Chapter 2

Limits of functions and continuity

2.1 Limits at infinity

Note it's not essential in Definition 1.1 that the quantity N is a positive **integer**. We could just have well decided that

> $\lim_{n\to\infty} a_n = L$ if, for each $\varepsilon > 0$, there exists $K \in \mathbb{R}$ such that, for all $n > K$, $|a_n - L| < \varepsilon$.

Both say that, given any positive number ε (no matter how small), there's a point in the sequence after which all terms lie within distance ε of L. We just use different means (a positive integer N or a real number K) to specify that point.

Having realized this, we see that the definition can be extended immediately to deal with *functions* of a real variable, not just sequences.

Definition 2.1 Let $D \subseteq \mathbb{R}$ be unbounded above and $f : D \to \mathbb{R}$. Then f has **limit L at infinity** if, for each $\varepsilon > 0$, there exists $K \in \mathbb{R}$ such that, for all $x \in D$ with $x > K$, $|f(x) - L| < \varepsilon$. The shorthand for this is $\lim_{x\to\infty} f(x) = L$.

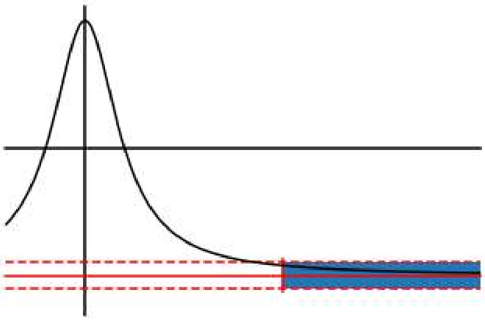
Example 2.2 Let $f : \mathbb{R} \to \mathbb{R}$, $f(x) = (1 - x^2)/(1 + x^2)$. Claim: $\lim_{x\to\infty} f(x) = -1$.

Proof: Let $\varepsilon > 0$ be given. Let $K = \sqrt{\frac{2}{\varepsilon}}$. Then, for all $x > K$,

$$\begin{aligned} |f(x) - (-1)| &= \left|\frac{1 - x^2}{1 + x^2} + 1\right| \\ &= \frac{2}{1 + x^2} \\ &< \frac{2}{x^2} \\ &< \frac{2}{K^2} = \varepsilon \end{aligned}$$

□

Again, it's useful to have a geometric picture of what this means:



Exercise 2.3 Let $f : [0, \infty) \to \mathbb{R}$ such that $f(x) = \dfrac{\sqrt{x}}{x+3}$. Prove that

$$\lim_{x\to\infty} f(x) = 0.$$

The definition says that, given any positive number ε (no matter how small), there's a real number K such that, for all $x > K$, $f(x)$ is closer than distance ε from L. This makes precise the informal idea that $f(x)$ gets "arbitrarily close" to L as x gets "sufficiently large." Note we must demand that D, the domain of f, is unbounded above, or else there will exist some real number K such that there *are* no values of $x \in D$ with $x > K$. If we allowed this situation, then *any* real number L would satisfy the definition of $\lim_{x\to\infty} f(x)$!

Exercise 2.4 Write down a precise definition of the limit of $f : D \to \mathbb{R}$ at **minus infinity**. What property must D have in order for this definition to make sense? Use your definition to prove that

$$\lim_{x\to-\infty} \frac{x^2-1}{x^2+1} = 1.$$

2.2 Cluster points

Limits at infinity are quite useful, but for the purposes of calculus, we really need to make proper sense of limits like

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

where a is a real number. For example,

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

Our first task is to clearly define the kind of point $a \in \mathbb{R}$ at which it makes sense to define the limit of a given function f. Informally, we know that the limit "as x tends to a" concerns the behaviour of $f(x)$ "sufficiently close" to (but *not at*) a. Just as $\lim_{x\to\infty} f(x)$ only makes sense if the domain of f is unbounded above (so that $f(x)$ is well-defined for x arbitrarily large), $\lim_{x\to a} f(x)$ only makes sense if the domain of f contains points which are "arbitrarily close" to (but different from) a.

Definition 2.5 Let $D \subseteq \mathbb{R}$. Then $a \in \mathbb{R}$ is a **cluster point** of D if, for each $\varepsilon > 0$, there exists $x \in D$ with $0 < |x - a| < \varepsilon$. Equivalently: for each $\varepsilon > 0$, the set $(D \backslash \{a\}) \cap (a - \varepsilon, a + \varepsilon)$ is nonempty.

Example 2.6 For each of the following subsets of $\mathbb{R}$, write down the set of all cluster points:

$$A = (0, 1], \qquad B = \mathbb{Q}, \qquad C = \{1/n : n \in \mathbb{Z}^+\}, \qquad D = \mathbb{Z}.$$

Solution Let's denote the set of cluster points of a set X by $\mathrm{cl}(X)$. Then

$$\begin{aligned}
\mathrm{cl}(A) &= [0, 1] \\
\mathrm{cl}(B) &= \mathbb{R} \\
\mathrm{cl}(C) &= \{0\} \\
\mathrm{cl}(D) &= \emptyset
\end{aligned}$$

From this we see that it's possible for a number that is *not* in a set to be a cluster point of that set (e.g. 0 is a cluster point of $(0, 1]$ and $\sqrt{2}$ is a cluster point of $\mathbb{Q}$), and that an element of a set may *fail* to be a cluster point of the set (e.g. 5 is not a cluster point of $\mathbb{Z}$).

