$ but shifted a units along the x axis and b units along the y axis. You can think of this as moving the center of the graph from $(0,0)$ to (a, b).
- The graph of $f(-x, y)$ is the graph of $f(x, y)$ reflected across the yz plane, inverting the x axis.
- The graph of $f(x,-y)$ is the graph of $f(x, y)$ reflected across the xz plane, inverting the y axis.
- The graph of $-f(x, y)$ is the graph of $f(x, y)$ reflected across the xy plane, inverting the z axis and drawing the graph "upside down".

Example 1.8. Let's consider the function $f(x, y)=x^2+y^2$ that we saw in figure 1.1. Then we can look at the following ways of shifting the function in figure 1.2:

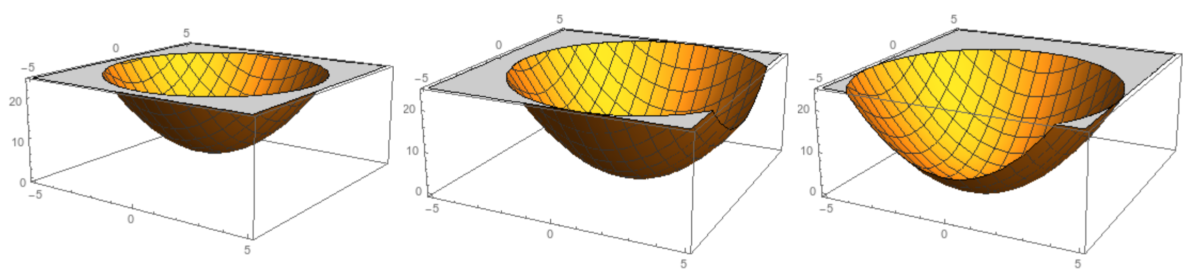

Figure 1.2: The graphs of $f(x, y)+5$, $f(x-1, y)$, and $f(x, y+2)$

Similarly, we can take the function $g(x, y)=x+5y$ and look at the following graphs in figure 1.3:

1.2.3 Graphing two-variable functions with cross-sections

We still don't have a good way to figure out what the graph of a two-variable function looks like if we don't already know. But the last section gives us an idea: look at each variable individually.

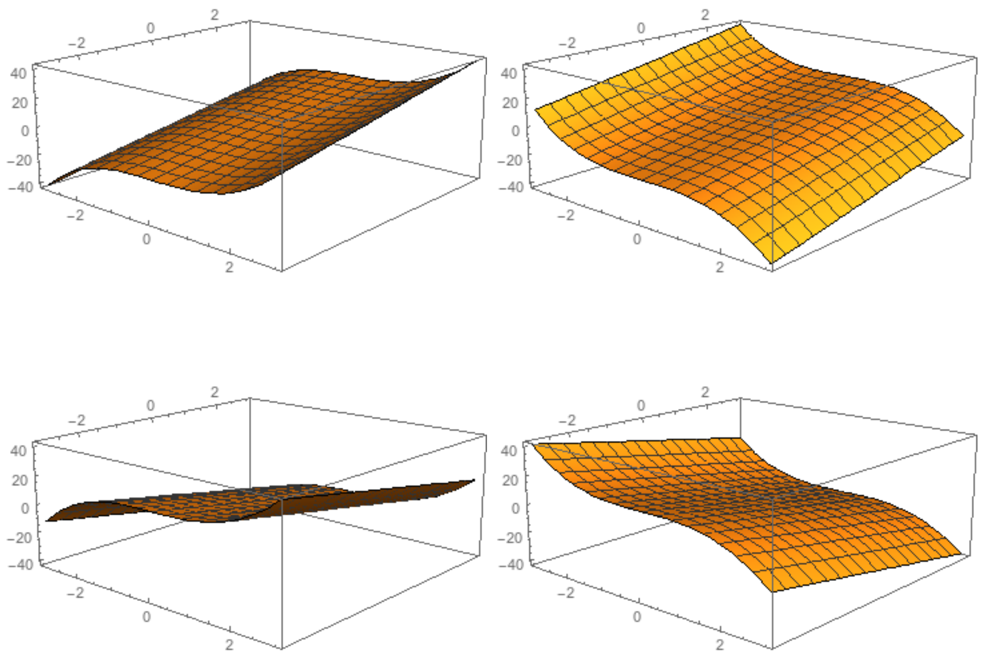

Figure 1.3: The graphs of $g(x, y)$, $g(-x, y)$, $g(x, -y)$, and $-g(x, y)$

Definition 1.9. If $f(x, y)$ is a function of two variables, then we can get a function of one-variable by fixing $x = c$ and considering the function $f(c, y)$. This function is called a *cross-section* of f with x fixed. The graph of this cross-section is the curve given by intersecting the plane $x = c$ with the graph of $f(x, y)$.

