Chapter 3

Limits of Functions

3.1 Functions on the Real Line

Recall that if A and B are two sets, their *Cartesian product* $A \times B$ is the set of all ordered pairs:

$$A \times B = \{(a, b); a \in A, b \in B\},$$

where $(a, b) = \{\{a\}, \{a, b\}\}$.

A *relation* is a subset R of $A \times B$. We write aRb if $(a, b) \in R$. The *domain* of the relation is the set $D_R = \{a \in A; aRb \text{ for some } b \in B\}$. We say that the relation R is *functional* if for all $a \in D_R$, there is exactly one $b \in B$ so that aRb. In this case we write $b = f(a)$ instead of aRb, and we write D_f instead of D_R. We call $f : A \to B$ a *function* or a *mapping* (with domain D_f). Sometimes we use the more concise notation $f : D_f \to B$. The *range* or *image* of the mapping f is the set $R_f = \{b \in B; b = f(a) \text{ for some } a \in A\}$.

If f and g are two functions from A to B with domains D_f and D_g, respectively, we say that $f = g$ if

- $D_f = D_g$,
- $f(x) = g(x)$ for all $x \in D_f$.

In this course, we always take $A = B = \mathbb{R}$. You should think of the domain D_f as being all the points x in $\mathbb{R}$ for which $f(x) \in \mathbb{R}$. Any $x \in \mathbb{R}$ for which we might be tempted to write "$f(x) = 0/0$" or "$f(x) = \pm\infty$" is not in D_f.

Example 3.1. A *polynomial* is a function of the form

$$f(x) = a_0 + a_1 x + \cdots + a_{n-1} x^{n-1} + a_n x^n,$$

where the *coefficients* $a_0, a_1, \ldots, a_{n-1}, a_n \in \mathbb{R}$, and the *degree* $n \in \mathbb{N}$. It is clear that we always have $D_f = \mathbb{R}$.

Example 3.2. A *rational function* is a function of the form $f(x) = p(x)/q(x)$, where p and q are polynomials. Its domain is $D_f = \{x \in \mathbb{R}; q(x) \neq 0\}$, e.g. $f(x) = \dfrac{x^2+5}{(x+1)(x-3)}$ has domain $D_f = \mathbb{R} \setminus \{-1, 3\}$, where we recall that if X and Y are arbitrary sets, then

$$X \setminus Y = X \cap Y^c = \{x \in X; x \notin Y\}.$$

Other important functions are $f(x) = e^{kx}, f(x) = \sin(kx)$ and $f(x) = \cos(kx)$, where $k \in \mathbb{R}$. These all have domain $\mathbb{R}$. In analysis, they are best defined as convergent power series, and that topic will be picked up in Semester 2.

Example 3.3. The *sign function* is defined by

$$\operatorname{sign}(x) = \begin{cases} -1 & \text{if } x < 0 \\ 0 & \text{if } x = 0 \\ 1 & \text{if } x > 0 \end{cases}$$

In this case, $D_f = \mathbb{R}, R_f = \{-1, 0, 1\}$.

Example 3.4. The *indicator function* of the closed interval $[a, b]$ is the function

$$\mathbf{1}_{[a,b]}(x) = \begin{cases} 1 & \text{if } x \in [a, b] \\ 0 & \text{if } x \notin [a, b] \end{cases}$$

In this case, $D_f = \mathbb{R}, R_f = \{0, 1\}$. In fact, we can talk about indicator functions of open intervals, or even arbitrary sets in $\mathbb{R}$. The indicator function of $[0, \infty)$ is called the *Heaviside function* by engineers.

We can add functions, multiply them by scalars, multiply them together and divide them, but we need to be careful with domains. So if f and g are functions from $\mathbb{R}$ to $\mathbb{R}$, and $\alpha \in \mathbb{R}$, we have

$$\begin{gathered} (f+g)(x) = f(x) + g(x), \\ (\alpha f)(x) = \alpha f(x), \\ (fg)(x) = f(x)g(x), \\ \left(\frac{f}{g}\right)(x) = \frac{f(x)}{g(x)}, \end{gathered}$$

with $D_{f+g} = D_f \cap D_g, D_{\alpha f} = D_f, D_{fg} = D_f \cap D_g, D_{f/g} = (D_f \cap D_g) \setminus \{x \in D_g; g(x) = 0\}$.