We can also characterize cluster points using sequences. This often turns out to be more convenient than the original definition.

Proposition 2.7 *$a \in \mathbb{R}$ is a cluster point of $D \subseteq \mathbb{R}$ if and only if there is a sequence in $D \backslash \{a\}$ that converges to a.*

Proof: Exercise (see Problem Set 1). □

Note that the sequence is in $D \backslash \{a\}$, that is, it never takes the value a.

2.3 Limits of functions

Cluster points of the domain of a function are exactly the points at which it makes sense to define a limit of the function.

Definition 2.8 Let $D \subseteq \mathbb{R}$, $f : D \to \mathbb{R}$ and $a \in \mathbb{R}$ be a cluster point of D. Then f has **limit** L **at** a if, for each $\varepsilon > 0$, there exists $\delta > 0$ such that, for all $x \in D$ with $0 < |x - a| < \delta$, $|f(x) - L| < \varepsilon$. The shorthand for this is $\lim_{x \to a} f(x) = L$.

Example 2.9 Let $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^2$. Claim: $\lim_{x \to 2} f(x) = 4$.

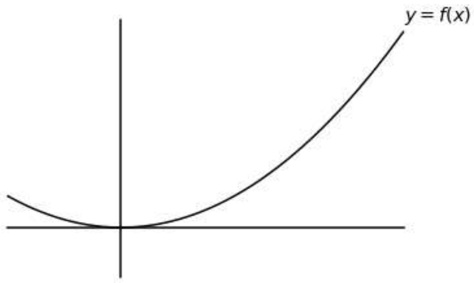
Proof: Let $\varepsilon \in (0, \infty)$ be given. Let $\delta = \boxed{\min\{1, \frac{\varepsilon}{5}\}}$.
Then, for all $x \in \mathbb{R}$ with $0 < |x-2| < \delta$,

$$\begin{aligned} |f(x) - 4| &= |x^2 - 4| \\ &= |x+2||x-2| \\ &\le 5|x-2| \qquad (\text{since } \delta \le 1 \text{ so } x \in (1,3)) \\ &< 5\delta \\ &\le \varepsilon \end{aligned}$$

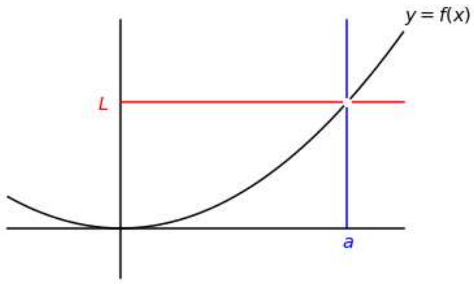
□

The definition says that, given any positive number ε (no matter how small), we can find another positive number δ which is so small that, whenever x is within distance δ of a, *and different from* a, $f(x)$ is within distance ε of L. This makes precise the informal idea that $f(x)$ is "arbitrarily close" to L for all x "sufficiently close" to, but different from, a.

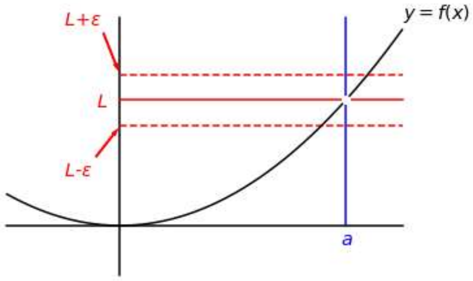
It's helpful to have a picture of what the definition means. Imagine the graph of the function f:



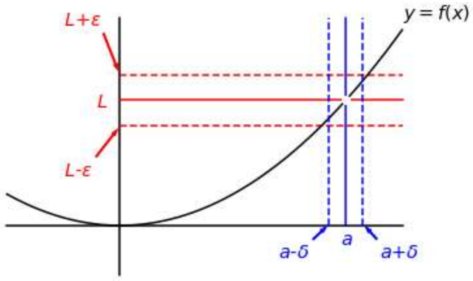
Delete from this the point $(a, f(a))$ (because the limit at a concerns the behaviour of f close to, but not at a) and mark on the lines $y = L$ and $x = a$, where L is the claimed limit:



Now, choose any $\varepsilon > 0$, and mark on the lines $y = L \pm \varepsilon$:



The claim is that, no matter how small we choose ε, there exists $\delta > 0$ such that the part of the graph between the vertical lines $x = a \pm \delta$ lies entirely between the horizontal lines $y = L \pm \varepsilon$:



A proof directly from the definition of limit (Definition 2.8), as in Example 2.9, is called a proof *from first principles*, or an ε—δ proof. It's very important that you get plenty of practice constructing such proofs, because this is the best way to properly understand the meaning of limits.

