Calculus 2: Multivariable Calculus & Differential Equations

Birkbeck 2016-2017

Module Code: BUEM001S5

Lecturer: Dan McVeagh

Contents:

Acknowledgements: These notes are based on lecture notes written by Amarpreet Rattan, but all errors are of course my responsibility. Please let me know if you find any mistakes.

1 Functions and differentiation

We begin with a review of many of the key notions in single variable calculus. As we shall see, some of the definitions and theorems carry through nicely to the multi-variable case, while others require some new ideas.

1.1 Functions of one variable

In this subsection we recall the ideas of a function of one variable and the derivative of such a function.

Let $U \subseteq \mathbb{R}$. A **(real) function** $f : U \to \mathbb{R}$ determines for each $x \in U$ a unique real number denoted by $f(x)$. We say that f is a function of one variable because the argument of f is one real number x. The set U is called the **domain of definition** of f. Usually such a function f is given by a rule or formula which allows us to compute $f(x)$ for every $x \in U$. Often we do not specify the domain of definition when describing a function f: we just give the rule for determining $f(x)$ and assume that U is the largest possible subset of $\mathbb{R}$ for which this rule makes sense.

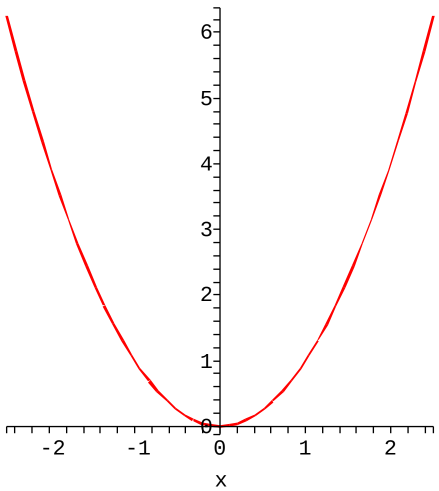
Example 1.1. If $f(x) = x^2$ then $U = \mathbb{R}$, since every real number has a square.

If $f(x) = \frac{1}{x}$ then $U = \mathbb{R} - \{0\}$.

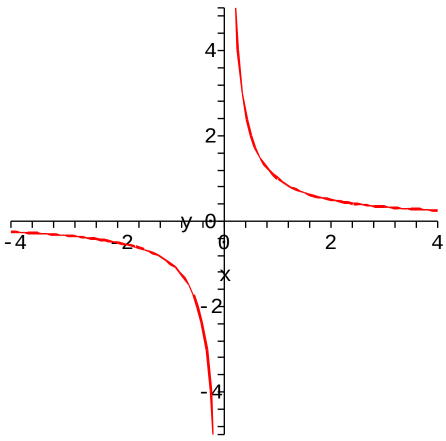
If $f(x) = \frac{(x+2)}{(x-3)(x+7)}$ then $U = \mathbb{R} - \{-7, 3\}$.

If $f(x) = \tan x$ then $U = \mathbb{R} - \left\{\frac{(2n+1)\pi}{2} : n \in \mathbb{Z}\right\}$.

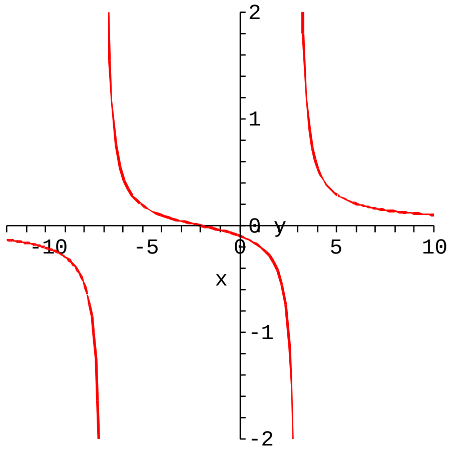
The **graph** associated with a function $f : U \to \mathbb{R}$ is the set of points $\{(x, f(x)) : x \in U\}$. Since the second coordinate of a point is usually denoted by y, we often express a function in the form $y = f(x)$. The following diagrams show the graphs of each of the functions given in Example 1.1.



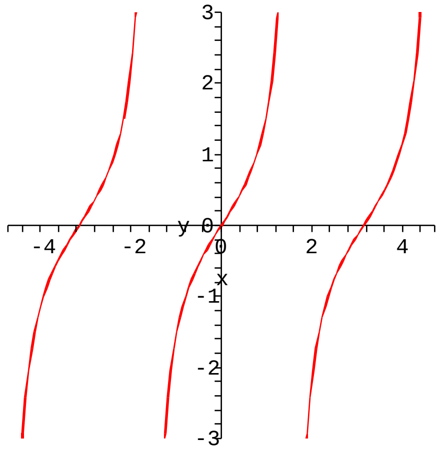
$y = x^2$



$y = \frac{1}{x}$



$y = \frac{(x+2)}{(x-3)(x+7)}$



$y = \tan x$

Differentiation

The **derivative** of a function $f : U \to \mathbb{R}$ at a point $a \in U$ is the gradient of the function at a.[1] It is denoted by $f'(a)$ or by $\dfrac{\mathrm{d}f}{\mathrm{d}x}$ at a.

The following table gives some standard functions and their derivative. (Remember when differentiating trigonometric functions the variable x is measured in radians.)

$f(x)$	$f'(x)$
x^n, $n \in \mathbb{R}$	nx^{n-1}
e^x	e^x
$\ln x$	$\frac{1}{x}$
$\sin x$	$\cos x$

$f(x)$	$f'(x)$
$\cos x$	$-\sin x$
$\tan x$	$\sec^2 x$
$\arcsin x$	$\frac{1}{\sqrt{1-x^2}}$
$\arctan x$	$\frac{1}{1+x^2}$

The derivative of more complicated functions can often be computed using the following rules of differentiation. Let $f, g : U \to \mathbb{R}$.

Linearity If $y(x) = f(x) + g(x)$ then $y'(x) = f'(x) + g'(x)$. If $y(x) = \alpha f(x)$ for a fixed number $\alpha \in \mathbb{R}$ then $y'(x) = \alpha f'(x)$. (These two statements are sometimes combined: if $y(x) = \alpha f(x) + \beta g(x)$ for $\alpha, \beta \in \mathbb{R}$ then $y'(x) = \alpha f'(x) + \beta g'(x)$.)

The product rule If $y(x) = f(x)g(x)$ then $y'(x) = f(x)g'(x) + g(x)f'(x)$ or alternatively $\dfrac{\mathrm{d}y}{\mathrm{d}x} = f\dfrac{\mathrm{d}g}{\mathrm{d}x} + g\dfrac{\mathrm{d}f}{\mathrm{d}x}$.

[1] Here we assume that the gradient at a exists. If the gradient does not exist at a then $f'(a)$ does not exist.

The quotient rule If $y(x) = \dfrac{f(x)}{g(x)}$, and $g(x) \neq 0$, then $y'(x) = \dfrac{g(x)f'(x) - f(x)g'(x)}{(g(x))^2}$ or alternatively $\dfrac{\mathrm{d}y}{\mathrm{d}x} = \dfrac{g\frac{\mathrm{d}f}{\mathrm{d}x} - f\frac{\mathrm{d}g}{\mathrm{d}x}}{g^2}$.

The chain rule If $y(x) = f(g(x))$ then $y'(x) = f'(g(x))g'(x)$ or alternatively $\dfrac{\mathrm{d}y}{\mathrm{d}x} = \dfrac{\mathrm{d}f}{\mathrm{d}g}\dfrac{\mathrm{d}g}{\mathrm{d}x}$. (For $f(g(x))$ to make sense, we must assume that $g(x)$ is contained in the domain of f.)

Exercises 1.1.

1. Find $\dfrac{\mathrm{d}y}{\mathrm{d}x}$ in each of the following cases.

   (a) $y = x^2 \sin x$

   (b) $y = \dfrac{x+1}{2x+3}$

   (c) $y = x^3 \ln x$

   (d) $y = \cot x$

   (e) $y = \ln(x^2 + 2)$

   (f) $y = \sin(3x - 5)$

   (g) $y = \dfrac{\sin 2x}{\cos x}$

   (h) $y = x^4 e^{2x}$

   (i) $y = \arctan(2\sqrt{x})$

   (j) $y = e^{\sin x}$

2. Find the derivative of each of the following functions.

   (a) $f(x) = x(x+1)^{10}$

   (b) $g(x) = \exp(x^2 + 2x + 3)$

   (c) $h(x) = x \ln(3x + 2)$

   (d) $k(x) = \dfrac{x \cos x}{x + 1}$

3. Let $y = x^4 \ln x$. Show that

$$x^2 \frac{\mathrm{d}^2 y}{\mathrm{d}x^2} - 7x\frac{\mathrm{d}y}{\mathrm{d}x} + 16y = 0.$$

1.2 Limits

Definition of limits

Let $f : U \to \mathbb{R}$, and let $a, l \in \mathbb{R}$. Informally, we say that $f(x)$ approaches l as x approaches a if the value of $f(x)$ gets closer to l as x gets closer to a. (Note that a does not necessarily have to be in U.) The formal definition is as follows.

