Getting acquainted with ε and δ

Coffee & Chalk Press

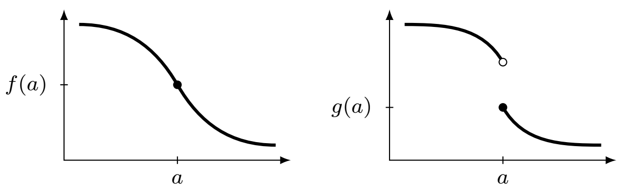
Ivica Smolić

Intro

After the long historical struggle with the notions of the infinitesimal, the continuity and the limit, modern definitions and notation appeared in the works of the 19th century mathematicians, most notably Bernard Bolzano and Karl Weierstraß. In this brief lecture we shall look under the hood of the modern notions of continuous functions and limits.

Continuity of functions

Below we give two examples, on the left is the graph of a function f, continuous at the point a, while on the right is the graph of a function g, which has a discontinuity (jump) at the point a.



So, how to precisely define the qualitative difference between the behaviour of functions f and g at the point a? In other words, how to precisely define when some function is *continuous* at a given point in its domain?

Prosaic descriptions

There are two vague ways to explain the continuity of a function f at the point a, either

"If we move a little bit away from a in the domain, then the values of the map f move a little bit away from $f(a)$ in the codomain."

or

"If we take some point in the codomain close to $f(a)$, then there is some point close to a which f maps to this shifted value."

None of this is true for the function g: If we move to the left from a it immediately "jumps" to some larger values and there are values just above the $g(a)$ to which none of the points close to a are mapped by g. In order to make all this more concrete we have to find an appropriate mathematical conceptualization of the phrases "little bit away" and "close to" that we've used above.

Neighbourhoods

Suppose that we are given some point of the real line and we want to describe which points are "close to it". One way to do this is to choose a fixed scale, say a real number $s > 0$, and then proclaim that all points $y \in \mathbb{R}$ for which $|y - x| < s$ holds are "close" to the given $x \in \mathbb{R}$. All definitions built from such a notion would suffer from a question that would immediately follow after any of them: Does this definition depend on the choice of s and, if yes, how exactly? We can avoid this inconvenience if we just give up on the *fixed* scale s, and go with the *arbitrary* neighbourhoods of a point.

Let us first recall what are the open subsets of the real numbers. We say that a set $O \subseteq \mathbb{R}$ is **open** if for any point $x \in O$ there is $\delta > 0$, such that $\langle x - \delta, x + \delta \rangle \subseteq O$. There are numerous examples, such as the empty set $\emptyset$, the whole $\mathbb{R}$, all intervals $\langle a, b \rangle$ with $a < b$, all unions $\langle a, b \rangle \cup \langle c, d \rangle$ with $a < b < c < d$, etc. On the other hand the sets $[a, b]$ and $[a, b\rangle$ are *not* open since in both cases there is no $\delta > 0$, such that the interval $\langle a - \delta, a + \delta \rangle$ is a subset of a given set.

A **neighbourhood** of a point $x \in \mathbb{R}$ is any *open subset* of $\mathbb{R}$, containing that point. For example, $O = \langle -1, 2 \rangle$, $U = \langle 0, 3 \rangle \cup \langle 5, 6 \rangle$ and $V = \mathbb{R}$ are three examples of neighbourhoods of the point $x = 1$. Note that a neighbourhood of a point doesn't have to be a *connected* open set! One particularly simple and useful type of a neighbourhood is an open interval centered at the given point: For any $x \in \mathbb{R}$ and $\varepsilon > 0$ we introduce

$$B(x, \varepsilon) := \langle x - \varepsilon, x + \varepsilon \rangle \ ,$$

with "B" standing for "ball" (here, an open 1-dimensional ball).

Building the definition

Now we can take a step closer to the precise mathematical definition, by utilizing the neighbourhoods in translation of the two vague descriptions from above.

$\star$ Attempt #1.

"Each neighbourhood of the point a is mapped by f to a neighbourhood of $f(a)$."

This definition, however, would exclude some functions which we would like to call "continuous". For example, the function $f(x) = x^2 : \mathbb{R} \to \mathbb{R}$ certainly looks continuous at $a = 0$, but there is no neighbourhood of this point which f maps to a neighbourhood of $f(0) = 0$.