Example 2.10 Let $f : [0, \infty) \to \mathbb{R}$ be defined such that $f(x) = \sqrt{x}$. Claim: $\lim_{x \to 1} f(x) = 1$.

Proof: Let $\varepsilon \in (0, \infty)$ be given. Then let $\delta = \boxed{\varepsilon}$. Then, for all $x \in [0, \infty)$ such that $0 < |x - 1| < \delta$,

$$
\begin{aligned}
|f(x) - 1| &= |\sqrt{x} - 1| \\
&= \left| (\sqrt{x} - 1) \frac{\sqrt{x} + 1}{\sqrt{x} + 1} \right| \\
&= \frac{|x - 1|}{\sqrt{x} + 1} \\
&\leq |x - 1| \\
&< \delta = \varepsilon.
\end{aligned}
$$

□

In Examples 2.9 and 2.10, the point a at which we defined the limit was actually in the domain of f (and the limit happened to be $f(a)$). Limits are more interesting when a is *not* in the domain of f.

Example 2.11 Let $f : \mathbb{R}\backslash\{-3, 3\} \to \mathbb{R}$, $f(x) = \dfrac{x-3}{x^2-9}$. Claim: $\lim_{x\to 3} f(x) = \dfrac{1}{6}$.

Proof: Let $\varepsilon > 0$ be given. Let $\delta = \min\{1, 30\varepsilon\}$. Then, for all $x \in \mathbb{R}\backslash\{-3, 3\}$ with $0 < |x - 3| < \delta$,

$$\begin{aligned} \left|f(x) - \frac{1}{6}\right| &= \left|\frac{x-3}{(x-3)(x+3)} - \frac{1}{6}\right| = \left|\frac{1}{x+3} - \frac{1}{6}\right| = \left|\frac{3-x}{6(x+3)}\right| \\ &\le \frac{1}{30}|x-3| \qquad \text{(since } \delta \le 1 \text{ so } x \in (2,4) \text{ whence } x+3 > 5\text{)} \\ &< \frac{1}{30}\delta \\ &\le \varepsilon \end{aligned}$$

□

You may be wondering where the heck the number

$$\delta = \min\{1, 30\varepsilon\}$$

came from in the above proof. The answer is that, before I started to write the proof down, I did the following rough work: I want to show that I can make $f(x)$ as close as I like to $1/6$ by making x sufficiently close to (but different from) 3. So I try to find an upper bound on $|f(x) - 1/6|$ in terms of $|x - 3|$:

$$\begin{aligned} \left|f(x) - \frac{1}{6}\right| &= \left|\frac{x-3}{(x-3)(x+3)} - \frac{1}{6}\right| = \left|\frac{1}{x+3} - \frac{1}{6}\right| \\ &= \left|\frac{3-x}{6(x+3)}\right| \\ &= \frac{1}{6}\frac{1}{|x+3|}|x-3|. \end{aligned}$$

I like the factor of $|x - 3|$: this is the thing I can make as small as I like. The factor of $1/6$ is no problem. The factor of $1/|x + 3|$ is not so congenial, however. It can get arbitrarily large – but only if x gets close to -3. So, if I demand that $|x - 3| < 1$, for example, then $x > 2$ so $|x + 3| \ge x + 3 > 5$. This is helpful, because it forces $1/|x + 3| < 1/5$. So, for all x with $0 < |x - 3| < 1$, I know that

$$\left|f(x) - \frac{1}{6}\right| \le \frac{1}{6 \times 5}|x-3|.$$

To make this less than ε, it's enough to also demand that $|x - 3| < 30\varepsilon$. So, I need $\delta \le 1$ and $\delta < 30\varepsilon$, hence my choice of δ.

It's important to understand that there are infinitely many other correct choices of δ. This is the δ that I arrived at by reasoning as above. You may argue differently and come up with something different.

2.3.1 Some basic properties of limits

Since we have a precise definition of limits, we are in a position to prove some basic facts about them.

Theorem 2.12 (Uniqueness of limits) *If f has a limit at a, this limit is unique.*

Proof: Let $f : D \to \mathbb{R}$ and a be a cluster point of D. Assume, towards a contradiction, that both L_1 and $L_2 \neq L_1$ satisfy the definition of limit of f at a. Let $\varepsilon = |L_2 - L_1|/2 > 0$. Since $\lim\limits_{x\to a} f(x) = L_1$, there exists $\delta_1 > 0$ such that, for all $x \in D$ with $0 < |x - a| < \delta_1$, $|f(x) - L_1| < \varepsilon$. Similarly, since $\lim\limits_{x\to a} f(x) = L_2$, there exists $\delta_2 > 0$ such that, for all $x \in D$ with $0 < |x - a| < \delta_2$, $|f(x) - L_2| < \varepsilon$. Let $\delta = \min\{\delta_1, \delta_2\}$. **Since a is a cluster point of** D, there exists $x_* \in D$ such that $0 < |x_* - a| < \delta$. But then $0 < |x_* - a| < \delta_1$, so $|f(x_*) - L_1| < \varepsilon$, and, furthermore, $0 < |x_* - a| < \delta_2$, so $|f(x_*) - L_2| < \varepsilon$. Hence

$$2\varepsilon = |L_2 - L_1| = |f(x_*) - L_1 - (f(x_*) - L_2)| \leq |f(x_*) - L_1| + |f(x_*) - L_2| < \varepsilon + \varepsilon$$

a contradiction. □

So it makes sense to talk about *the* limit of a function at a. Note that this proof made essential use of the condition that a is a cluster point of D, the domain of f (see the text highlighted in bold above). Without this assumption, we can't be sure that the point $x_* \in D\backslash\{a\}$ distance less than δ from a actually exists.