Similarly, the function of one variable given by $f(x, c)$ is a cross-section of f with y fixed. The graph of this function is the curve given by intersecting the plane $y = c$ with the graph of $f(x, y)$.

Each cross-section is a single-variable function, and thus straightforward to graph. By graphing a number of cross sections we can get a good idea what the graph of the whole function looks like.

Example 1.10. Let $f(x, y) = x^2 - y^2$. First we'll take cross-sections holding y constant. We can plot these below in figure 1.4:

Thus we see that the cross-sections holding y constant are parabolas, which start lower and lower the further away we get from the $y = 0$ plane.

We can also take cross-sections holding x constant. We get the similar graphs in figure 1.5:

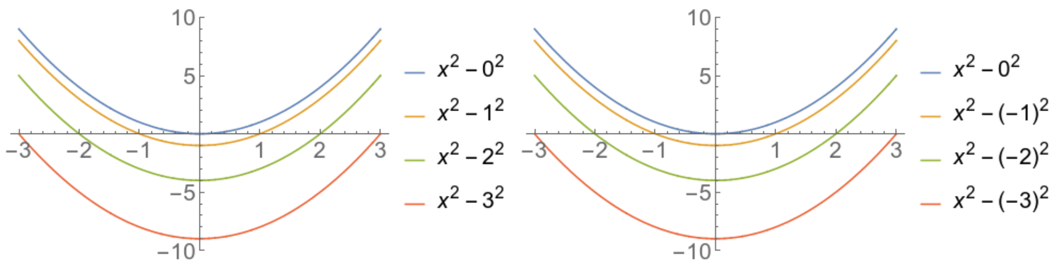

Figure 1.4: Cross sections of $x^2 - y^2$ holding y constant

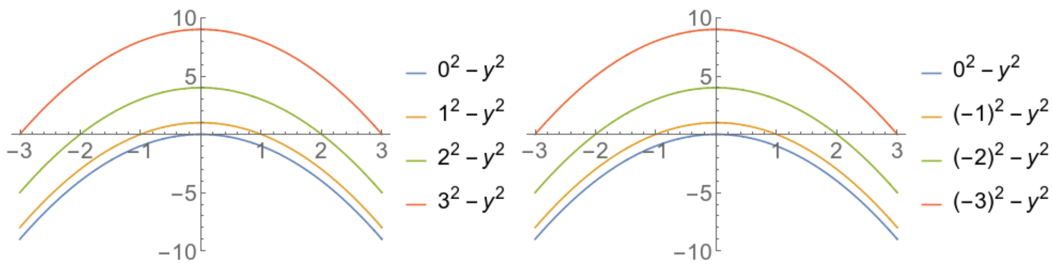

Figure 1.5: Cross sections of $x^2 - y^2$ holding x constant

These show us that holding x constant, we get upside-down parabolas, with the peak being higher and higher the farther we are from the plane $x = 0$.

Putting this together, we can assemble a picture of the real function:

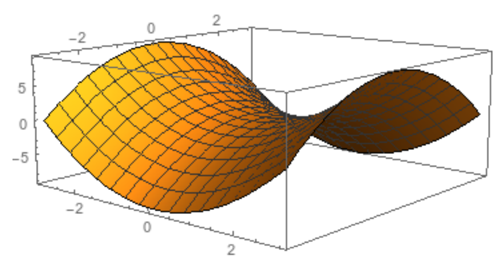

Figure 1.6: The graph of $x^2 - y^2$

Example 1.11. Let $g(x, y) = x^3 + \sin(y)$. We can again take cross sections, holding x and y constant in turn:

From the left, we see that holding x constant, we have a gentle sine wave along the y axis. From the right, we see that holding y constant, x is increasing in a cubic. Putting this information together, we can get a graph for the whole surface:

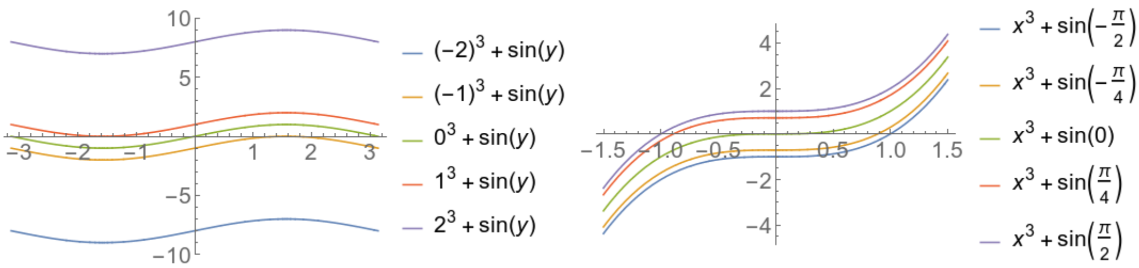

Figure 1.7: Cross sections of $x^3 + \sin(y)$, holding x constant on the left and y constant on the right

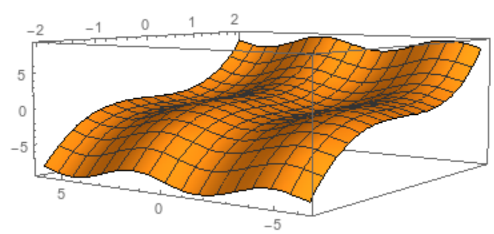

Figure 1.8: The graph of $x^3 + \sin(y)$

1.2.4 Graphing two-variable functions with level sets

Sometimes we want to approach the same question from a different direction (literally!). Instead of holding x constant or y constant, we will hold z constant.

Definition 1.12. If $f(x, y)$ is a function of two variables, then the *level set* of f at level c is the set of all points (x, y) such that $f(x, y) = c$.

A *contour diagram* for f is a graph with several level sets for f at different levels.

Importantly, the level set is not a function, and doesn't need to pass any vertical line tests or anything similar.

Contour diagrams show up commonly in topographical maps.

Example 1.13. The contour plots in figure 1.9 look very similar, but the contour heights make them very different. We can see the corresponding graphs in figure 1.10.

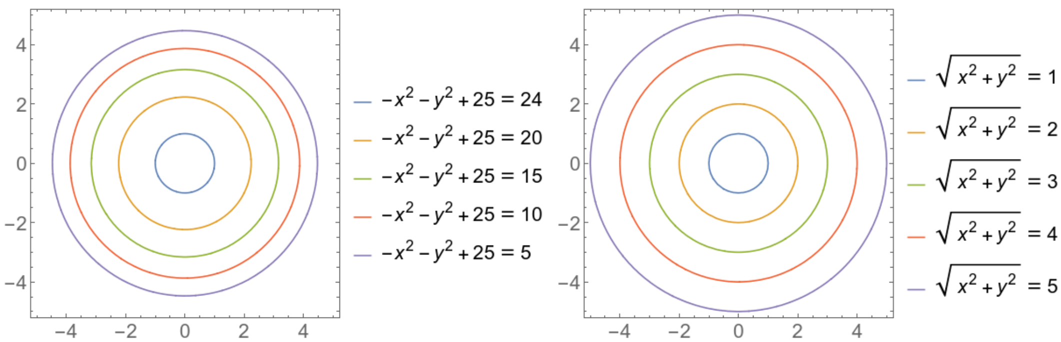

Figure 1.9: Contour diagrams for $f(x, y) = 25 - x^2 - y^2$ and $g(x, y) = \sqrt{x^2 + y^2}$

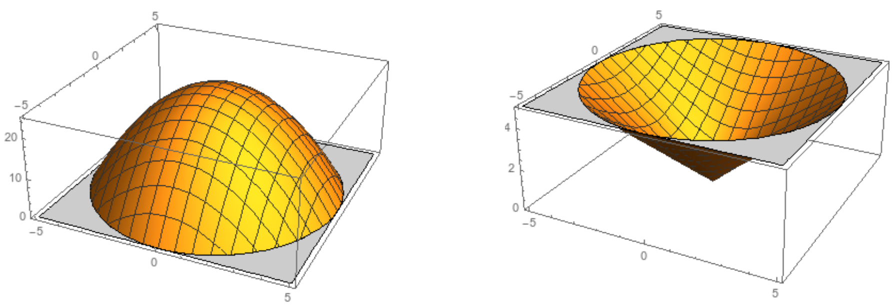

Figure 1.10: The graphs of $f(x, y) = 25 - x^2 - y^2$ and $g(x, y) = \sqrt{x^2 + y^2}$

Example 1.14. We can also draw contour plots for some of our earlier functions. The contour plot for the saddle from example 1.10 and the sine function from example 1.11 appear in figure 1.11.