Definition 1.2. We say that $f(x)$ approaches l as x approaches a if for all $\epsilon > 0$, there is an interval $(s,t) \subseteq U \cup \{a\}$, with $s < a < t$, such that $|f(x) - l| < \epsilon$ for all $x \in (s,t) - \{a\}$. If $f(x)$ approaches l as x approaches a then we write either $\lim\limits_{x \to a} f(x) = l$ or $f(x) \to l$ as $x \to a$.

We say that $\lim\limits_{x \to a} f(x)$ does not exist if there is no l satisfying the above definition.

Remark 1. The requirement that $|f(x) - l| < \epsilon$ can be rewritten as $-\epsilon < f(x) - l < \epsilon$. Thus if $f(x)$ approaches l as x approaches a then for all $\epsilon > 0$, there is an interval $(s,t) \subseteq U \cup \{a\}$ so that $f(x)$ is within ϵ of l for all $x \in (s,t) - \{a\}$.

Remark 2. Normally the interval (s,t) is chosen to be symmetrical about a. That is, $(s,t) = (a - \delta, a + \delta)$, for some $\delta > 0$. However, this is not strictly necessary.

Remark 3. The definition states that $|f(x) - l| < \epsilon$ for all $x \in (s,t) - \{a\}$, but it does not say anything about the function at $x = a$. Thus, when evaluating limits, we are interested in how the function behaves *near* a, not at a. In fact, when evaluating a limit of a function as x approaches a, the function doesn't even need to be defined at a itself. For example, for both functions in Figure 1, the limit as x approaches a exists and is equal to b. This is true even for the second function in Figure 1; even though the value of the function is not b at a, when x is close to a, the value of the function is close to b.

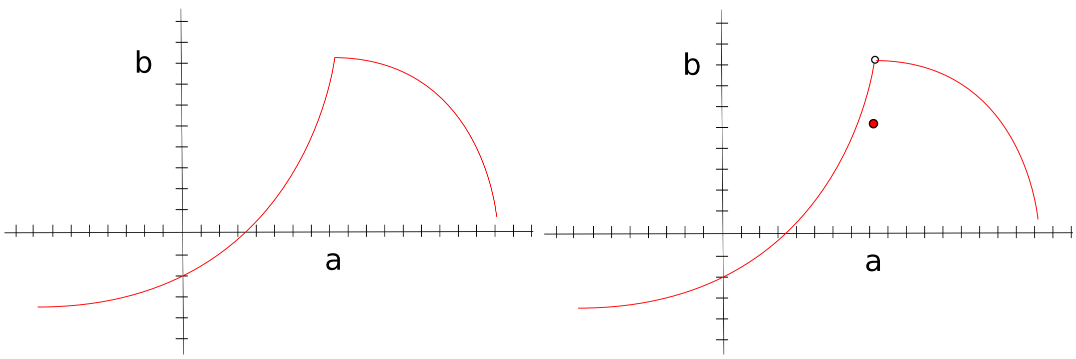


Figure 1: In the above two graphs of two slightly different functions, the limit as $x \to a$ does exist and it is b.

Remark 4. With the exception of the following example, we will not work with the technical definition of a limit in this module. We will, however, gain an understanding how to apply various rules to evaluate limits.

Example 1.3. Let $f : \mathbb{R} \to \mathbb{R}$ be the function $f(x) = x^2$. We want to use the definition of the limit to show that $\lim_{x\to 3} f(x) = 9$. To do this, for every $\epsilon > 0$ we have to find an interval $(s,t) \subseteq \mathbb{R}\cup\{3\} = \mathbb{R}$ such that $|f(x)-9| < \epsilon$ for all $x \in (s,t)-\{3\}$.

If $\epsilon \leq 9$ then we can take $(s,t) = (\sqrt{9-\epsilon}, \sqrt{9+\epsilon})$. Then for $x \in (s,t) - \{3\}$ we have $\sqrt{9-\epsilon} < x < \sqrt{9+\epsilon}$ and hence $9-\epsilon < x^2 < 9+\epsilon$. This implies $|x^2-9| < \epsilon$ as required.

If $\epsilon > 9$ then we can take $(s,t) = (0, 3\sqrt{2})$. Then for $x \in (s,t) - \{3\}$ we have $0 < x < 3\sqrt{2}$ and hence $0 < x^2 < 18$. This implies $-9 < x^2 - 9 < 9$ and hence $|x^2-9| < 9 < \epsilon$ as required.

Example 1.4. We have seen that $\lim_{x\to 3} x^2 = 9$. Note that if

$$g(x) = \begin{cases} 0 & \text{if } x = 3 \\ x^2 & \text{if } x \neq 3 \end{cases} \tag{1}$$

then also $\lim_{x\to 3} g(x) = 9$. This is because *near* 3, the function $g(x)$ behaves exactly like x^2, and the value of $g(x)$ at $x = 3$ is not important for the limit $\lim_{x\to 3} g(x)$.

A typical reason why a limit might fail to exist is because the function approaches different values when x approaches a from the left and right. For example this is the case for the function in Figure 2, so for this function the limit as $x \to a$ does not exist.

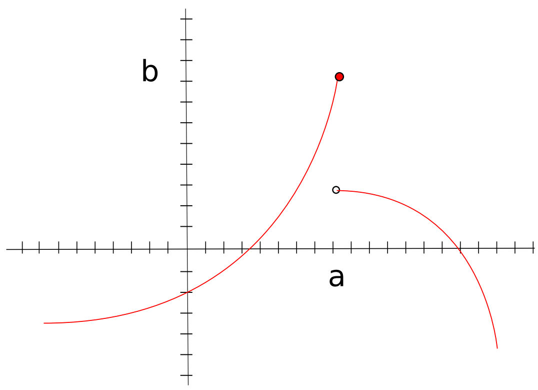


Figure 2: In the above graph, the limit as $x \to a$ does not exist.

Note that the above definition only applies when the limit l is finite; we write that

$$\lim_{x\to a} f(x) \to \infty \text{ (or } -\infty)$$

if $f(x)$ becomes arbitrarily large (or arbitrarily large negative) as x approaches a. For example, if $f(x) = 1/x^2$, then $f(x)$ gets larger and larger as $x \to 0$, therefore $\lim_{x\to 0} 1/x^2 \to \infty$. Note that $1/x^2$ is not defined at $x = 0$, but we could still evaluate the limit as $x \to 0$. However if $f(x) = 1/x$, then as x approaches 0 from the left, $f(x)$ approaches $-\infty$ and as x approaches 0 from the right, $f(x)$ approaches $+\infty$. Since these one sided limits disagree, we see that $\lim_{x\to 0} 1/x$ does not exist.

Some rules for evaluating limits

Let $f, g : U \to \mathbb{R}$ and suppose $\lim_{x\to a} f(x) = l$ and $\lim_{x\to a} g(x) = m$. Then

L1 $\lim_{x\to a} f(x) \pm g(x) = l \pm m$;

L2 $\lim_{x\to a} f(x)g(x) = lm$;

L3 $\lim_{x\to a} \dfrac{f(x)}{g(x)} = \dfrac{l}{m}$ provided $m \neq 0$.

L4 $\lim_{x\to a} \sqrt[n]{f(x)} = \sqrt[n]{l}$ (if n is even we assume that $f(x) \geq 0$ for all x near a).

Note that $\lim_{x\to a} f(x) = \lim_{x\to 0} f(x + a)$. All these properties can be proved using the definition of a limit; however, we will not do this in this module.

The observation in Remark 3 and Example 1.4 is also very useful for evaluating limits. We can formulate it as a rule as follows: Suppose that we wish to evaluate $\lim_{x\to a} f(x)$ and there is some function $g(x)$ and some interval $I = (b, c)$ containing a such that $f(x)$ and $g(x)$ agree on $I - \{a\}$. Then, $\lim_{x\to a} f(x) = \lim_{x\to a} g(x)$. Again, this is because it is immaterial what happens to f at a.

Evaluating limits in certain cases is straight forward. For example, we have $\lim_{x\to a} x = a$ and $\lim_{x\to a} c = c$ for any constant c. Using L2 repeatedly with $\lim_{x\to a} x = a$ gives $\lim_{x\to a} x^n = a^n$ for any positive integer n. From this, and using L1, it follows that if $f(x)$ is any polynomial or rational function (a ratio of two polynomials) and $f(a)$ is defined, then $\lim_{x\to a} f(x) = f(a)$.

Less straight forward is determining the limit of a rational function when the denominator has limit 0 as x approaches a, since the quotient rule L3 does not apply.