Remark. Let $\mathcal{F}(\mathbb{R})$ denote the set of all functions from $\mathbb{R}$ to $\mathbb{R}$ having domain $\mathbb{R}$. Then the first two of the list tell us that $\mathcal{F}(\mathbb{R})$ is a vector space over $\mathbb{R}$, (but it turns out that it is not finite-dimensional). The first and third tell us that $\mathcal{F}(\mathbb{R})$ is a ring, and the first three that $\mathcal{F}(\mathbb{R})$ is an algebra. These structural properties are important in higher analysis.

The *composition* $f \circ g$ of two functions f and g is defined only when $R_g \subseteq D_f$, and we then have

$$(f \circ g)(x) = f(g(x)),$$

with $D_{f \circ g} = D_g$.

Note that we typically do not have $f \circ g = g \circ f$, even where the domains of these functions have points in common, e.g. $f(x) = \sqrt{x}$ and $g(x) = e^x$. $R_g = (0, \infty) \subset D_f = [0, \infty)$ and $(f \circ g)(x) = e^{x/2}$ with $D_{f \circ g} = \mathbb{R}$; but $R_f = [0, \infty) \subset D_g = \mathbb{R}$ and $(g \circ f)(x) = e^{\sqrt{x}}$ with $D_{g \circ f} = [0, \infty)$.

3.2 The Limit of a Function

We want to make sense of the notion of $\lim_{n \to \infty} f(x)$ for a function $f : \mathbb{R} \to \mathbb{R}$. We do this by using sequences.

Definition. We say that f has a *limit* l at a point $a \in \mathbb{R}$, if $\lim_{n \to \infty} f(x_n) = l$ for every sequence (x_n) which satisfies

1. For all $n \in \mathbb{N}, x_n \in D_f \setminus \{a\}$,
2. $\lim_{n \to \infty} x_n = a$.

In this case, we write

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

Notes.

1. We must have convergence for *every* sequence (x_n).[1]
2. We do not need to have $a \in D_f$. In fact, if $a \notin D_f$, then $D_f \setminus \{a\} = D_f$.

[1] This is a technical point. How do we know that $D_f \setminus \{a\}$ contains any sequences (x_n) that converge to a? In order to ensure this, we will always impose a *domain condition* on our function f at a, namely that there exists $\delta > 0$ so that $(a - \delta, a) \cup (a, a + \delta) \subseteq D_f$. If this condition is in place, then we can find many sequences in $D_f \setminus \{a\}$ that converge to a, e.g. any sequence of the form $(a \pm y_n \delta)$ where (y_n) is a null sequence with $0 < |y_n| < 1$ for all $n \in \mathbb{N}$. This domain condition will be satisfied for all of the examples of functions that we consider in the notes.

Example 3.5. $f(x) = \dfrac{x-5}{(x^2-25)(x-3)}$ has domain $D_f = \mathbb{R}\backslash\{-5, 3, 5\}$.

Observe that $f(x) = \dfrac{1}{(x+5)(x-3)}$ for all $x \in D_f$.

It is fairly easy to calculate limits at points in D_f. For example,

$\lim_{x\to1} f(x) = -\frac{1}{12}$. (Check this carefully, using the definition). The more interesting problem is to find out what happens at points that are not in D_f.

To investigate the point $x = 5$. Choose an arbitrary sequence (x_n) with $x_n \in D_f$ satisfying $\lim_{n\to\infty} x_n = 5$ (examples of such sequences are $x_n = 5 + 1/n, 5 + 7/n^2, 5 - 108/n^4$ etc.). Then $f(x_n) = \dfrac{1}{(x_n+5)(x_n-3)}$ and $\lim_{n\to\infty} f(x_n) = \dfrac{1}{10.2} = \dfrac{1}{20}$, by the algebra of limits. So we conclude that

$$\lim_{x\to5} f(x) = \frac{1}{20}.$$

Thus the limit exists at the point $x = 5$, even though it is not in the domain of f.

