G12RAN: Real Analysis

4. Sequences and Continuous functions

Continuous functions

Definition Let I be an interval in $\mathbb{R}$ such that I has more than one point in it and let f be a function from I to $\mathbb{R}$. Let $a \in I$. If a is not an endpoint of I then we say that f is *continuous at* a if

$$\lim_{x \to a-} f(x) = \lim_{x \to a+} f(x) = f(a), \qquad (*)$$

i.e. both one-sided limits exist and are equal to the value of the function at the point, $f(a)$.

If, instead, a is an endpoint of the interval I (e.g. if $I = [a, b]$) then only one of the one-sided limits makes sense, and we say that f is *continuous at* a if this one-sided limit exists and equals $f(a)$.

Note also that $(*)$ above is the same as saying that

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

This time the value of the function at a does matter (unlike in the definition of limit values in Chapter 3).

If f is not continuous at a then f is *discontinuous at* a. The function f is *continuous* (from I to $\mathbb{R}$) if it is continuous at every point of I. Otherwise the function f is *discontinuous* (i.e. there is at least one point of I at which f is discontinuous).

An equivalent definition of continuity in terms of sequences is as follows. With I, a as above, we say that f is *continuous at* a if the following condition holds: for every sequence (x_n) in I which converges to a, we have

$$\lim_{n \to \infty} f(x_n) = f(a).$$

(Note that, unlike with the definition of limit values from Chapter 3, some or all of the x_n may equal a: does this make any difference?)

This sequence version of the definition is valid both for points a which are endpoints of I and for points which are not endpoints of I, so it may be easier to use. On the other hand, if I is an open interval or $I = \mathbb{R}$ then none of the points in I are endpoints of I and in this case both definitions are very easy to work with.

In terms of sequences, the function f is continuous from I to $\mathbb{R}$ if and only if the following condition holds: for all sequences $(x_n) \subseteq I$ which converge to some point of I, the sequence $(f(x_n))$ also converges and

$$\lim_{n \to \infty} f(x_n) = f(\lim_{n \to \infty} x_n).$$

You should also be aware of the $\epsilon - \delta$ definition of continuity used in most standard texts. This is equivalent to our definition: we state it here for information.

Definition Let A, B be subsets of $\mathbb{R}$ and let f be a function from A to B.

Let $a \in A$. Then f is *continuous at a* if for all $\epsilon > 0$ there is a $\delta > 0$ such that for all x in A with $|x - a| < \delta$ we have $|f(x) - f(a)| < \epsilon$ (so, for $x \in A \cap (a - \delta, a + \delta)$ we have $f(x) \in (f(a) - \epsilon, f(a) + \epsilon)$).

Typical examples of continuous functions include all polynomial functions, and standard functions such as $\sin(x)$, $\cos(x)$, $\exp(x)$ *etc.* (where they are defined: e.g. the function $\tan(x)$ is continuous except at the points where it is undefined. The function $\tan(x)$ is undefined whenever $\cos(x) = 0$).

In order to prove the theorems we want about continuous functions, we will need some more results about sequences.

Sequences revisited

Definition Let $(x_n) \subseteq \mathbb{R}$. Then a *subsequence* of (x_n) is a sequence of the form $x_{n_1}, x_{n_2}, x_{n_3}, \dots$ (or $(x_{n_k})_{k=1}^{\infty}$) with $n_1 < n_2 < n_3 < \dots$ (and where all the n_k are in $\mathbb{N}$). If we say that (y_n) is a subsequence of (x_n), it means that there is a strictly increasing sequence of positive integers $(n_k)_{k=1}^{\infty}$ as above with $y_k = x_{n_k}$ for $k = 1, 2, 3, \dots$.

Theorem 4.1 (Bolzano-Weierstrass Theorem) Let (x_n) be a bounded sequence of real numbers. Then (x_n) has a convergent subsequence.