Example 1.5. Determine $\lim_{x\to 3} f(x)$, where $f(x) = \frac{x^2-x-6}{x-3}$.

We would like to use the quotient limit rule, but cannot immediately since the limit as $x \to 3$ of the denominator is 0. Notice that if the limit of the numerator were not zero, then the limit would be $\pm\infty$ or not exist. However, the limit of the numerator is 0, so we must dig deeper. Notice that we can factor the numerator into $(x+2)(x-3)$ and we can cancel the two factors of $x - 3$. Thus, we have $f(x) = x + 2$, $x \neq 3$. That is, $f(3)$ is not defined. However, using this we have

$$\lim_{x\to 3} \frac{x^2 - x - 6}{x - 3} = \lim_{x\to 3} x + 2 = 5.$$

Computations like the one in Example 1.5 frequently occur when using the definition of a derivative.

Note that in Example 1.5 (and the examples below), one can make a guess of the limit by simply substituting numbers close to 3 for x into a calculator and see what you get. While this method can sometimes lead to the right answer, there are dangers with it! If we do this in the case of Example 1.5 we get the values in Table 1. In this case, we approach the right answer.

x	$f(x)$
3.1	5.1
3.01	5.01
3.001	5.001
3.0001	5.0001
3.00001	5.00001

Table 1: Here is the table of values we get when substituting values close to 3 in $f(x) = \frac{x^2-x-6}{x-3}$. As you can see, we get closer and closer to 5.

Example 1.6. Find $\lim\limits_{x\to 0} \dfrac{\sqrt{t^2+9}-3}{t^2}$.

Again, we want to apply the quotient limit rule, but the limit of the denominator is 0, so we should start with some algebraic manipulation. We attempt to solve the problem by rationalizing the numerator:

$$\begin{aligned}
\lim_{t\to 0} \frac{\sqrt{t^2+9}-3}{t^2} &= \lim_{t\to 0} \frac{\sqrt{t^2+9}-3}{t^2} \cdot \frac{\sqrt{t^2+9}+3}{\sqrt{t^2+9}+3} \\
&= \lim_{t\to 0} \frac{t^2+9-9}{t^2(\sqrt{t^2+9}+3)} \\
&= \lim_{t\to 0} \frac{1}{\sqrt{t^2+9}+3} \\
&= \frac{1}{\lim_{t\to 0} \sqrt{t^2+9}+3} \qquad \text{(using L3)} \\
&= \frac{1}{\sqrt{\lim_{t\to 0} t^2+9}+3} \qquad \text{(using L4)} \\
&= \frac{1}{6}.
\end{aligned}$$

Example 1.7. Sometimes it is possible to compute the limit of $f(x)$ as x approaches a from the left and as x approaches a from the right separately. We remarked earlier, that if those two values are not equal, then the limit $\lim\limits_{x\to a} f(x)$ does not exist. However if the limits from the left and from the right exist separately and are equal, then the limit $\lim\limits_{x\to a} f(x)$ also exists.

We now use this observation to compute $\lim\limits_{x\to a} f(x)$ where $f(x) = |x|$. Notice that $|x|$ is equal to x when $x \geq 0$, and is equal to $-x$ when $x < 0$. Thus, except for $a = 0$, the limit as $x \to a$ of $f(x)$ agrees with functions we have dealt with before. At $a = 0$, the right limit of $|x|$ as $x \to 0$ is equal to the right limit of x as $x \to 0$

and is therefore 0. Similarly, the left limit of $|x|$ as $x \to 0$ is equal to the left limit of $-x$ as $x \to 0$ and is therefore 0. Hence the left and right limits of $|x|$ as $x \to 0$ exist separately and are both 0, so $\lim_{x\to 0}|x| = 0$. Thus, we have

$$\begin{aligned}\lim_{x\to a}|x| &= \begin{cases} -a & \text{if } a < 0 \\ 0 & \text{if } a = 0 \\ a & \text{if } a > 0 \end{cases} \\ &= |a|.\end{aligned}$$

Limits as $x \to \infty$

We can also talk about the limit of a function as x approaches infinity. We write $\lim_{x\to\infty} f(x) = l$ if the value of $f(x)$ can be made arbitrarily close to l by taking x sufficiently large. For example, $\lim_{x\to\infty} 1/x = 0$ because $1/x$ gets closer and closer to 0 the larger x gets.

The four rules L1 - L4 all hold if we replace $x \to a$ with $x \to \infty$. Thus, using L2 repeatedly with $\lim_{x\to\infty} 1/x = 0$ we see that $1/x^n \to 0$ as $x \to \infty$ for any positive integer n.

Example 1.8. Let $f(x) = \frac{2x^2-1}{x^2+1}$. Again, in its current form, we can't use L3 since both numerator and denominator both go to ∞ as $x \to \infty$. However, applying some algebra, we get

$$\begin{aligned}\lim_{x\to\infty} \frac{2x^2-1}{x^2+1} &= \lim_{x\to\infty} \frac{\frac{2x^2-1}{x^2}}{\frac{x^2+1}{x^2}} \\ &= \lim_{x\to\infty} \frac{2-\frac{1}{x^2}}{1+\frac{1}{x^2}} \\ &= \frac{\lim_{x\to\infty} 2-\frac{1}{x^2}}{\lim_{x\to\infty} 1+\frac{1}{x^2}} \qquad \text{(using L3)} \\ &= \frac{2-0}{1+0} = 2.\end{aligned}$$

We finish this section with two more sophisticated rules for computing limits.

The squeeze rule

The squeeze rule, also known as the sandwich rule, is as follows. Let $f, g, h : U \to \mathbb{R}$ and let $(s,t) \subseteq U \cup \{a\}$. Suppose $f(x) \leq g(x) \leq h(x)$ for all $x \in (s,t) - \{a\}$, $\lim_{x\to a} f(x) = l$ and $\lim_{x\to a} h(x) = l$. Then $\lim_{x\to a} g(x) = l$.

Example 1.9. Let $g(x) = \sin x$. An easy geometric argument[2] shows that $-|x| \leq \sin x \leq |x|$ for all x. Since $\lim\limits_{x\to 0} -|x| = 0$ and $\lim\limits_{x\to 0}|x| = 0$, it follows from the squeeze rule that $\lim\limits_{x\to 0} \sin x = 0$.

From $\lim\limits_{x\to 0} \sin x = 0$ we can also deduce the limit $\lim\limits_{x\to 0} \cos x$ as follows. Since $\sin^2 x + \cos^2 x = 1$, we have $\cos x = \sqrt{1 - \sin^2 x}$ for $-\frac{\pi}{2} \leq x \leq \frac{\pi}{2}$. Thus

$$\lim_{x\to 0} \cos x = \sqrt{\lim_{x\to 0}(1 - \sin^2 x)} = \sqrt{1 - 0} = 1.$$

Example 1.10. Let $g(x) = \frac{\sin x}{x}$. Using a (less easy) geometric argument one can show that

$$\cos x < \frac{\sin x}{x} < 1$$

for $x \in (-\frac{\pi}{2}, \frac{\pi}{2}) - \{0\}$. Since $\lim\limits_{x\to 0} \cos x = 1$ and $\lim\limits_{x\to 0} 1 = 1$, the squeeze rule implies that

$$\lim_{x\to 0} \frac{\sin x}{x} = 1.$$

L'Hôpital's rule

If $\lim\limits_{x\to a} f(x) = \lim\limits_{x\to a} g(x) = 0$ or $\pm\infty$ and $\lim\limits_{x\to a} \dfrac{f'(x)}{g'(x)}$ exists then

$$\lim_{x\to a} \frac{f(x)}{g(x)} = \lim_{x\to a} \frac{f'(x)}{g'(x)}.$$

Example 1.11. Using the squeeze rule we have shown that $\lim\limits_{x\to 0} \dfrac{\sin x}{x} = 1$. Assuming that we know the derivative of $\sin x$, we can also compute this limit using L'Hôpital's rule: since $\lim\limits_{x\to 0} \sin x = 0$ and $\lim\limits_{x\to 0} x = 0$, L'Hôpital's rule gives

$$\lim_{x\to 0} \frac{\sin x}{x} = \lim_{x\to 0} \frac{\cos x}{1} = 1.$$

Remark: All the conditions in the statement of L'Hôpital's rule are essential. To see this, take $f(x) = \cos x$ and $g(x) = x$. Then, $\lim\limits_{x\to 0} \cos x = 1$ and $\lim\limits_{x\to 0} \dfrac{\cos x}{x}$ does not exist. However, if we attempted to apply L'Hôpital's rule, we would get

$$\lim_{x\to 0} \frac{-\sin x}{1} = 0.$$

[2] If $0 \leq \alpha \leq \pi/2$, then α is the length of the arc along the unit circle from the point $(\cos\alpha, \sin\alpha)$ to the point $(1, 0)$ on the x-axis, and $\sin\alpha$ is the distance of the point $(\cos\alpha, \sin\alpha)$ from the x-axis. It is clear that the arc is always at least as long as the distance, i.e. $\sin\alpha \leq \alpha = |\alpha|$. Also for such α we have $\sin\alpha \geq 0 \geq -|\alpha|$, so we have shown that $-|\alpha| \leq \sin\alpha \leq |\alpha|$. A similar argument shows the inequalities for negative values of α.