To investigate the point $x = -5$. In this case, numerical experiments may lead you to doubt that a limit exists. So consider the sequence (y_n), where $y_n = -5 + 1/n$, for each $n \in \mathbb{N}$. Then $\lim_{n\to\infty} y_n = -5$, and we find that

$$f(y_n) = \frac{1}{1/n(1/n-8)} = \frac{n}{1/n-8},$$

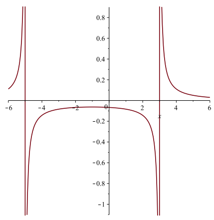
and so $\lim_{n\to\infty} f(y_n) = -\infty$. So we've found a sequence (y_n) such that (1) and (2) of the definition are satisfied, but $(f(y_n))$ diverges. Hence we conclude that f has no limit at $x = -5$. Its an *exercise* for you to figure out what happens at $x = 3$.

Note that, in this example, if you take the limit at a point $a \in D_f$, you always find that

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

e.g. if $a = 1, \lim_{x\to1} f(x) = -\frac{1}{12} = f(1)$. This tells us that the function f is *continuous* at these points. We will discuss the concept of continuity in more detail in Chapter 4. Note that the function f is not continuous at the points $x = -5, 3$ or 5; indeed $f(a)$ is not defined when a takes these values.

Figure 3.1: Graph of $f(x) = \dfrac{x-5}{(x^2-25)(x-3)}$



Theorem 3.2.1. *[Algebra of Limits] Suppose that $f, g : \mathbb{R} \to \mathbb{R}$, and $a \in \mathbb{R}$ is such that $\lim_{x\to a} f(x) = l$ and $\lim_{x\to a} g(x) = m$, then*

1. $\lim_{x\to a}(f+g)(x) = l + m$,
2. $\lim_{x\to a}(fg)(x) = lm$,
3. $\lim_{x\to a}(\alpha f)(x) = \alpha l$, *for all* $\alpha \in \mathbb{R}$,
4. $\lim_{x\to a}\left(\dfrac{f}{g}\right)(x) = \dfrac{l}{m}$, *if* $m \neq 0$.

Proof. These all follow from the algebra of limits for sequences (Theorem 2.2.1), e.g. for (1), if (x_n) is an arbitrary sequence in $(D_f \cap D_g) \setminus \{a\}$ that converges to a, then

$$\begin{aligned}\lim_{n\to\infty}(f+g)(x_n) &= \lim_{n\to\infty}(f(x_n) + g(x_n)) \\ &= \lim_{n\to\infty} f(x_n) + \lim_{n\to\infty} g(x_n) \\ &= l + m,\end{aligned}$$

as required. Note that Theorem 2.2.1 is used here to get to the second line of the display. □

From a geometric perspective, the idea of a limit is that as x gets closer and closer to a, so $f(x)$ should get closer and closer to $f(a)$. More insight to this is given by the following theorem, which establishes the important $(\epsilon - \delta)$ criterion for existence of limits.

Theorem 3.2.2. *Let f be a mapping from $\mathbb{R}$ to $\mathbb{R}$, then $\lim_{x\to a} f(x) = l$, if and only if the $(\epsilon - \delta)$ criterion holds, i.e. given any $\epsilon > 0$ there exists $\delta > 0$ so that whenever $x \in D_f$ with $0 < |x - a| < \delta$, then $|f(x) - l| < \epsilon$.*

Proof. First assume that the $(\epsilon - \delta)$ criterion holds and suppose $\epsilon > 0$. Now let (x_n) be an arbitrary sequence in $D_f \setminus \{a\}$, which converges to a. Then for any $\eta > 0$, there exists $N \in \mathbb{N}$, so that if $n > N$, we have $0 < |x_n - a| < \eta$. Now choose η to be the δ which comes from the criterion. Then for all $n > N$, we have $|f(x_n) - l| < \epsilon$, and so $\lim_{x\to a} f(x) = l$, as was required.