Remark. Of course, the sequence (x_n) need not itself converge, as the sequence $x_n = (-1)^n$ shows. The same example shows that a sequence may have different convergent subsequences with different limits. For this example, some subsequences converge to 1, some converge to -1, and the remaining subsequences diverge.

Main theorems about continuous functions

Here are two of the most important results about continuous functions.

Theorem 4.2 (Boundedness Theorem) Let a, b be real numbers with $a \leq b$, and let $f : [a, b] \to \mathbb{R}$ be a continuous function. Then f is bounded on $[a, b]$ and attains its bounds.

Remark: this means that, under these conditions,there must exist points x_1, x_2 in $[a, b]$ such that, for all $x \in [a, b]$,

$$f(x_1) \leq f(x) \leq f(x_2).$$

Theorem 4.3 (Intermediate Value Theorem) Let a, b be real numbers with $a < b$, and let f be a continuous function from $[a, b]$ to $\mathbb{R}$. Then for every c between $f(a)$ and $f(b)$ there exists d in $[a, b]$ such that $f(d) = c$.

More advanced results on sequences (NON-EXAMINABLE)

We conclude with some more advanced concepts, which are the *lower and upper limits* of a sequence. These are of vital importance throughout more advanced Analysis, but it is possible to understand this module without them.

If we have a bounded sequence of real numbers x_n, but we do not yet know whether it converges, then we can not use the notation $\lim_{n\to\infty}(x_n)$. But we do know that (x_n) has at least one convergent subsequence.

If you look at the set of all possible limits of subsequences of your sequence (x_n), you obtain a non-empty, bounded subset of $\mathbb{R}$. It turns out that this set always has a greatest element (the maximum possible limit of a convergent subsequence of (x_n)) and a least element (the minimum possible limit of a convergent subsequence of (x_n)). These are called, respectively, the *upper limit* and the *lower limit* of the sequence (x_n). The upper limit of the sequence is also called the *limit supremum* of (x_n) and is denoted by $\limsup_{n\to\infty}(x_n)$. The lower limit is also called the *limit infimum* of the sequence (x_n) and is denoted by $\liminf_{n\to\infty}(x_n)$. A more explicit definition is the following.

Definition Let (x_n) be a bounded sequence of real numbers. For each $n \in \mathbb{N}$, set $E_n = \{x_n, x_{n+1}, x_{n+2}, \ldots\}$, (a non-empty bounded subset of $\mathbb{R}$) and set

$$s_n = \inf(E_n), \quad S_n = \sup(E_n).$$

It is easy to see that the sequences (s_n) and (S_n) are both bounded and monotone, so must converge, and we can define

$$\liminf_{n\to\infty}(x_n) = \lim_{n\to\infty}(s_n), \quad \limsup_{n\to\infty}(x_n) = \lim_{n\to\infty}(S_n).$$

When $x_n = (-1)^n$, for example, we have

$$\liminf_{n\to\infty}(x_n) = -1 \qquad \text{and} \qquad \limsup_{n\to\infty}(x_n) = 1.$$

Note that the lower and upper limits of a bounded sequence *always exist*, whether or not the sequence converges, and this makes them very useful when dealing with sequences whose convergence we are unsure of. The following result gathers together some standard facts about lim inf and lim sup. As an exercise, you may convince yourselves of the details.

Theorem 4.4 Let (x_n) be a bounded sequence of real numbers. Then

(i) $\liminf_{n\to\infty}(x_n) \le \limsup_{n\to\infty}(x_n)$,

(ii) $\limsup_{n\to\infty}(x_n) = -\liminf_{n\to\infty}(-x_n)$

(iii) The sequence (x_n) converges if and only if

$$\liminf_{n\to\infty}(x_n) = \limsup_{n\to\infty}(x_n).$$

If (x_n) converges, then

$$\lim_{n\to\infty}(x_n) = \liminf_{n\to\infty}(x_n) = \limsup_{n\to\infty}(x_n).$$