Thus, we would not get the correct answer. Also, the condition that $\lim_{x\to a} \frac{f'(x)}{g'(x)}$ exists is crucial. To see this, if $f(x) = x + \sin x$ and $g(x) = x$ then we see that $\lim_{x\to\infty} \frac{f(x)}{g(x)} = 1$, since

$$\lim_{x\to\infty} \frac{x + \sin x}{x} = \lim_{x\to\infty} \frac{1 + \frac{\sin x}{x}}{1} = \frac{1+0}{1} = 1.$$

The limit $\lim_{x\to\infty} \frac{\sin x}{x} = 0$ can be shown using the squeeze rule (which functions would you use in the squeeze rule?). However, if you attempted to apply L'Hôpital's rule, you would get

$$\lim_{x\to\infty} \frac{1+\cos x}{1},$$

which does not exist, so you would get the wrong answer again.

Exercises 1.2.

1. Evaluate the following limits

   (a) $\lim_{x\to -2} x^2$

   (b) $\lim_{x\to 3} \frac{x^2-9}{x-3}$

   (c) $\lim_{x\to 0} \frac{\frac{5}{x}-1}{3+\frac{2}{x}}$

   (d) $\lim_{h\to 0} \frac{\sqrt{2-h}-\sqrt{2}}{h}$

   (e) $\lim_{h\to\infty} \frac{3h^2-4h+6}{5h^2+2}$

   (f) $\lim_{h\to 0} \frac{h^2+5h}{h}$

2. Use L'Hôpital's rule, or otherwise, to determine the following limits.

   (a) $\lim_{x\to 0} \frac{\sin 4x}{\tan 5x}$

   (b) $\lim_{x\to\frac{\pi}{2}} \frac{\pi-2x}{\cos x}$

   (c) $\lim_{x\to 2} \frac{x^3-3x-2}{x^2-3x+2}$

   (d) $\lim_{x\to 1} \frac{x-1}{\ln x}$

   (e) $\lim_{x\to 1} \frac{x-\exp(x-1)}{(x-1)^2}$

3. Use the squeeze rule to evaluate the limit $\lim_{x\to 0} x \sin \frac{1}{x}$.

4. (Exam question, June 2005)

   (a) Without using L'Hôpital's rule, evaluate $\lim_{x\to 3} \dfrac{x^2 - x - 6}{x^2 - 9}$.

   (b) Using L'Hôpital's rule, or otherwise, evaluate $\lim_{x\to 0} \dfrac{\sin 3x}{e^{2x} - 1}$.

1.3 Continuous functions

Let $U \subseteq \mathbb{R}$. A function $f : U \to \mathbb{R}$ is **continuous at** $a \in U$ if $\lim_{x\to a} f(x) = f(a)$. The function f is **continuous on** U (or simply continuous) if it is continuous at a for all $a \in U$. Informally, a function f is continuous at $a \in U$ if its graph doesn't jump at this point.

We discussed earlier that any polynomial or rational function $f(x)$ defined at a satisfies $\lim_{x\to a} f(x) = f(a)$. **Thus, all such functions are continuous on their domains.** Furthermore we have shown that the function $f(x) = |x|$ is continuous.

In the previous section we saw that

$$\lim_{x\to 0} \sin x = 0 \quad \text{and} \quad \lim_{x\to 0} \cos x = 1.$$

Since both of these agree with the value of the function at 0, it follows that the functions sin and cos are continuous at 0.

Using the formula $\sin(x + a) = \sin(x)\cos(a) + \sin(a)\cos(x)$, we see that

$$\begin{aligned} \lim_{x\to a} \sin x = \lim_{x\to 0} \sin(x + a) &= \lim_{x\to 0} \big(\sin(x)\cos(a) + \sin(a)\cos(x)\big) \\ &= \lim_{x\to 0} \sin(x)\cos(a) + \lim_{x\to 0} \sin(a)\cos(x) = \sin(a). \end{aligned}$$

Thus, $\lim_{x\to a} \sin x = \sin a$. Since the limit as x approaches a of $\sin x$ is $\sin a$, the sine function is continuous everywhere. The above argument can be repeated for $\cos x$ as well. **Thus, the sine and cosine functions are continuous everywhere**. On first sight the tangent function looks like it is not continuous because its graph jumps at $\pi/2, 3\pi/2, \ldots$. However these values are not in the domain of definition $U = \mathbb{R} - \{\frac{2n+1}{2}\pi : n \in \mathbb{Z}\}$ of tan, and in fact the tangent function is continuous on its domain.

In fact, most of the nice functions we know are continuous on their domains.

Exercises 1.3.

1. Let $f : \mathbb{R} \to \mathbb{R}$ with

$$f(x) = \begin{cases} 1 - x & \text{if } x < -1 \\ 3 & \text{if } x = -1 \\ 3 + x & \text{if } -1 < x \leq 2 \\ 3x - 1 & \text{if } x > 2 \end{cases}$$

   Find all $a \in \mathbb{R}$ such that f is continuous at a.

1.4 The formal definition of the derivative

Let $U \subseteq \mathbb{R}$ and let $f : U \to \mathbb{R}$. Then the **derivative** of f at $x \in U$ is

$$\lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

if this limit exists. Otherwise f does not have a derivative at x.

The idea for the limit comes from considering the gradient of the chord through the points $(x, f(x))$ and $(x+h, f(x+h))$. As $h \to 0$, the chord approaches the tangent at $(x, f(x))$ and so the gradient of the chord approaches the gradient of the tangent at this point. See Figure 3 for an illustration.

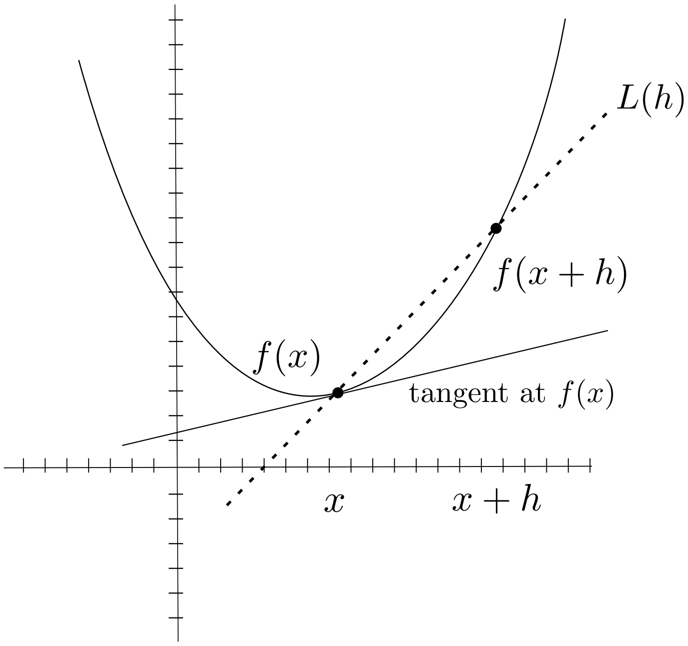


Figure 3: We illustrate why the definition of a derivative gives the gradient of the tangent line. Notice that the line $L(h)$ gets closer and closer to becoming the tangent line at $f(x)$ as h approaches 0.

Example 1.12. Let $f(x) = x^2$. Then

$$\begin{aligned}
f'(x) &= \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} \\
&= \lim_{h \to 0} \frac{(x+h)^2 - x^2}{h} \\
&= \lim_{h \to 0} \frac{(x^2 + 2hx + h^2) - x^2}{h} \\
&= \lim_{h \to 0} \frac{2hx + h^2}{h} \\
&= \lim_{h \to 0} 2x + h = 2x.
\end{aligned}$$

Thus $f'(x) = 2x$.

Example 1.13. We can *define* the number e as the unique number such that

$$\lim_{h \to 0} \frac{e^h - 1}{h} = 1.$$

(We would first have to show that $\lim\limits_{h \to 0} \dfrac{a^h - 1}{h}$ is always finite for any a, but we won't do that here.) Using this, we find the derivative of the exponential function using the definition of a derivative. We see that if $f(x) = e^x$, we have

$$\begin{aligned} f'(x) &= \lim_{h \to 0} \frac{e^{x+h} - e^x}{h} \\ &= \lim_{h \to 0} e^x \frac{e^h - 1}{h} \\ &= e^x \lim_{h \to 0} \frac{e^h - 1}{h} \\ &= e^x. \end{aligned}$$

Thus, $f'(x) = e^x$.