We can also establish that limits of functions obey some basic arithmetic rules:

Theorem 2.13 (Algebra of Limits) *Let a be a cluster point of D, $f : D \to \mathbb{R}$, $g : D \to \mathbb{R}$, $\lim\limits_{x\to a} f(x) = L$ and $\lim\limits_{x\to a} g(x) = K$. Then*

(i) $\lim\limits_{x\to a}(f(x) + g(x)) = L + K$,

(ii) $\lim\limits_{x\to a} f(x)g(x) = LK$,

(iii) $\lim\limits_{x\to a} \dfrac{1}{f(x)} = \dfrac{1}{L}$ *if, in addition, $0 \notin f(D)$ and $L \neq 0$.*

We *could* give direct ε—δ proofs of each part of this theorem. They would be structurally very similar to the analogous proofs of Proposition 1.9 (the Algebra of Limits for sequences). This is actually quite a good exercise, and I encourage you to have a go. However, we're going to prove Theorem 2.13 by exploiting a sneaky and *very* powerful trick: we can reformulate limits of functions entirely in terms of convergence of sequences:

Theorem 2.14 *Let a be a cluster point of D and $f : D \to \mathbb{R}$. Then $\lim\limits_{x\to a} f(x) = L$ if and only if, for all sequences (x_n) in $D\backslash\{a\}$ such that $x_n \to a$, $f(x_n) \to L$.*

Proof: We must prove two things:

(i) If $\lim_{x\to a} f(x) = L$ then, for all sequences (x_n) in $D\backslash\{a\}$ such that $x_n \to a$, $f(x_n) \to L$, and

(ii) If for all sequences (x_n) in $D\backslash\{a\}$ such that $x_n \to a$, $f(x_n) \to L$, then $\lim_{x\to a} f(x) = L$.

(i) Assume $\lim_{x\to a} f(x) = L$ and let (x_n) be any sequence in $D\backslash\{a\}$ such that $x_n \to a$. We must show that $f(x_n) \to L$. So, let $\varepsilon > 0$ be given. By assumption, there exists $\delta > 0$ such that, for all $x \in D$ with $0 < |x-a| < \delta$, $|f(x) - L| < \varepsilon$. But $x_n \to a$, and δ is a positive number, so there exists $N \in \mathbb{Z}^+$ such that, for all $n \geq N$, $|x_n - a| < \delta$. Furthermore, $x_n \in D\backslash\{a\}$, so for all n, $|x_n - a| > 0$. Hence, for all $n \geq N$, $0 < |x_n - a| < \delta$. But then (by the definition of δ), for all $n \geq N$, $|f(x_n) - L| < \varepsilon$. Since $\varepsilon > 0$ was arbitrary, it follows that $f(x_n) \to L$.

(ii) We will prove the *contrapositive*: if f does **not** have limit L at a, then there exists a sequence (x_n) in $D\backslash\{a\}$ such that $x_n \to a$ but $f(x_n) \nrightarrow L$. So, assume $\lim_{x\to a} f(x) = L$ is **false**. Our first job is to figure out precisely what this condition means. It is the **negation** of the statement $\lim_{x\to a} f(x) = L$, that is

$$\neg\left[\forall \varepsilon \in (0,\infty), \exists \delta \in (0,\infty), \forall x \in (D\backslash\{a\}) \cap (a-\varepsilon, a+\varepsilon), |f(x) - L| < \varepsilon\right]$$

$$\Leftrightarrow \quad \exists \varepsilon \in (0,\infty), \neg\left[\exists \delta \in (0,\infty), \forall x \in (D\backslash\{a\}) \cap (a-\varepsilon, a+\varepsilon), |f(x) - L| < \varepsilon\right]$$

$$\Leftrightarrow \quad \exists \varepsilon \in (0,\infty), \forall \delta \in (0,\infty), \neg\left[\forall x \in (D\backslash\{a\}) \cap (a-\varepsilon, a+\varepsilon), |f(x) - L| < \varepsilon\right]$$

$$\Leftrightarrow \quad \exists \varepsilon \in (0,\infty), \forall \delta \in (0,\infty), \exists x \in (D\backslash\{a\}) \cap (a-\varepsilon, a+\varepsilon), \neg\left[|f(x) - L| < \varepsilon\right]$$

$$\Leftrightarrow \quad \exists \varepsilon \in (0,\infty), \forall \delta \in (0,\infty), \exists x \in (D\backslash\{a\}) \cap (a-\varepsilon, a+\varepsilon), |f(x) - L| \geq \varepsilon.$$