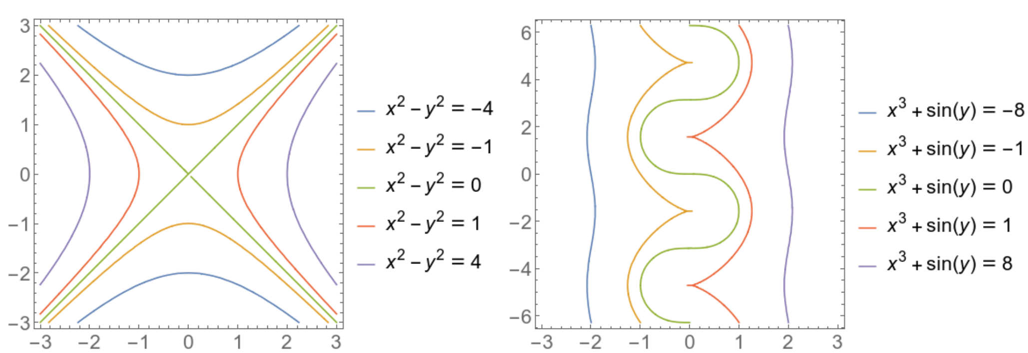

Figure 1.11: Contour plots for $x^2 - y^2$ and $x^3 + \sin(y)$

1.2.5 Graphing three-variable functions with level surfaces

We've now established a few approaches to graphically representing functions of two variables. What can we do with functions of three variables?

Simply graphing the entire function isn't really a plausible solution. As a mathematical object, the graph of a three-variable function as a subset of $\mathbb{R}^4$ is perfectly well defined; but it's almost impossible to draw or visualize these graphs, so they don't help us with our problem of visually representing three-variable functions.

In contrast, cross-sections and level sets are both useful tools. They are much tricker to implement here, because the cross-sections and level sets will themselves be two-variable functions, and thus give us two-dimensional surfaces inside threespace.

Definition 1.15. If $f(x, y, z)$ is a function of two variables, then the *level set* of f at level c is the set of all points (x, y, z) such that $f(x, y, z) = c$.

It's much harder to draw a contour diagram in this case, but we can sort of make an attempt still.

Example 1.16. Find the level surfaces of $f(x, y, z) = x^2 + y^2 + z^2$.

There are no surfaces for $c < 0$, and for $c = 0$ the level surface is a point. For larger c we get a sphere of radius $\sqrt{c}$. Thus the level sets for $c = 1, 4, 9$ are shown in figure 1.12.

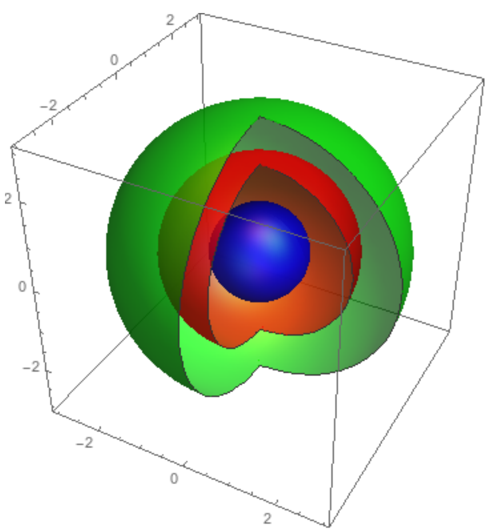

Figure 1.12: Level sets for $x^2 + y^2 = z^2$ at the levels $c = 1, 4, 9$

Example 1.17. We can see the level surfaces of $g(x, y, z) = x^2 + y^2$ and $h(x, y, z) = x + z$, at the levels $1, 2, 3, 4$, in figure 1.13. Thbe level surfaces for g are cylinders of radius $\sqrt{c}$, and the level surfaces of h are all parallel planes.

Example 1.18. We'd like to understand the level surfaces of $f(x, y, z) = x^2 + y^2 - z^2$. These will look different depending on the level of c.