Example 1.14. Using the squeeze rule, we showed in Example 1.10 that

$$\lim_{x \to 0} \frac{\sin x}{x} = 1. \tag{2}$$

We now use this limit to show that the derivative of $\sin x$ is $\cos x$, using the definition of a derivative. We have

$$\begin{aligned} (\sin x)' &= \lim_{h \to 0} \frac{\sin(x+h) - \sin x}{h} \\ &= \lim_{h \to 0} \frac{\sin x \cos h + \sin h \cos x - \sin x}{h} \\ &= \sin x \lim_{h \to 0} \frac{\cos h - 1}{h} + \cos x \lim_{h \to 0} \frac{\sin h}{h}. \end{aligned} \tag{3}$$

The second term in (3) is $\cos(x) \cdot 1 = \cos(x)$ from our sine limit. What remains is to show that $\lim\limits_{h \to 0} \frac{\cos h - 1}{h} = 0$. However, we can do that using the sine limit:

$$\begin{aligned} \lim_{h \to 0} \frac{\cos h - 1}{h} &= \lim_{h \to 0} \frac{\cos h - 1}{h} \cdot \frac{\cos h + 1}{\cos h + 1} \\ &= \lim_{h \to 0} \frac{\cos^2 h - 1}{h(\cos h + 1)} \\ &= - \lim_{h \to 0} \frac{\sin^2 h}{h(\cos h + 1)} \end{aligned}$$

$$
\begin{aligned}
&= -\lim_{h\to 0} \frac{\sin h}{h} \cdot \frac{\sin h}{\cos h + 1} \\
&= -\lim_{h\to 0} \frac{\sin h}{h} \cdot \lim_{h\to 0} \frac{\sin h}{\cos h + 1} \qquad \text{(using L2)} \\
&= -1 \cdot \frac{\lim_{h\to 0} \sin h}{\lim_{h\to 0} \cos h + 1} \\
&= -1 \cdot \frac{0}{1+1} = 0
\end{aligned}
$$

(The previous limit can be shown using L'Hôpital's rule, but let's not assume we know the derivative of $\cos x$! You can find the derivative of $\cos x$ in the exercises below.) Thus, $(\sin x)' = \cos x$.

We end this section with an important theorem.

Theorem 1.15. *Suppose that f is a real-valued function. If f is differentiable at a point a, then f is continuous at a.*

The converse of Theorem 1.15 is false, and the classic example of a function continuous at a point a but not differentiable there is $|x|$. This function is continuous everywhere and differentiable everywhere but 0.

Exercises 1.4.

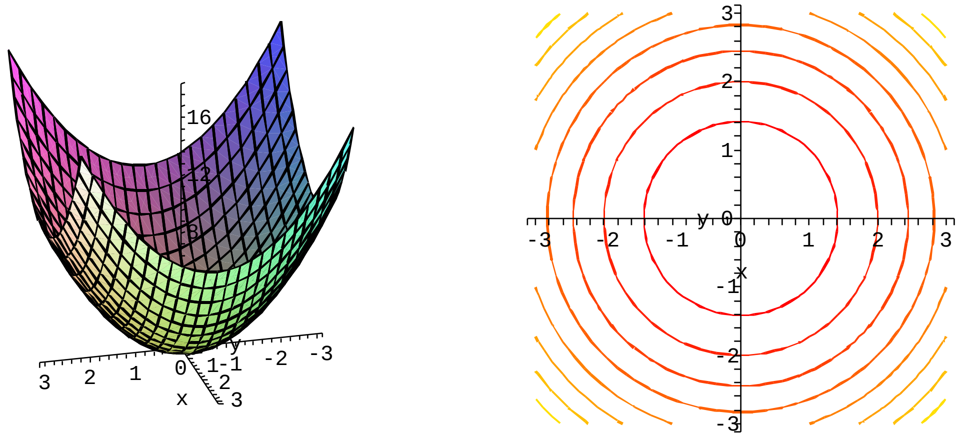
1. Use the formal definition to find the derivative of each of the following functions.
   (a) $f(x) = 3x$
   (b) $g(x) = x^3$
   (c) $h(x) = \frac{1}{x}$
   (d) $k(x) = x^n$ (Hint: use the binomial theorem)
   (e) $l(x) = \cos x$
2. Use the formal definition of the derivative and the properties of limits to prove the product rule for differentiation.

1.5 Functions of two variables

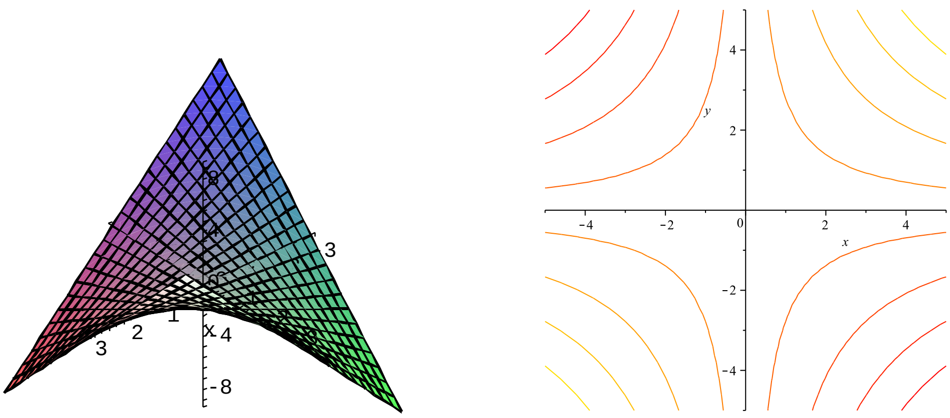
A (real) function of two variables is a function $f : U \to \mathbb{R}$ where $U \subseteq \mathbb{R}^2 = \{(x, y) : x, y \in \mathbb{R}\}$. Such a function determines for each $(x, y) \in U$ a unique real number denoted by $f(x, y)$. We call U the **domain of definition** of f and usually take it to be the largest set for which f is defined (more precisely, the largest set for which the rule or formula defining f makes sense).

We often represent $f(x, y)$ with z. In this case the function defines a subset of $\mathbb{R}^3$, namely $\{(x, y, z) : (x, y) \in U, z = f(x, y)\}$. Using this approach, we can think of a function of two variables as defining a surface in $\mathbb{R}^3$, and this is often a useful way of visualizing such a function. Alternatively, for $a \in \mathbb{R}$ the equation $f(x, y) = a$ defines a set of points in $\mathbb{R}^2$. Plotting the curves corresponding to these points for several values of a gives the contour-plot of the function.[3]

Example 1.16. The surface and contour plot, respectively, of $z = x^2 + y^2$.



Example 1.17. The surface and contour plot, respectively, of $z = xy$.



Exercises 1.5.

1. Determine the domain of definition for each of the following functions.

   (a) $f(x, y) = \dfrac{x}{x + y}$

   (b) $f(x, y) = \ln |xy|$

2. Sketch the contour plots for each of the following functions.

[3]The contour-plot is like a map of the surface representing the function. The curve $f(x, y) = a$ represents the points that are at distance a above the plane $z = 0$.

(a) $z = 2x + y$

(b) $z = 4x^2 + y^2$

(c) $z = \frac{y}{x}$

1.6 Partial derivatives

Let $U \subseteq \mathbb{R}^2$ and let $f : U \to \mathbb{R}$. The **partial derivative** of f with respect to x, denoted by $\frac{\partial f}{\partial x}$, is defined as

$$\frac{\partial f}{\partial x} = \lim_{h \to 0} \frac{f(x+h, y) - f(x, y)}{h}.$$

Similarly, the partial derivative of f with respect to y, denoted by $\frac{\partial f}{\partial y}$, is defined as

$$\frac{\partial f}{\partial y} = \lim_{h \to 0} \frac{f(x, y+h) - f(x, y)}{h}.$$

We also use the notation f_x and f_y for the partial derivative of f with respect to x and y, respectively.

If, in either case, the required limit does not exist then the corresponding partial derivative does not exist.

Note: To find $\frac{\partial f}{\partial x}$ we differentiate f with respect to x, and treat y as a constant. Similarly, to find $\frac{\partial f}{\partial y}$ we differentiate f with respect to y, and treat x as a constant.

Example 1.18. Let $f(x, y) = \frac{1}{3}x^3 - 3x^2y + y^2$. Then $\frac{\partial f}{\partial x} = x^2 - 6xy$ and $\frac{\partial f}{\partial y} = -3x^2 + 2y$.