So *there exists* $\varepsilon > 0$ such that, for all $\delta > 0$, there is some $x \in D$ with $0 < |x-a| < \delta$ such that $|f(x) - L| \geq \varepsilon$. Since this holds for each and every $\delta > 0$, it holds, in particular, if $\delta = 1/n$, where $n \in \mathbb{Z}^+$. That is, for each $n \in \mathbb{Z}^+$, there exists $x_n \in D$ with $0 < |x_n - a| < 1/n$ such that $|f(x_n) - L| \geq \varepsilon$. Now $x_n \to a$ (by the **Squeeze Rule**), and $x_n \in D\backslash\{a\}$ (since $0 < |x_n - a|$), but $|f(x_n) - L| \geq \varepsilon$ for all n so certainly $f(x_n) \nrightarrow L$. $\square$

Armed with Theorem 2.14, we can now make short work of proving Theorem 2.13 – it follows almost immediately from Proposition 1.9:

Proof of Theorem 2.13: Let (x_n) be any sequence in $D\backslash\{a\}$ converging to a. Since $\lim_{x\to a} f(x) = L$, $f(x_n) \to L$ by Theorem 2.14. Since $\lim_{x\to a} g(x) = K$, $g(x_n) \to L$ by Theorem 2.14. Hence

(i) $f(x_n) + g(x_n) \to L + K$ by Proposition 1.9(i) and

(ii) $f(x_n)g(x_n) \to LK$ by Proposition 1.9(ii).

This holds for any such sequence (x_n) so $f(x) + g(x)$ has limit $L + K$ at a, and $f(x)g(x)$ has limit LK at a by Theorem 2.14.

If, in addition, $f(x) \neq 0$ for all $x \in D$ and $L \neq 0$, then $f(x_n) \neq 0$ for all n, so

(iii) $1/f(x_n) \to 1/L$ by Proposition 1.9(iii).

This holds for any such sequence (x_n) so $1/f(x)$ has limit $1/L$ at a by Theorem 2.14. □

Note that this proof used both directions of Theorem 2.14: since $\lim_{x\to a} f(x) = L$ we know that $f(x_n) \to L$ (the "only if" direction), and since $f(x_n) + g(x_n) \to L + K$ we know that $\lim_{x\to a}(f(x) + g(x)) = L + K$ (the "if" direction).

Exercise 2.15 Read through the proof of Theorem 2.13 and, at every place where Theorem 2.14 is cited, determine whether the "if" direction or the "only if" direction is being used.

2.3.2 Continuity and limits

You may have been told (by some shameless charlatan) that a function is continuous "if you can draw its graph without taking your pen off the paper." It should by now be clear that this is *not* an acceptable mathematical definition. What if you have no pen, or paper, or hands? What if you're simply not very good at drawing? Does the continuity of a function depend on your skill as a draughtsman? Clearly this is utter piffle. A mathematical definition should be formulated only in terms of the object being defined, not how we are able to interact with (or think about) it.

Luckily, we can use convergence of sequences to give a precise (and rather elegant) definition of continuity for functions:

Definition 2.16 Let $f : D \to \mathbb{R}$ and $a \in D$. Then f is **continuous at** a if, for all sequences (x_n) in D such that $x_n \to a$, $f(x_n) \to f(a)$. f is **continuous** if it is continuous at a for all $a \in D$. If f is not continuous, it is **discontinuous**.

For example, all polynomial functions

$$f : \mathbb{R} \to \mathbb{R}, \qquad f(x) = a_0 + a_1 x + \cdots + a_k x^k$$

are continuous. This follows from the Algebra of Limits for sequences and the rather obvious fact that $f(x) = x$ is continuous.[1]

Example 2.17 A simple example of a discontinuous function is a "step function":

$$f : \mathbb{R} \to \mathbb{R}, \qquad f(x) = \begin{cases} 0 & x < 0 \\ 1 & x \geq 0 \end{cases}$$

To see this, note that the sequence $x_n = -1/n \to 0$, but $f(x_n) = f(-1/n) = 0$, which does **not** converge to $f(0) = 1$.

[1] Proof: If $x_n \to a$ then $f(x_n) = x_n \to a = f(a)$.

A step function is discontinuous at a single point. Somewhat counterintuitively, it's possible for a function $f : \mathbb{R} \to \mathbb{R}$ to be *continuous* at only a single point:

Example 2.18 Let $f : \mathbb{R} \to \mathbb{R}$ be such that $f(x) = x$ if $x \in \mathbb{Q}$ and $f(x) = 0$ if $x \notin \mathbb{Q}$. Claim: f is continuous as 0 and discontinuous everywhere else.