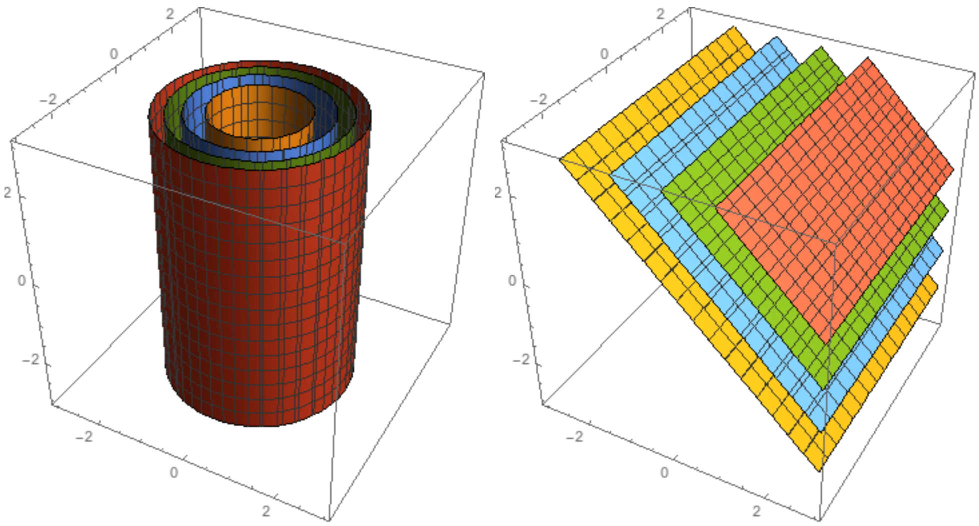

Figure 1.13: The level surfaces of g and h at the levels $1, 2, 3, 4$

It's probably easiest to think about these level surfaces by thinking about their own contour plots as z varies. If $c = 0$, then our equation is $x^2 + y^2 = z^2$. We see that for each z we get a circle of radius z in the plane perpendicular to the z-axis, and in fact at $z = 0$ we have a single point. Stacking these all together gives us two cones.

If c is positive, then we have the equation $x^2 + y^2 = z^2 + c$. Then we see that for each z we get a circle of radius $\sqrt{z^2 + c} > z$, and the radius will always be positive. If instead we take, say, the $x = 0$ cross-section, we get $y^2 - z^2 = c$, which is a hyperbola. The resulting surface is a *hyperboloid of one sheet*.

Finally, if c is negative, we get $x^2+y^2 = z^2+c$, where there is no solution when $z^2+c < 0$. Thus we'll have a stack of increasing radius circles, but it will start at $z = \pm\sqrt{c}$. This surface is a *hyperboloid of two sheets*.

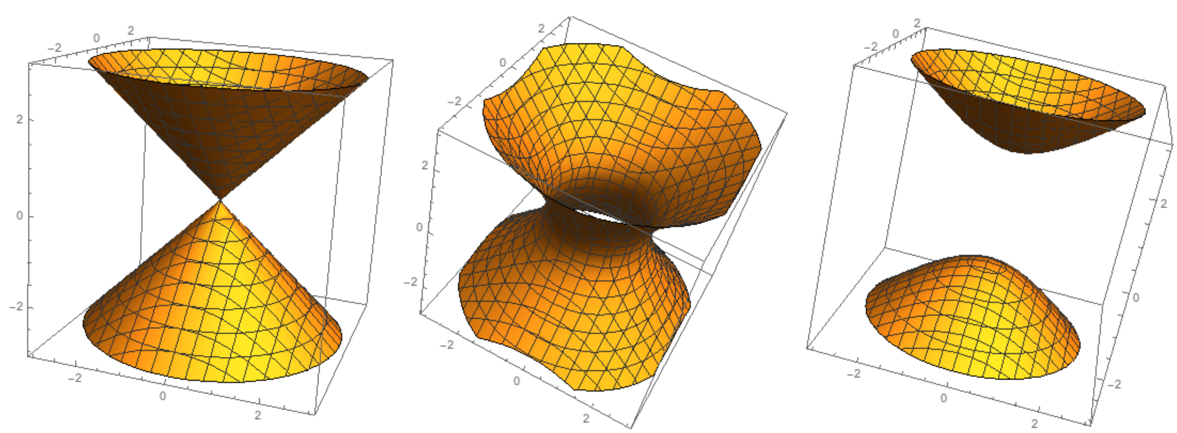

Figure 1.14: Level surfaces of $x^2 + y^2 - z^2$ at the levels $0, 2, -2$

You can see some helpful and common surfaces on page 702 of the textbook, at the end of section 12.5.

Remark 1.19. We've used surfaces to represent the full graph of two-variable functions, and also to represent the level surfaces of three-variable functions. These surfaces are at least somewhat related, and in fact if we have the graph of a function $f(x, y)$, then it is also the level surface at zero of the function $f(x, y) - z$.

Thus every graph of a two-variable function is also a level surface of some three-variable function. The converse, however, is not true; many of the level surfaces we have seen cannot be the graphs of two-variable functions, since they fail the vertical line test.

1.3 Linear Functions

In single-variable calculus, we used the derivative to approximate functions with their tangent lines. We want to do the same sort of approximation for functions of multiple variables; first we need to understand the analogue of a line. We focus on the two-variable case, but our ideas here have obvious generalizations.