Now we must establish the converse, namely if f has a limit at a, then the $\epsilon - \delta$ criterion follows. In fact, we seek a proof by contrapositive[2], so we will assume that the $\epsilon - \delta$ criterion fails, and then show that f cannot have a limit at a. If the $\epsilon - \delta$ criterion fails, then there exists $\epsilon > 0$ such that for all $\delta > 0$, there exists $x \in D_f$ with $0 < |x - a| < \delta$, but $|f(x) - l| \geq \epsilon$.

Now for the given $\epsilon > 0$, choose successively $\delta = 1, 1/2, 1/3, \ldots$ and construct a sequence (x_n) as follows:

$$\begin{array}{c} x_1 \in D_f \text{ satisfies } 0 < |x_1 - a| < 1 \text{ and } |f(x_1) - l| \geq \epsilon. \\ x_2 \in D_f \text{ satisfies } 0 < |x_2 - a| < 1/2 \text{ and } |f(x_2) - l| \geq \epsilon. \\ \cdots\cdots\cdots\cdots \\ \cdots\cdots\cdots\cdots \\ \cdots\cdots\cdots\cdots \\ x_n \in D_f \text{ satisfies } 0 < |x_n - a| < 1/n \text{ and } |f(x_n) - l| \geq \epsilon. \end{array}$$

Then by the Archimedean property of the real numbers (or alternatively, the sandwich rule), we have $\lim_{n\to\infty} x_n = a$, but by the above construction the sequence $(f(x_n))$ does not converge to l.

So we have shown that if the $(\epsilon - \delta)$ criterion fails, then the function f does not have a limit at a. So it follows that if f does have a limit at a, then the $(\epsilon - \delta)$ criterion must hold. □

We have met the sandwich rule for sequences (Theorem 2.2.2). The next result gives a sandwich rule for functions.

Theorem 3.2.3 (Sandwich Rule for Functions)**.** *Suppose that $f, g, h : \mathbb{R} \to \mathbb{R}$ and suppose that there exists an interval $(a, b) \subseteq D_f \cap D_g \cap D_h$ so that for all $x \in (a, b)$*

$$f(x) \leq g(x) \leq h(x).$$

If there exists $c \in (a, b)$ such that $\lim_{x\to c} f(x) = \lim_{x\to c} h(x) = l$, then $\lim_{x\to c} g(x)$ exists and equals l.

[2] i.e. we use the fact that if P and Q are propositions, then $P \Rightarrow Q$ if and only if $\sim Q \Rightarrow \sim P$ (see Proposition 4.14 in MAS114(Sem 1)notes).

Proof. Let (x_n) be any sequence that converges to c, and note that by taking $N \in \mathbb{N}$ sufficiently large, we can ensure that $x_n \in (a, b)$ for all $n > N$. Then the result follows by the sandwich rule for sequences (Theorem 2.2.2), by taking the sequences there to be $a_n = f(x_{n+N}), b_n = g(x_{n+N})$ and $c_n = h(x_{n+N})$ for all $n \in \mathbb{N}$. $\square$

3.3 Divergence

We say that a function $f : \mathbb{R} \to \mathbb{R}$ *diverges* at $x = a$ if $\lim_{x\to a} f(x)$ does not exist. There are some interesting types of behaviour involved here. For example, we say that the function f *diverges to infinity* at $x = a$ and we write $\lim_{x\to a} f(x) = \infty$ if for every sequence (x_n), with $x_n \in D_f \setminus \{a\}$, which satisfies $\lim_{n\to\infty} x_n = a$, we have $\lim_{n\to\infty} f(x_n) = \infty$, i.e. the sequence $(f(x_n))$ diverges to infinity in the sense given in Chapter 2. The notion of *divergence to minus infinity* is defined similarly, and in that case we write $\lim_{x\to a} f(x) = -\infty$. For an example of divergence to infinity, consider e.g. $\lim_{x\to 0} 1/x^2$. The details are left to you.