Proof: Assume that f is continuous at a. Let (r_n) be any rational sequence converging to a. Then $f(r_n) = r_n \to a$. Let (i_n) be any irrational sequence converging to a. Then $f(i_n) = 0 \to 0$. But f is continuous at a, so $f(r_n) \to f(a)$ and $f(i_n) \to f(a)$. Hence $f(a) = a$ and $f(a) = 0$, that is $a = 0$. So *if* f is continuous at a *then* $a = 0$. That is, f is discontinuous at every $a \neq 0$.

It remains to show that f is continuous at 0 (careful: we have **not** proved this yet!). Let (x_n) be any sequence converging to 0. Then

$$|f(x_n)| = \left\{ \begin{array}{cl} |x_n|, & \text{if } x_n \in \mathbb{Q}, \\ 0, & \text{if } x_n \notin \mathbb{Q}, \end{array} \right\} \leq |x_n| \to 0$$

so $f(x_n) \to 0 = f(0)$ by the Squeeze Rule. Hence f is continuous at 0. □

Here's an even more counterintuitive example: a function $f : \mathbb{R} \to \mathbb{R}$ which is continuous at every irrational number, but discontinuous at every rational number! It was invented by Carl Johannes Thomae but, in the anglosphere, is usually referred to by the (somewhat mysterious) name given to it by John Horton Conway:

Example 2.19 (Stars Over Babylon) Let $f : \mathbb{R} \to \mathbb{R}$ be the function which maps every irrational number to 0, and maps each rational number p/q, expressed in lowest terms (that is, p and q have greatest common divisor 1), to $1/q$. So

$$\begin{aligned} f(\sqrt{2}) &= 0 \\ f(3/12) &= = f(1/4) = 1/4 \\ f(5\frac{1}{2}) &= f(11/2) = \frac{1}{2} \\ f(0) &= f(0/1) = 1, \end{aligned}$$

for example. Note that $f(x) \geq 0$ for all $x \in \mathbb{R}$, and $f(x) = 0$ if **and only if** x is irrational.

Claim 1: f is discontinuous at every $a \in \mathbb{Q}$.

Proof: Assume, towards a contradiction, that $a \in \mathbb{Q}$ and f is continuous at a. Then, by the density of the irrationals in the reals, for each $n \in \mathbb{Z}^+$, there exists $i_n \in \mathbb{R}\backslash\mathbb{Q}$ such that $a < i_n < a + \frac{1}{n}$. By the Squeeze Rule, $i_n \to a$. Since f is continuous at a, $f(i_n) \to f(a)$ and $f(a) > 0$ (since a is rational). But $f(i_n)$ is the constant sequence 0, so $f(i_n) \to 0$, a contradiction. □

Claim 2: f is continuous at every $a \in \mathbb{R}\backslash\mathbb{Q}$.

Proof: Again, assume towards a contradiction, that $a \in \mathbb{R}\backslash\mathbb{Q}$ and f is *discontinuous* at a. Then there exists a sequence x_n such that $x_n \to a$ but $f(x_n) \nrightarrow f(a) = 0$. Hence, there exists $\varepsilon > 0$ such that, for each $k \in \mathbb{Z}^+$, there exists $n_k \geq k$ such that $f(x_{n_k}) = |f(x_{n_k}) - 0| \geq \varepsilon$. That is, there is a subsequence $y_k = x_{n_k}$ of (x_n) such that $f(y_k) \geq \varepsilon > 0$. Clearly, (y_k) is a sequence of *rational numbers* (since $f(x) = 0$ if x is *irrational*).

Consider the rational sequence (y_k). Since it's a subsequence of a sequence converging to a, it converges to a. Since it's rational, we can write it in the form

$$y_k = \frac{p_k}{q_k}$$

where $p_k \in \mathbb{Z}$, $q_k \in \mathbb{Z}^+$ and $\gcd(p_k, q_k) = 1$. Then $f(y_k) = 1/q_k$. But $f(y_k) \geq \varepsilon$, so

$$0 < q_k \leq \frac{1}{\varepsilon},$$

that is, the sequence (q_k) is bounded. Hence, by the Bolzano-Weierstrass Theorem, it has a convergent subsequence, (q_{k_l}) say. But any convergent sequence of integers is eventually constant, that is, there exists $L \in \mathbb{Z}^+$ such that, for all $l \geq L$, $q_{k_l} = q_{k_L} = q$ say. (To see this, note that (q_{k_l}) is Cauchy , so there exists $L \in \mathbb{Z}^+$ such that, for all $m \geq L$, $|q_{k_L} - q_{k_m}| < 1/2$, and the only integer within distance $1/2$ of q_{k_L} is q_{k_L} itself.) Consider now the subsequence

$$y_{k_l} = \frac{p_{k_l}}{q_{k_l}} = \frac{p_{k_l}}{q},$$

for all $l \geq L$. This converges to a, so $p_{k_l} \to qa$, which is irrational (note $q > 0$). But (p_{k_l}) is also a sequence of integers, so if it converges, it is eventually constant, and hence its limit is an integer. This is a contradiction. $\square$

It's not hard to show that every function $f : D \to \mathbb{R}$ is automatically continuous at a if a is *not* a cluster point of D.