$\star$ Attempt #2.

"For each neighbourhood V of $f(a)$, there is a neighbourhood O of a which is mapped by f to a subset of V."

Although seemingly more complicated than the previous one (it sounds somehow "backwards"), this definition will turn out to be more useful and closer to our intuitive notion of "continuity".

We can make it a little bit more economical by replacing completely arbitrary neighbourhoods with the simple ones, $V = B(f(a), \varepsilon)$ and $O = B(a, \delta)$. Do we lose any generality with such a choice? No, since for any neighbourhood W of $f(a)$, by definition of an open set, there is an open ball $B(f(a), \varepsilon) \subseteq W$ and, by our economized definition, an associated neighbourhood $B(a, \delta)$, open set which is mapped by f into a subset of $B(f(a), \varepsilon)$, hence also a subset of W.

Polishing the phrasing

We're almost there, but first we'll introduce one more useful notation. For any map $f : X \to Y$ and a subset $A \subseteq X$, by $f(A)$ we denote the **image** of A, that is the *set* of all points to which f maps points from A,

$$f(A) := \{f(a) \in Y \mid a \in A\} \ .$$

For example, if $f(x) = x^2 : \mathbb{R} \to \mathbb{R}$, then $f([-1, 2]) = [0, 4]$.

Finally, we say that the function $f : \mathbb{R} \to \mathbb{R}$ is **continuous at the point** $a \in \mathbb{R}$ if for each real $\varepsilon > 0$ (this number measures the size of the neighbourhood in the codomain) there exists a real $\delta > 0$ (this number measures the size of the neighbourhood in the domain), such that

$$f(B(a, \delta)) \subseteq B(f(a), \varepsilon) \, .$$

There are various ways to rephrase this definition. First note that $f(A) \subseteq Z$ means that $f(a) \in Z$ holds for any $a \in A$. Thus, we may say that f is continuous at a if for all $\varepsilon > 0$ there exists $\delta > 0$, such that $x \in B(a, \delta)$ implies $f(x) \in B(f(a), \varepsilon)$, that is

$$(\forall \, \varepsilon > 0)(\exists \, \delta > 0) : \ x \in B(a, \delta) \ \Rightarrow \ f(x) \in B(f(a), \varepsilon) \, .$$

Yet another equivalent formulation of this definition, the one which is more frequently found in textbooks on real analysis, says that f is continuous at a if

$$(\forall\, \varepsilon > 0)(\exists\, \delta > 0) : \; |x - a| < \delta \;\Rightarrow\; |f(x) - f(a)| < \varepsilon \, .$$

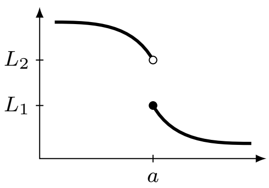
One technical remark: What if the domain of f is not the entire $\mathbb{R}$ but some other nonempty subset $D \subseteq \mathbb{R}$? Well, we just have to be careful that all the objects in the definition make sense, most notably, that we always insert into $f(x)$ a point x from the domain D. Hence, we say a function $f : D \to \mathbb{R}$ is continuous at the point $a \in D$ if

$$(\forall\, \varepsilon > 0)(\exists\, \delta > 0)(\forall\, x \in D) : \; x \in B(a, \delta) \;\Rightarrow\; f(x) \in B(f(a), \varepsilon) \, .$$

One must not be misled by the seemingly "broken appearance" of some function. For example, the function $h(x) = 1/x : \mathbb{R}^\times \to \mathbb{R}$, where $\mathbb{R}^\times = \langle -\infty, 0 \rangle \cup \langle 0, +\infty \rangle$, is continuous at all points of its domain, despite the "jump at the origin" ($x = 0$ is not an element of its domain!).

Limits

Now, let us look back again at the graph of the function g from above.