Definition 1.20. A *plane* is defined by the equation $z = z_0 + m(x - x_0) + n(y - y_0)$. This plane passes through the point (x_0, y_0, z_0), and we say it has slope m in the x-direction and slope n in the y-direction.

Equivalently, a plane is given by the equation $z = c + mx + ny$ (where $c = z_0 - mx_0 - ny_0$.

The plane is the graph of the *linear function* $f(x, y) = z_0 + m(x - x_0) + n(y - y_0)$.

Remark 1.21. If you have taken linear algebra, you will notice that this is somewhat different from the definition of a linear function given there. A linear function in the linear algebra sense must also pass through the origin, and thus the equation can always be written $f(x, y) = mx + ny$.

Thus technically we have defined an "affine transformation" rather than a linear transformation. But under the same technicality, most lines in single-variable calculus are not linear functions. We'll mostly ignore that here.

Just as a line in the plane is determined by two points, a plane in threespace is determined by three points. The first point gives us (x_0, y_0, z_0), and the other two points tell us the slopes m and n.

Example 1.22. Find the equation of hte plane passing through the points $(1, 0, 1), (1, -1, 3), (3, 0, -1)$.

The first two points have the same x-coordinate, so we can use them to find the y slope. We see that $n = \frac{3-1}{-1-0} = -2$. Then we have

$$\begin{aligned} f(x,y) &= 1 + m(x-1) - 2(y-0) \\ -1 &= 1 + m(3-1) - 2(0-0) = 2m+1 \end{aligned}$$

so $m = -1$ and the equation for the plane is

$$f(x,y) = 1 - (x-1) - 2y = 2 - x - 2y.$$

We can check our work by plugging all three points back into this equation and confirming that they work.

Example 1.23. Find an equation for the plane going through $(-1,2,3)$, $(1,5,2)$, and $(3,4,1)$.

This time we don't have any conveniently unchanged coordinates. So instead we write

$$\begin{aligned} f(x,y) &= 3 + m(x+1) + n(y-2) \\ 2 &= 3 + m(1+1) + n(5-2) = 3 + 2m + 3n \\ 1 &= 3 + m(3+1) + n(4-2) = 3 + 4m + 2n \\ 2m + 3n &= -1 \\ 4m + 2n &= -2 \\ -4n &= 0 \\ n &= 0 \\ m &= -1/2 \\ f(x,y) &= 3 - \frac{1}{2}(x+1). \end{aligned}$$

The important thing about linear functions is that changes in x and changes in y change the output completely independently. This makes everything about the functions very simple.

Example 1.24.

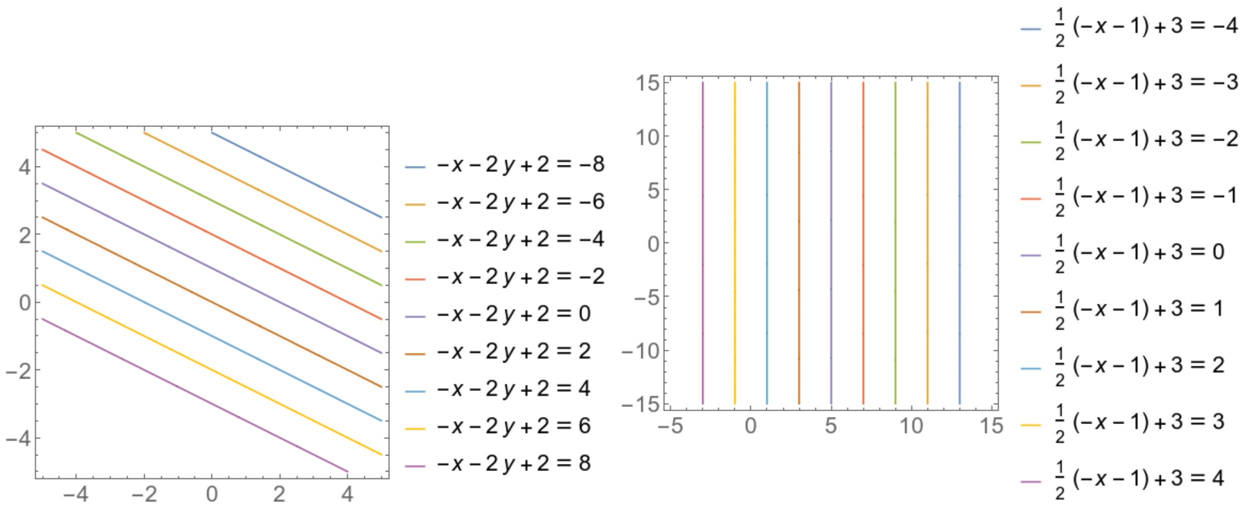

Figure 1.15: Contour diagrams for $f(x, y) = 2 - x - 2y$ and $g(x, y) = 3 - (x + 1)/2$

1.4 Limits and Continuity

In calculus 1, we learned about limits, which tell us in some sense the value a function "should" have at a point—which may or may not be the value it does have, and it may not have a value at all. We can extend the same idea to multivariable functions.