To investigate divergence further, suppose that the following occurs: whenever (x_n) is a sequence in D_f converging to a, with $x_n < a$ for all $n \in \mathbb{N}$, then there exists $l_1 \in \mathbb{R}$, so that $\lim_{n\to\infty} f(x_n) = l_1$. In this case, we call l_1 the *left limit* of f at a, and we write $\lim_{x\uparrow a} f(x) = l_1$. Similarly, it can be the case that whenever (x_n) is a sequence in D_f converging to a, with $x_n > a$ for all $n \in \mathbb{N}$, then there exists $l_2 \in \mathbb{R}$, so that $\lim_{n\to\infty} f(x_n) = l_2$. In that case, we call l_2 the *right limit* of f at a, and we write $\lim_{x\downarrow a} f(x) = l_2$.[3] We will show below that a function f for which both left and right limits exist at a is divergent there if and only if $l_1 \neq l_2$.

For example, consider Heaviside's indicator function $H(x) = \mathbf{1}_{[0,\infty)}$, as defined in Example 3.4. It is not difficult to verify that $\lim_{x\uparrow 0} H(x) = 0$ and $\lim_{x\downarrow 0} H(x) = 1$. This is an example of a *discontinuity*; we will have more to say about continuous functions, and discontinuities, in Chapter 4.

There is an $(\epsilon - \delta)$ criterion for left and right limits:

Theorem 3.3.1. *Let f be a mapping from $\mathbb{R}$ to $\mathbb{R}$, then*

1. $\lim_{x\uparrow a} f(x) = l_1$, *if and only given any $\epsilon > 0$ there exists $\delta > 0$ so that whenever $x \in D_f$ with $0 < a - x < \delta$, then $|f(x) - l_1| < \epsilon$.*

2. $\lim_{x\downarrow a} f(x) = l_2$, *if and only given any $\epsilon > 0$ there exists $\delta > 0$ so that whenever $x \in D_f$ with $0 < x - a < \delta$, then $|f(x) - l_2| < \epsilon$.*

[3]Alternative notation that you may meet is to write $\lim_{x\to a-}$ instead of $\lim_{x\uparrow a}$, and $\lim_{x\to a+}$ instead of $\lim_{x\downarrow a}$.

The proof is very similar to that of Theorem 3.2.2, and is left to you to do as Problem 52. Note that $0 < a - x < \delta$ if and only if $a - \delta < x < a$ and $0 < x - a < \delta$ if and only if $a < x < a + \delta$.

Theorem 3.3.2. *A function $f : \mathbb{R} \to \mathbb{R}$ has a limit at a point $x = a$ if and only if both left and right limits exist at $x = a$ and they are equal (in which case the limit takes this same common value).*

Proof. Suppose f has a limit l at $x = a$, then for every sequence (x_n) with $x_n \in D_f \setminus \{a\}$ and $\lim_{n\to\infty} x_n = a$, we have $\lim_{n\to\infty} f(x_n) = l$. This includes all sequences that satisfy $x_n < a$, and also those for which $x_n > a$, for all $n \in \mathbb{N}$. It follows that left and right limits both exist, and are equal to l.

For the converse, it is best to use an $(\epsilon - \delta)$ argument. By Theorem 3.3.1, given any $\epsilon > 0$ there exists $\delta_1 > 0$ so that whenever $x \in D_f$ with $0 < a - x < \delta_1$, then $|f(x) - l| < \epsilon$, and there exists $\delta_2 > 0$ so that whenever $x \in D_f$ with $0 < x - a < \delta_2$, then $|f(x) - l| < \epsilon$. Now take $\delta = \min\{\delta_1, \delta_2\}$. We then have

$$\begin{aligned} 0 < |x - a| = \max\{a - x, x - a\} < \delta \Leftrightarrow &\quad 0 < a - x < \delta \text{ or } 0 < x - a < \delta \\ \Rightarrow &\quad 0 < a - x < \delta_1 \text{ or } 0 < x - a < \delta_2 \\ \Rightarrow &\quad |f(x) - l| < \epsilon, \end{aligned}$$

and so $\lim_{x\to\infty} f(x) = l$, by Theorem 3.2.2. □

We also meet functions that have divergent left and right limits. For example consider $f(x) = 1/x$, with $D_f = \mathbb{R} \setminus \{0\}$. Then it is easily verified that $\lim_{x\uparrow 0} f(x) = -\infty$, and $\lim_{x\downarrow 0} f(x) = \infty$.

Figure 3.2: Graph of $f(x) = 1/x$