Example 1.19. Let $z = (x^2 + y^2)\exp(-x^2 - 4y^2)$. Then[4]

$$\begin{aligned} \frac{\partial z}{\partial x} &= 2x\exp(-x^2 - 4y^2) + (x^2 + y^2)\exp(-x^2 - 4y^2)(-2x) \\ &= 2x\exp(-x^2 - 4y^2)(1 - x^2 - y^2), \end{aligned}$$

$$\begin{aligned} \frac{\partial z}{\partial y} &= 2y\exp(-x^2 - 4y^2) + (x^2 + y^2)\exp(-x^2 - 4y^2)(-8y) \\ &= 2y\exp(-x^2 - 4y^2)(1 - 4x^2 - 4y^2). \end{aligned}$$

Example 1.20. Let $f(x, y) = (x^2 + 2y)\ln(x^2 + 2y)$. Then

$$\begin{aligned} f_x(x, y) &= 2x\ln(x^2 + 2y) + (x^2 + 2y)\left(\frac{1}{x^2 + 2y}\right)2x \\ &= 2x(1 + \ln(x^2 + 2y)), \end{aligned}$$

[4] Recall that $\exp(X)$ is just an alternative way of writing e^X.

$$\begin{aligned} f_y(x,y) &= 2\ln(x^2+2y) + (x^2+2y)\left(\frac{1}{x^2+2y}\right)2 \\ &= 2(1+\ln(x^2+2y)). \end{aligned}$$

Example 1.21. Let $f(x,y) = \begin{cases} 1 & \text{if } x>0 \text{ and } y>0, \\ 0 & \text{otherwise.} \end{cases}$ Then, at the point $(0,0)$,

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

Similarly, at the point $(0,0)$, $\frac{\partial f}{\partial y} = 0$. Thus both $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ exist at $(0,0)$, even though f is not continuous at this point. Note that this contrasts sharply with functions of one variable, for if $U \subseteq \mathbb{R}$ and $f : U \to \mathbb{R}$ is differentiable at $a \in U$ then f is continuous at a.

Sometimes it is useful to consider the partial derivatives $f_x(x,y)$ and $f_y(x,y)$ together. This can be achieved by working with the **derivative** of f at (x,y) which is defined to be the row vector $(f_x(x,y), f_y(x,y))$. We denote the derivative of f at (x,y) by $Df(x,y)$ or $f'(x,y)$. For example for the function f in Example 1.20 we have

$$f'(x,y) = \big(2x(1+\ln(x^2+2y)), 2(1+\ln(x^2+2y))\big)$$

Exercises 1.6.

1. Let $f(x,y) = x^3y - 2x^2y^2 + 3xy - y^3$. Find f_x and f_y.
2. Let $z = e^{(x+y^2)}\sin(xy)$. Find $\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial y}$.

1.7 Products of vectors

In this section we first recall the definition of the dot product of two vectors. We then study the cross product which a useful tool for 3-dimensional geometry.

The dot product

Let $\mathbf{u} = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}$ and $\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$ be two vectors in $\mathbb{R}^n$. Then their dot product (also called scalar product) is defined by

$$\mathbf{u}.\mathbf{v} = u_1v_1 + u_2v_2 + \cdots + u_nv_n.$$

Thus the dot product of two vectors is a number and not a vector. The dot product is used to measure lengths and angles of vectors. The length of a vector $\mathbf{u}$, denoted by $|\mathbf{u}|$, is

$$\begin{aligned} |\mathbf{u}| &= \sqrt{u_1^2 + u_2^2 + \cdots + u_n^2} \\ &= \sqrt{\mathbf{u}.\mathbf{u}}. \end{aligned}$$

A vector is called a unit vector if it has length 1. The angle φ between two non-zero vectors $\mathbf{u}$ and $\mathbf{v}$ satisfies

$$\cos\varphi = \frac{\mathbf{u}.\mathbf{v}}{|\mathbf{u}||\mathbf{v}|}.$$

In particular, two non-zero vectors $\mathbf{u}$ and $\mathbf{v}$ are perpendicular (i.e. the angle between these vectors is $\frac{\pi}{2}$) if and only if $\mathbf{u}.\mathbf{v} = 0$.

The cross product

The cross product is only defined for vectors in 3-dimensional space. If $\mathbf{u} = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}$ and $\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}$ are two vectors in $\mathbb{R}^3$, then the **cross product**[5] $\mathbf{u} \times \mathbf{v}$ is defined by

$$\mathbf{u} \times \mathbf{v} = \begin{pmatrix} u_2v_3 - u_3v_2 \\ u_3v_1 - u_1v_3 \\ u_1v_2 - u_2v_1 \end{pmatrix}.$$

So the cross product of two 3-dimensional vectors is again a 3-dimensional vector. The following theorem shows that the cross product can be used to construct perpendiculars in 3-dimensional space.

Theorem 1.22. *Assume that neither* $\mathbf{u}$ *nor* $\mathbf{v}$ *is the zero vector and that* $\mathbf{u}$ *and* $\mathbf{v}$ *are not parallel. Then the vector* $\mathbf{u} \times \mathbf{v}$ *is not zero and is perpendicular to* $\mathbf{u}$ *and to* $\mathbf{v}$. *Furthermore the three vectors* $\mathbf{u}$, $\mathbf{v}$ *and* $\mathbf{u} \times \mathbf{v}$ *form a right-handed triad.*

The last statement of the theorem means that one can use the right hand to point into the directions of the three vectors: thumb pointing in the direction of $\mathbf{u}$, index finger pointing in the direction of $\mathbf{v}$ and middle finger pointing in the direction of $\mathbf{u} \times \mathbf{v}$.

Proof of Theorem 1.22. We will only show the statement that $\mathbf{u} \times \mathbf{v}$ is perpendicular

[5] The cross product is sometimes also called vector product because its result is again a vector and not a number.

to $\mathbf{u}$ and to $\mathbf{v}$. We have

$$\begin{aligned}(\mathbf{u}\times\mathbf{v}).\mathbf{u} &= (u_2v_3 - u_3v_2)u_1 + (u_3v_1 - u_1v_3)u_2 + (u_1v_2 - u_2v_1)u_3\\ &= u_1u_2v_3 - u_1u_3v_2 + u_2u_3v_1 - u_1u_2v_3 + u_1u_3v_2 - u_2u_3v_1\\ &= 0\end{aligned}$$

which shows that $\mathbf{u}\times\mathbf{v}$ is perpendicular to $\mathbf{u}$. A similar computation shows that $\mathbf{u}\times\mathbf{v}$ is perpendicular to $\mathbf{v}$. □

Example 1.23. If $\mathbf{u} = \begin{pmatrix}1\\0\\-3\end{pmatrix}$ and $\mathbf{v} = \begin{pmatrix}-1\\3\\2\end{pmatrix}$ then

$$\mathbf{u}\times\mathbf{v} = \begin{pmatrix}0\cdot 2 - (-3)\cdot 3\\ (-3)\cdot(-1) - 1\cdot 2\\ 1\cdot 3 - 0\cdot(-1)\end{pmatrix} = \begin{pmatrix}9\\1\\3\end{pmatrix}.$$

Exercises 1.7.

1. Let $\mathbf{u} = \begin{pmatrix}1\\2\\3\end{pmatrix}$, $\mathbf{v} = \begin{pmatrix}2\\2\\3\end{pmatrix}$ and $\mathbf{w} = \begin{pmatrix}-1\\0\\0\end{pmatrix}$. Compute the cross products $\mathbf{u}\times\mathbf{v}$, $(\mathbf{u}\times\mathbf{v})\times\mathbf{w}$, $\mathbf{v}\times\mathbf{w}$ and $\mathbf{u}\times(\mathbf{v}\times\mathbf{w})$. Is the cross product associative?

2. Show that $\mathbf{u}\times\mathbf{v} = -(\mathbf{v}\times\mathbf{u})$ for any vectors $\mathbf{u}$ and $\mathbf{v}$. This property of the cross product is called anticommutativity.

1.8 Tangent planes

Before discussing tangent planes for a function of two variables we recall how to find tangent lines for a function of one variable.

Tangent lines

If $U \subseteq \mathbb{R}$ and $f : U \to \mathbb{R}$ then for $a \in U$ the tangent (line) at the point $P = (a, f(a))$ is the straight line passing through P with the same gradient as f at P. The equation of a straight line has the form $y = mx + c$ for some constants m and c, where m is the gradient of the line. Now the gradient of f at P is $f'(a)$, and so for the tangent at P, $m = f'(a)$. Also, the tangent at P passes through the point $(a, f(a))$, and so putting $x = a$ and $y = f(a)$ into the equation of the tangent gives

$$f(a) = f'(a)a + c.$$

Hence $c = f(a) - f'(a)a$, and so the equation of the tangent at P is

$$\begin{aligned} y &= f'(a)x + f(a) - f'(a)a\\ &= f'(a)(x-a) + f(a).\end{aligned}$$

Thus the tangent at P is

$$f'(a)(x-a) + f(a) - y = 0,$$

which can be rewritten in vector form as

$$\begin{pmatrix} x-a \\ y-f(a) \end{pmatrix} \cdot \begin{pmatrix} f'(a) \\ -1 \end{pmatrix} = 0.$$

Here we have assumed that $f'(a)$ exists, and that the graph of f is not vertical at $P = (a, f(a))$; if the graph of f was vertical at P then the tangent at P is $x = a$.