Proposition 2.20 *Let $f : D \to \mathbb{R}$ and assume $a \in D$ is not a cluster point of D. Then f is continuous at a.*

Proof: Let (x_n) be any sequence in D converging to a. Since a is *not* a cluster point of D, there exists $\varepsilon > 0$ such that no element of D has $0 < |x - a| < \varepsilon$. Hence, if $|x-a| < \varepsilon$ and $x \in D$, then $x = a$. Now $x_n \to a$, so there exists $N \in \mathbb{Z}^+$ such that, for all $n \geq N$, $|x_n - a| < \varepsilon$ and hence $x_n = a$. Hence, for all $n \geq N$, $f(x_n) = f(a)$, that is, the sequence $(f(x_n))$ is constant and equals $f(a)$, for all $n \geq N$. Hence $f(x_n) \to f(a)$. $\square$

So in considering whether a function is continuous, we may restrict attention only to cluster points of (and in) its domain. For such points we have two useful ways to reformulate continuity:

Theorem 2.21 *Let $f : D \to \mathbb{R}$, $a \in D$ and a be a cluster point of D. Then the following are equivalent:*

(i) f is continuous at a;

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

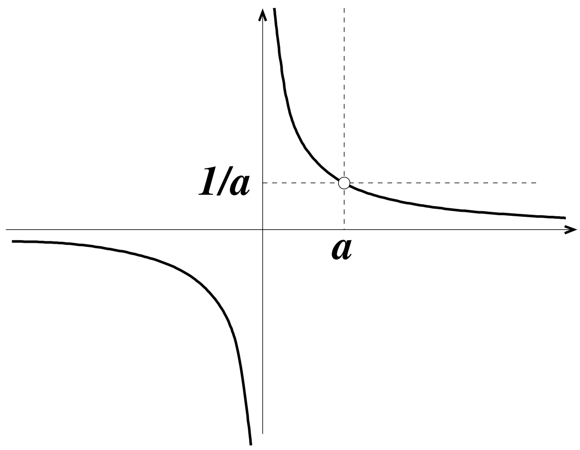
(iii) For each $\varepsilon > 0$ there exists $\delta > 0$ such that, for all $x \in D$ with $|x - a| < \delta$, $|f(x) - f(a)| < \varepsilon$.

Proof: (i) $\Rightarrow$ (ii) follows immediately from Theorem 2.14.
(ii) $\Rightarrow$ (iii): assume $\lim_{x \to a} f(x) = f(a)$ and let $\varepsilon > 0$ be given. Then (by Definition 2.8) there exists $\delta > 0$ such that, for all $x \in D$ with $0 < |x - a| < \delta$, $|f(x) - f(a)| < \varepsilon$. Clearly, if $x = a$ then $|f(x) - f(a)| = 0 < \varepsilon$. Hence, for all $x \in D$ with $|x - a| < \delta$, $|f(x) - f(a)| < \varepsilon$.
(iii) $\Rightarrow$ (i): assume (iii) holds, and let (x_n) be any sequence in D converging to a. We must show that $f(x_n) \to L$. So, let $\varepsilon > 0$ be given. By assumption, there exists $\delta > 0$ such that, for all $x \in D$ with $|x - a| < \delta$, $|f(x) - f(a)| < \varepsilon$. Now $x_n \to a$ and $\delta > 0$, so there exists $N \in \mathbb{Z}^+$ such that, for all $n \geq N$, $|x_n - a| < \delta$. But then, for all $n \geq N$, $|f(x_n) - f(a)| < \varepsilon$, by definition of δ. Hence $f(x_n) \to f(a)$. Since this holds for any sequence (x_n) in D converging to a, f is continuous at a. $\square$

Many mathematicians take condition (iii) in Theorem 2.21 as the *definition* of continuity at a, instead of the sequential definition, Definition 2.16. It's important to be able to use both. Here's an example which demonstrates that the whole "continuous if you can draw its graph without taking your pen off the paper" pseudo-definition isn't just embarrassingly jejeune, it's also plain wrong.

Example 2.22 Let $f : \mathbb{R}\backslash\{0\} \to \mathbb{R}$, $f(x) = \dfrac{1}{x}$. Claim: f is continuous.[2]

Proof: We must show that f is continuous at a for every $a \neq 0$. To do this , we will use Theorem 2.21, that is, show that f, a satisfy the ε—δ criterion of part (iii).



So, choose and fix $a \neq 0$ and let some positive number ε be given. We must show that there exists $\delta > 0$ such that, for all $x \in \mathbb{R}\backslash\{0\}$ with $|x - a| < \delta$, $|f(x) - f(a)| < \varepsilon$.

[2]Try drawing the graph of that without taking your pen off the paper!