Definition 1.25. If $f : \mathbb{R}^2 \to \mathbb{R}$ is a function, then it has a limit at the point (a, b) of L, and we write

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

if we can make $f(x, y)$ as close as we want to L, purely by requiring the distance from (x, y) to (a, b) to be small enough (but not zero).

Remark 1.26. Formally we'd write something like: for every $\epsilon > 0$, there is a $\delta > 0$ such that if $\sqrt{(x-a)^2 + (y-b)^2} < \delta$ then $|f(x, y) - L| < \epsilon$. We won't be drilling down into that level of rigor in this class, though. If you want to see more of this kind of thing, take Math 310.

Example 1.27. When the function doesn't do anything weird, limits won't do anything surprising.

$\lim_{(x,y)\to(3,4)} x^2 + y^2 = 3^2 + 4^2 = 25$
$\lim_{(x,y)\to(1,-1)} x^2 - y^2 = 1^2 - (-1)^2 = 0$

Definition 1.28. A function f is *continuous* at a point (a, b) if

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

A function is *continuous on a region* R if it is continuous at each point in R.

If f is not continuous at a point (a, b) then it is *discontinuous* there.

Fact 1.29. *A function defined entirely from algebraic, trigonometric, and exponential functions is continuous anywhere it is defined.*

(No function is ever continuous anywhere it is not defined).

Example 1.30. Let $f(x, y) = \frac{x^2 y}{x^2+y^2}$. This function is continuous everywhere it is defined, which is everywhere except $(0, 0)$. So it's easy to compute, for instance, that $\lim_{(x,y)\to(1,1)} f(x, y) = \frac{1^2 \cdot 1}{1^2+1^2} = \frac{1}{2}$.

Now let's consider $\lim_{(x,y)\to(0,0)} f(x, y)$. We can't just plug $(0, 0)$ in here, so we need to do something else.

First, we can look at the graph and contour diagram of f.

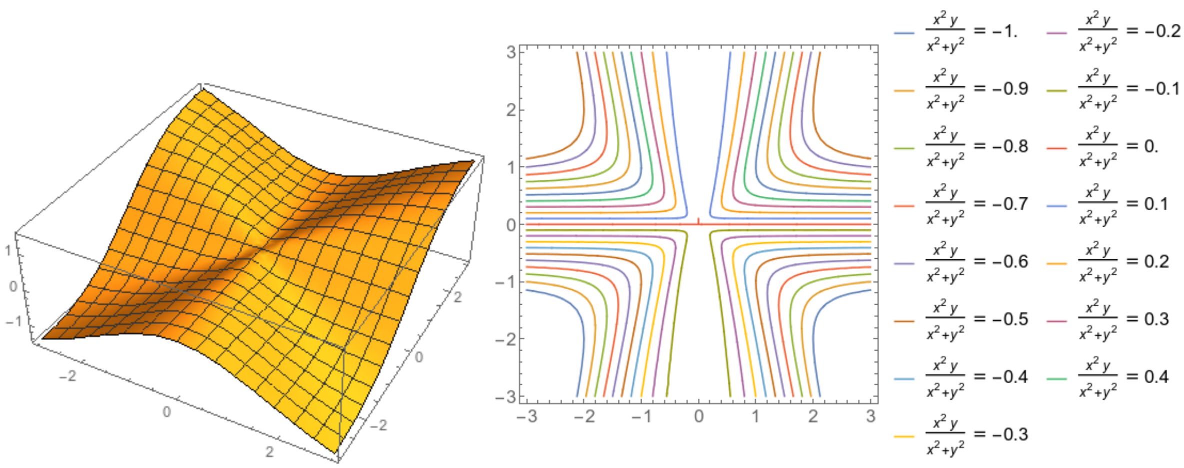

Figure 1.16: The graph and contour plot for $f(x, y) = \frac{x^2 y}{x^2+y^2}$

We see that the function seems to smoothly approach 0 as (x, y) approaches $(0, 0)$, so we suspect the limit is in fact 0.