We want to give a geometric interpretation of the vector $\begin{pmatrix} f'(a) \\ -1 \end{pmatrix}$ which appears in the equation of the tangent line. Since $f'(a)$ is the gradient of the tangent line at $P = (a, f(a))$, the vector $\begin{pmatrix} 1 \\ f'(a) \end{pmatrix}$ is parallel to the tangent line (i.e. it points in the same direction as the line). Since $\begin{pmatrix} 1 \\ f'(a) \end{pmatrix} \cdot \begin{pmatrix} f'(a) \\ -1 \end{pmatrix} = 0$, we see that the vectors $\begin{pmatrix} 1 \\ f'(a) \end{pmatrix}$ and $\begin{pmatrix} f'(a) \\ -1 \end{pmatrix}$ are perpendicular. Thus the vector $\begin{pmatrix} f'(a) \\ -1 \end{pmatrix}$ is **normal** to the tangent line at P, and so is normal to the curve $y = f(x)$ at the point $(a, f(a))$.

Tangent planes

We now want to generalize the idea of a tangent at a point to functions of two variables. Let $f : U \to \mathbb{R}$ for some $U \subseteq \mathbb{R}^2$ and let $(a, b) \in U$. We look for a plane through the point $P = (a, b, f(a, b))$, whose gradient in the x and y directions agrees with $f_x(a, b)$ and $f_y(a, b)$, respectively.

Now, the equation of any non-vertical plane in $\mathbb{R}^3$ has the form

$$z = m_1 x + m_2 y + c,$$

for some $m_1, m_2, c \in \mathbb{R}$. For this plane,

$$\frac{\partial z}{\partial x} = m_1 \text{ and } \frac{\partial z}{\partial y} = m_2.$$

Since these two gradients of the tangent plane agree with the partial derivatives of f at P, we require that

$$\begin{aligned} m_1 &= f_x(a, b), \\ m_2 &= f_y(a, b). \end{aligned}$$

In order to determine c, we note that when $(x, y) = (a, b)$ then $z = f(a, b)$. Thus

$$f(a, b) = f_x(a, b)a + f_y(a, b)b + c,$$

which gives

$$c = f(a,b) - f_x(a,b)a - f_y(a,b)b.$$

From this we get that the tangent plane of f at (a,b) is

$$\begin{aligned} z &= f_x(a,b)x + f_y(a,b)y + f(a,b) - f_x(a,b)a - f_y(a,b)b \\ &= f_x(a,b)(x-a) + f_y(a,b)(y-b) + f(a,b). \end{aligned}$$

This can be rewritten as

$$f_x(a,b)(x-a) + f_y(a,b)(y-b) + f(a,b) - z = 0,$$

or in vector form as

$$\begin{pmatrix} x-a \\ y-b \\ z-f(a,b) \end{pmatrix} \cdot \begin{pmatrix} f_x(a,b) \\ f_y(a,b) \\ -1 \end{pmatrix} = 0.$$

Again we want to give a geometric interpretation of the vector $\begin{pmatrix} f_x(a,b) \\ f_y(a,b) \\ -1 \end{pmatrix}$. Since $f_x(a,b)$ is the gradient of the tangent plane at $(a,b,f(a,b))$ in the x-direction, the vector $\begin{pmatrix} 1 \\ 0 \\ f_x(a,b) \end{pmatrix}$ is parallel to the tangent plane. Similarly, the vector $\begin{pmatrix} 0 \\ 1 \\ f_y(a,b) \end{pmatrix}$ is parallel to the tangent plane. The cross product of these two vectors is $\begin{pmatrix} 0 \\ 1 \\ f_y(a,b) \end{pmatrix} \times \begin{pmatrix} 1 \\ 0 \\ f_x(a,b) \end{pmatrix} = \begin{pmatrix} f_x(a,b) \\ f_y(a,b) \\ -1 \end{pmatrix}$, and from the properties of the cross product it follows that the vector $\begin{pmatrix} f_x(a,b) \\ f_y(a,b) \\ -1 \end{pmatrix}$ is perpendicular to $\begin{pmatrix} 0 \\ 1 \\ f_y(a,b) \end{pmatrix}$ and $\begin{pmatrix} 1 \\ 0 \\ f_x(a,b) \end{pmatrix}$. Hence $\begin{pmatrix} f_x(a,b) \\ f_y(a,b) \\ -1 \end{pmatrix}$ is normal to the tangent plane, and so is normal to the surface of $z = f(x,y)$ at the point $(a,b,f(a,b))$.

Comparing the expression for the tangent plane at a point with the tangent line at a point, we observe that $\begin{pmatrix} f_x(a,b) \\ f_y(a,b) \end{pmatrix}$ in the tangent plane is playing the role of $f'(a)$ in the tangent line.

Example 1.24. Let $f(x,y) = x^2 - 3xy + 2y^2$. Then $f_x = 2x - 3y$ and $f_y = -3x + 4y$. To find the tangent plane at the point $(a,b) = (1,2)$ we calculate $f_x(1,2) = -4$, $f_y(1,2) = 5$ and $f(1,2) = 3$. Then the tangent plane at $(1,2)$ is

$$\begin{pmatrix} x-1 \\ y-2 \\ z-3 \end{pmatrix} \cdot \begin{pmatrix} -4 \\ 5 \\ -1 \end{pmatrix} = 0.$$

Thus the tangent plane at $(1, 2)$ is

$$-4x + 4 + 5y - 10 - z + 3 = 0.$$

That is

$$4x - 5y + z = -3.$$

Note that the vector $\begin{pmatrix} -4 \\ 5 \\ -1 \end{pmatrix}$ is normal to the surface of f at $(1, 2, 3)$.

Exercises 1.8.

1. Find the equation of the tangent to $y = x^2 - 4x + 4$ when $x = 3$.
2. Let $f(x, y) = -\frac{1}{4}(3xy^2 - 5x^3y + 2x^4)$. Find the equation of the tangent plane to f at the point $(2, 4)$.
3. (Exam question, June 2005)
   Let $f : \mathbb{R}^2 \to \mathbb{R}$, with $f(x, y) = \dfrac{xy}{1 + 2x + y^2}$. Find the equation of the tangent plane to f at the point $(-2, 1)$. Use your answer to estimate $f(-1.9, 1.1)$.

1.9 The gradient

Let $f : U \to \mathbb{R}$ for some $U \subseteq \mathbb{R}^2$ and let $(a, b) \in U$. Consider the tangent plane of f at (a, b):

$$f(a, b) - z + f_x(a, b)(x - a) + f_y(a, b)(y - b) = 0.$$

When $z = f(a, b)$ we get

$$f_x(a, b)(x - a) + f_y(a, b)(y - b) = 0.$$

That is,

$$\begin{pmatrix} x - a \\ y - b \end{pmatrix} \cdot \begin{pmatrix} f_x(a, b) \\ f_y(a, b) \end{pmatrix} = 0.$$

This is the equation of a straight line in the plane. In fact it is the tangent to the contour $f(x, y) = f(a, b)$ at the point (a, b). The vector $\begin{pmatrix} f_x(a, b) \\ f_y(a, b) \end{pmatrix}$ is normal to this tangent, and so is normal to the contour $f(x, y) = f(a, b)$.

The vector $\begin{pmatrix} f_x \\ f_y \end{pmatrix}$ is called the **gradient** of f, and is denoted by ∇f. (Note that ∇f is just the transpose of the derivative of f.)

Example 1.25. For $f(x, y) = x^2 - 3xy + 2y^2$, $\nabla f = \begin{pmatrix} 2x - 3y \\ -3x + 4y \end{pmatrix}$.

Exercises 1.9.

1. Let $f(x, y) = -\frac{1}{4}(3xy^2 - 5x^3y + 2x^4)$. Find $\nabla f(x, y)$.

1.10 Directional derivatives

Let $f : U \to \mathbb{R}$ be a function of two variables x and y and consider a point $(a, b) \in U$. Then $f_x(a, b)$ gives the gradient of f at (a, b) in the x direction; that is, in the direction of the vector $\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$. Similarly, $f_y(a, b)$ gives the gradient of f at (a, b) in the y direction; that is, in the direction of the vector $\mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. We now determine how to find the gradient of f at (a, b) in the direction of an arbitrary vector $\mathbf{u} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$.