Now

$$
\begin{aligned}
|f(x) - f(a)| &= \left| \frac{1}{x} - \frac{1}{a} \right| && (2.1) \\
& && (2.2) \\
&= \frac{|x-a|}{|a||x|} && (2.3)
\end{aligned}
$$

So first let's insist that $\delta \leq |a|/2$. Then, for all $x \in (a-\delta, a+\delta)$, $|x| \geq |a|/2$, and hence

$$\frac{1}{|x|} \leq \frac{2}{|a|}$$

Given this, we see from (2.3) that, to ensure $|f(x) - f(a)| < \varepsilon$, it suffices to require that $|x-a| < |a|^2\varepsilon/2$. We may now write out the argument explicitly.

Given any $\varepsilon > 0$, let $\delta = \min\{|a|/2, |a|^2\varepsilon/2\} > 0$. Then for all $x \in \mathbb{R}\backslash\{0\}$ with $|x-a| < \delta$,

$$
\begin{aligned}
|f(x) - f(a)| &= \frac{|x-a|}{|a||x|} \\
&\leq \frac{2}{|a|^2}|x-a| \qquad (\text{since } |x| > |a| - \delta \geq |a|/2) \\
&< \frac{2}{a^2}\delta \\
&\leq \varepsilon.
\end{aligned}
$$

Hence, by Theorem 2.21, f is continuous at a. Since $a \in \mathbb{R}\backslash\{0\}$ was arbitrary, f is continuous. □

This kind of argument is called an ε—δ proof of continuity. Often an argument making direct use of Definition 2.16 is simpler.

Example 2.22 revisited Prove that $f : \mathbb{R}\backslash\{0\} \to \mathbb{R}$, $f(x) = \dfrac{1}{x}$ is continuous directly from Definition 2.16.

Proof: Let $a \in \mathbb{R}\backslash\{0\}$ and (x_n) be any sequence in $\mathbb{R}\backslash\{0\}$ such that $x_n \to a$. Then

$$f(x_n) = \frac{1}{x_n} \to \frac{1}{a} = f(a)$$

by the Algebra of Limits. Hence f is continuous at a. □

Here's another result which follows very quickly from the sequential definition of continuity (the composition of two continuous functions is continuous):

Theorem 2.23 *Let D and E be subsets of $\mathbb{R}$, $f : D \to E$, $g : E \to \mathbb{R}$, f be continuous at a and g be continuous at $f(a)$. Then $g \circ f : D \to \mathbb{R}$ is continuous at a.*

Proof: We will show directly that $g \circ f : D \to \mathbb{R}$, $(g \circ f)(x) = g(f(x))$ satisfies Definition 2.16. So, let (x_n) be any sequence in D converging to a. Since f is continuous at a, $y_n := f(x_n) \to f(a)$. Hence, since g is continuous at $f(a)$, $g(y_n) \to g(f(a))$. Hence $g(f(x_n)) \to g(f(a))$. $\square$

Exercise 2.24 Give an ε—δ proof of Theorem 2.23.

We finish this chapter by recalling two famous and important theorems about continuous functions on a closed, bounded *interval*,

$$[a, b] = \{x \in \mathbb{R} \,:\, a \leq x \leq b\}.$$

Theorem 2.25 (Intermediate Value Theorem) *Let $f : [a, b] \to \mathbb{R}$ be continuous and y be any number between $f(a)$ and $f(b)$. Then there exists $c \in [a, b]$ such that $f(c) = y$.*

Theorem 2.26 (Extreme Value Theorem) *Let $f : [a, b] \to \mathbb{R}$ be continuous. Then f is bounded (above and below) and attains both a minimum and a maximum value.*

These two Theorems may look obvious, but it's important to realize that we can (and, in MATH1026, *did*) prove them rigorously from Definitions 2.16 and 1.1. The whole point of Real Analysis is to go beyond intuitive, hand waving reasoning, to back up all our assertions with precise, rigorous reasoning. After all, as we shall see, not everything that is "obvious" is true!

Summary

- A real number a is a **cluster point** of a set $D \subseteq \mathbb{R}$ if, for each $\delta > 0$, there exists $x \in D$ with $0 < |x - a| < \delta$.
- A function $f : D \to \mathbb{R}$ has **limit** L at a, a cluster point of D, if, for each $\varepsilon > 0$, there exists $\delta > 0$ such that, for all $x \in D$ with $0 < |x - a| < \delta$, $|f(x) - L| < \varepsilon$. The shorthand for this is $\lim\limits_{x \to a} f(x) = L$.
- We can reformulate limits in terms of convergence of sequences:
  Theorem $\lim\limits_{x \to a} f(x) = L$ if and only if, for all sequences (x_n) in $D \backslash \{a\}$ converging to a, $(f(x_n))$ converges to L.
- $f : D \to \mathbb{R}$ is **continuous** at $a \in D$ if, for all sequences (x_n) in D converging to a, $f(x_n)$ converges to $f(a)$.
- **Theorem** If a is a cluster point of D (the domain of f), the following are equivalent:
  (i) f is continuous at a
  (ii) $\lim\limits_{x \to a} f(x) = f(a)$
  (iii) For each $\varepsilon > 0$ there exists $\delta > 0$ such that, for all $x \in D$ with $|x - a| < \delta$, $|f(x) - f(a)| < \varepsilon$.