Informally, we see that the denominator "goes to zero" "twice", while the numerator goes to zero "three times". Thus we would expect the limit to be zero.

If we want to be more rigorous, we calculate the distance between $f(x, y)$ and the guessed limit 0. Then we have

$$|f(x, y) - L| = \left| \frac{x^2 y}{x^2 + y^2} - 0 \right| = \left| \frac{x^2}{x^2 + y^2} \right| |y| \leq |y| \leq \sqrt{x^2 + y^2}.$$

Thus the distance between $f(x, y)$ and 0 is less than the distance between (x, y) and $(0, 0)$. Clearly by making (x, y) closer to $(0, 0)$ we can make $f(x, y)$ as close as we want to 0.

Since the limit exists, we can extend this function to be continuous at the origin: the function

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

is continuous at $(0,0)$.

Example 1.31. Now let's consider the similar function $g(x,y) = \frac{x^2}{x^2+y^2}$. Like in example 1.30, this is continuous everywhere it is defined, which is everywhere except at $(0,0)$.

But at $(0,0)$ things are tricker. The graph has a noticeable spike, and the contour plot looks terrible near $(0,0)$, with all the contours converging onto that single point.

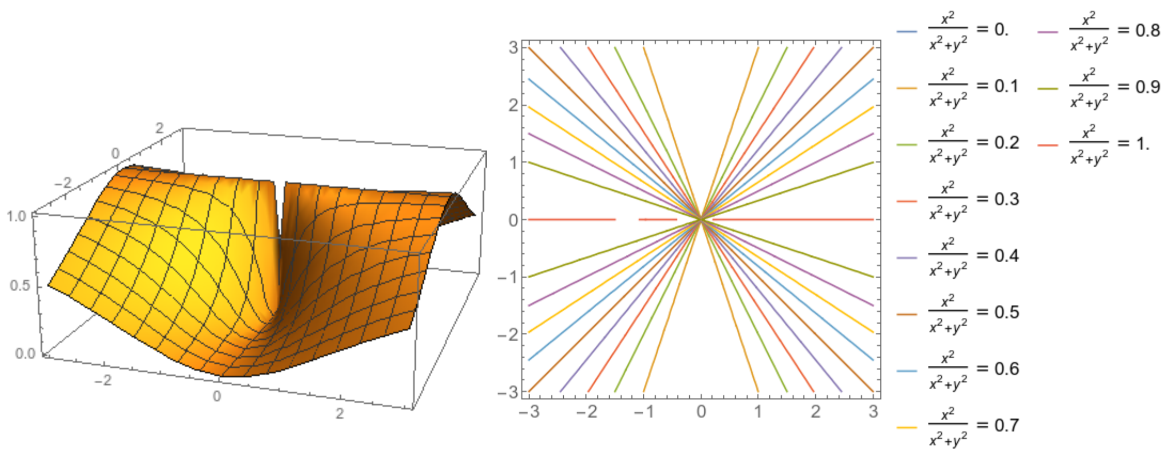

Figure 1.17: The graph and contour plot for $g(x,y) = \frac{x^2}{x^2+y^2}$

The informal algebraic argument we gave before doesn't help: both the top and the bottom go to zero "twice". So this doesn't help us find any limit.

Formally, we want to show that no limit exists, so we want to show that you can be as close to $(0,0)$ as you want and still get very different answers for $g(x,y)$.

So first let's consider points that look like $(a,0)$. Then $g(a,0) = \frac{a^2}{a^2+0^2} = 1$. Since a can be as small as we want, this tells us that we can be as close to the origin as we want and have $g(x,y) = 1$.

But this doesn't mean the limit is 1! As an example, take points that look like $(0,b)$. Then $g(0,b) = \frac{0^2}{0^2+b^2} = 0$. Since b can be anything, this also tells us that we can be as close to the origin as we want, and have $g(x,y) = 0$. Thus no limit exists.

In fact, by approaching from the right direction, we can get any value between 0 and 1. And we can see this behavior both in the graph (which has an abrupt spike or dip near the origin), and in the contour plot (which shows us different directions of approach, and the values they will give).

We just saw that we can show that limits don't exist by approaching the same point from different directions. This should remind you of the one-variable case, where we might check the right- and left-sided limits and show they differ.

But the multivariable case is considerably more complex, because there are infinitely many directions. (In fact it's more complicated than that: there are functions that have a consistent limit as long as you approach along any straight-line path, but that break down when you approach along the right curve).

But if we want to understand this better, we'll need a language for talking about directions.