We suppose that $\mathbf{u}$ is a **unit** vector. That is, $u_1^2 + u_2^2 = 1$. We denote by $f_{\mathbf{u}}(a, b)$ the gradient of f at (a, b) in the direction of $\mathbf{u}$. To be consistent with the idea that $f_{\mathbf{e}_1}(a, b) = f_x(a, b)$ and $f_{\mathbf{e}_2}(a, b) = f_y(a, b)$, we expect $\begin{pmatrix} u_1 \\ u_2 \\ f_{\mathbf{u}}(a, b) \end{pmatrix}$ to be parallel to the tangent plane of f at the point (a, b). Thus we get

$$\begin{pmatrix} u_1 \\ u_2 \\ f_{\mathbf{u}}(a, b) \end{pmatrix} . \begin{pmatrix} f_x(a, b) \\ f_y(a, b) \\ -1 \end{pmatrix} = 0.$$

This gives

$$\begin{aligned} f_{\mathbf{u}}(a, b) &= u_1 f_x(a, b) + u_2 f_y(a, b) \\ &= \begin{pmatrix} f_x(a, b) \\ f_y(a, b) \end{pmatrix} . \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \\ &= \nabla f(a, b).\mathbf{u}. \end{aligned}$$

The number $f_{\mathbf{u}}(a, b)$ is called the **directional derivative** of f at (a, b) in the direction of the (unit) vector $\mathbf{u}$.

Example 1.26. Let $f(x, y) = x^2 - 3xy + 2y^2$ and $\mathbf{u} = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$. To find the directional derivative of f at $(1, 2)$ in the direction of $\mathbf{u}$, we first need to replace $\mathbf{u}$ with a unit vector. Since the length of $\mathbf{u}$ is $\sqrt{3^2 + 1^2} = \sqrt{10}$ we replace $\mathbf{u}$ with $\hat{\mathbf{u}} = \frac{1}{\sqrt{10}} \begin{pmatrix} 3 \\ 1 \end{pmatrix}$. Now

$$\begin{aligned} f_{\mathbf{u}}(1, 2) &= \nabla f(1, 2).\hat{\mathbf{u}} \\ &= \frac{1}{\sqrt{10}} \begin{pmatrix} f_x(1, 2) \\ f_y(1, 2) \end{pmatrix} . \begin{pmatrix} 3 \\ 1 \end{pmatrix} \\ &= \frac{1}{\sqrt{10}} \begin{pmatrix} -4 \\ 5 \end{pmatrix} . \begin{pmatrix} 3 \\ 1 \end{pmatrix} \\ &= \frac{1}{\sqrt{10}}(-12 + 5) = -\frac{7}{\sqrt{10}} \approx -2.2136. \end{aligned}$$

Exercises 1.10.

1. Let $f(x,y) = -\frac{1}{4}(3xy^2 - 5x^3y + 2x^4)$. Find the directional derivative $f_{\mathbf{u}}(2,4)$, where $\mathbf{u} = \begin{pmatrix} \frac{8}{17} \\ \frac{15}{17} \end{pmatrix}$.
2. Let $f(x,y) = 3x + xy^3 - 5x^2y$. Find the directional derivative $f_{\mathbf{u}}(1,2)$, where $\mathbf{u} = \begin{pmatrix} -3 \\ 4 \end{pmatrix}$.
3. (Exam question, June 2005)
   Let $f : \mathbb{R}^2 \to \mathbb{R}$, with $f(x,y) = y^2 - 5xy + 2x^2y - x^3$.
   (a) Find $\nabla f(x,y)$.
   (b) Find $f_u(3,4)$, when $u = \begin{pmatrix} -\frac{12}{13} \\ \frac{5}{13} \end{pmatrix}$.

1.11 Functions of n variables

In this section we sketch how the ideas developed for functions of two variables generalize to functions of n variables.

Let $U \subseteq \mathbb{R}^n = \{(x_1, x_2, \ldots, x_n) : x_1, x_2, \ldots, x_n \in \mathbb{R}\}$. A (real) function $f : U \to \mathbb{R}$ of n variables determines for each $(x_1, x_2, \ldots, x_n) \in U$ a unique real number denoted by $f(x_1, x_2, \ldots, x_n)$.

For each variable x_i $(i = 1, 2, \ldots, n)$ we define the partial derivative of f with respect to x_i by

$$\frac{\partial f}{\partial x_i} = \lim_{h \to 0} \frac{f(x_1, x_2, \ldots, x_i + h, \ldots, x_n) - f(x_1, x_2, \ldots, x_i, \ldots, x_n)}{h},$$

where the limit exists; otherwise the partial derivative does not exist. We also denote $\frac{\partial f}{\partial x_i}$ by f_{x_i}.

When dealing with functions of n variables we often abbreviate $(x_1, x_2, \ldots, x_i, \ldots, x_n)$ to $\mathbf{x}$; in particular we write $f(\mathbf{x})$ instead of $f(x_1, x_2, \ldots, x_i, \ldots, x_n)$.

We can generalize the ideas developed for functions of two variables to functions of n variables. So, if $U \subseteq \mathbb{R}^n$ and $f : U \to \mathbb{R}$, then the **tangent hyperplane** at $\mathbf{a} = (a_1, a_2, \ldots, a_n) \in U$ is

$$y - f(\mathbf{a}) = f_{x_1}(\mathbf{a})(x_1 - a_1) + f_{x_2}(\mathbf{a})(x_2 - a_2) + \cdots + f_{x_n}(\mathbf{a})(x_n - a_n),$$

which can be rewritten as

$$\begin{pmatrix} x_1 - a_1 \\ x_2 - a_2 \\ \vdots \\ x_n - a_n \\ y - f(\mathbf{a}) \end{pmatrix} \cdot \begin{pmatrix} f_{x_1}(\mathbf{a}) \\ f_{x_2}(\mathbf{a}) \\ \vdots \\ f_{x_n}(\mathbf{a}) \\ -1 \end{pmatrix} = 0.$$

The vector $\begin{pmatrix} f_{x_1}(\mathbf{a}) \\ f_{x_2}(\mathbf{a}) \\ \vdots \\ f_{x_n}(\mathbf{a}) \\ -1 \end{pmatrix}$ is **normal** to the **hypersurface** of $y = f(\mathbf{x})$ at the point $(a_1, a_2, \ldots, a_n, f(\mathbf{a}))$.

The **gradient** $\nabla f(\mathbf{x})$ is given by

$$\nabla f(\mathbf{x}) = \begin{pmatrix} f_{x_1}(\mathbf{x}) \\ f_{x_2}(\mathbf{x}) \\ \vdots \\ f_{x_n}(\mathbf{x}) \end{pmatrix},$$

and the **derivative** of $f(\mathbf{x})$ is just the transpose of $\nabla f(\mathbf{x})$.

For a unit vector $\mathbf{u} \in \mathbb{R}^n$, the **directional derivative** $f_{\mathbf{u}}(\mathbf{x})$ is given by

$$f_{\mathbf{u}}(\mathbf{x}) = \nabla f(\mathbf{x}).\mathbf{u}.$$

Example 1.27. Let $f(x, y, z) = 3xy^2z^2 - 4x^3z$. Then $\frac{\partial f}{\partial x} = 3y^2z^2 - 12x^2z$, $\frac{\partial f}{\partial y} = 6xyz^2$ and $\frac{\partial f}{\partial z} = 6xy^2z - 4x^3$. At the point $(\frac{1}{2}, 1, -1)$, $f = 2$, $\frac{\partial f}{\partial x} = 6$, $\frac{\partial f}{\partial y} = 3$ and $\frac{\partial f}{\partial z} = -\frac{7}{2}$. Thus the tangent hyperplane of f at the point $(\frac{1}{2}, 1, -1)$ is

$$\begin{pmatrix} x - \frac{1}{2} \\ y - 1 \\ z + 1 \\ w - f(\frac{1}{2}, 1, -1) \end{pmatrix} . \begin{pmatrix} f_x(\frac{1}{2}, 1, -1) \\ f_y(\frac{1}{2}, 1, -1) \\ f_z(\frac{1}{2}, 1, -1) \\ -1 \end{pmatrix} = 0.$$

That is,

$$\begin{pmatrix} x - \frac{1}{2} \\ y - 1 \\ z + 1 \\ w - 2 \end{pmatrix} . \begin{pmatrix} 6 \\ 3 \\ -\frac{7}{2} \\ -1 \end{pmatrix} = 0,$$

which gives

$$6x - 3 + 3y - 3 - \frac{7}{2}z - \frac{7}{2} - w + 2 = 0.$$

Simplifying the last equation, the tangent hyperplane of f at the point $(\frac{1}{2}, 1, -1)$ is

$$12x + 6y - 7z - 2w = 15.$$

Exercises 1.11.

1. Consider the function

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

Compute the gradient ∇f and the directional derivative $f_{\mathbf{u}}(2\pi, \pi, 0)$ where $\mathbf{u} = \frac{1}{3}\begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}$.