We see that, as we approach the point a from the *right*, the values of the function g approach the value L_1. On the other hand, as we approach the point a from the *left*, the values of the function g approach the value L_2 (despite the fact that $g(a) = L_1$). The process of "approaching" lies at the heart of the notion of the "limit" and we would like to crystallize it into a proper mathematical definition. Again, we can reach for the concept of neighbourhoods to build the definition that we aspire to. For example, if we introduce the right neighbourhood $\langle a, a + \delta \rangle$ and the left neighbourhood $\langle a - \delta, a \rangle$ of a point a, then the definition would run as follows,

"The L_1 is the *limit from the right* of g at a, in a sense that for any neighbourhood V of L_1 there is a right neighbourhood of a which is mapped by g into a subset of V."

More precisely, for any function $g : D \to \mathbb{R}$, defined on a nonempty domain $D \subseteq \mathbb{R}$, we write

$$L_1 = \lim_{x \to a^+} g(x) := \lim_{\varepsilon \to 0^+} g(a + \varepsilon) \qquad \text{if}$$

$$(\forall\, \varepsilon > 0)(\exists\, \delta > 0)(\forall\, x \in D) : \; 0 < x - a < \delta \;\Rightarrow\; g(x) \in B(L_1, \varepsilon) \, .$$

Here the notation "$\to a^+$" and "$\to 0^+$" emphasizes that we are approaching to a and to 0 from the right. Completely analogously, we may define the limit from the left, by writing

$$L_2 = \lim_{x \to a^-} g(x) := \lim_{\varepsilon \to 0^+} g(a - \varepsilon) \qquad \text{if}$$

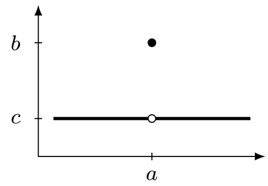
$$(\forall\, \varepsilon > 0)(\exists\, \delta > 0)(\forall\, x \in D) : \; 0 < a - x < \delta \;\Rightarrow\; g(x) \in B(L_2, \varepsilon) \, .$$

If these two one-sided limits exist and coincide then we speak of **the limit** of a function at a given point. More precisely, for any function $f : D \to \mathbb{R}$, defined on a nonempty domain $D \subseteq \mathbb{R}$, we write

$$L = \lim_{x \to a} f(x) \qquad \text{if}$$

$$(\forall\, \varepsilon > 0)(\exists\, \delta > 0)(\forall\, x \in D) : \; 0 < |x - a| < \delta \;\Rightarrow\; f(x) \in B(L, \varepsilon) \, .$$

A function may have a well-defined limit at a point, which is still different from the value of the function at that point. For example, suppose that $f : \mathbb{R} \to \mathbb{R}$ is defined as $f(a) = b$ and $f(x) = c \neq b$ for any $x \neq a$, as shown on the picture below,



Then

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

If we compare the definitions of the continuity of a function at a point with the definition of the limit at that point, it is not difficult to notice a striking similarity: The limit of a *continuous* function at a point of its domain is always well-defined and equal to the value of such a function at that point; Conversely, if the limit of a function f exists at a given point a of its domain and it is equal to the value $f(a)$ of that function, then f is continuous at the point a.

Note that in all these definitions of limits the point a doesn't have to be an element of the domain of the function. For example, if $f(x) = x + c : \langle 0, 1 \rangle \to \mathbb{R}$ with some $c \in \mathbb{R}$, then $\lim_{x \to 0^+} f(x) = c$. This may be cause of some confusion if we erroneously try to evaluate a limit just by inserting the value of the limiting argument in the function. For example, if $f(x) = \sin(x)/x : \mathbb{R}^\times \to \mathbb{R}$, then $\lim_{x \to 0} f(x) = 1$ is a well-defined limit, in contrast with the undefined expression $0/0$.

Curious examples

Here we take a look at some more involved examples. It is not difficult to invent a function which is continuous on all of its domain except at one point. But, how about the function which is nowhere continuous? A famous example is the **Dirichlet function** $\mathcal{D} : \mathbb{R} \to \mathbb{R}$, with value 1 at all the rational points of the real line, and zero at all the others,

$$\mathcal{D}(x) = \begin{cases} 1 \, , & x \in \mathbb{Q} \\ 0 \, , & x \in \mathbb{R} - \mathbb{Q} \end{cases}$$

A more intriguing example is a function which is continuous at *exactly one* point of its domain: The function $x\mathcal{D}(x) : \mathbb{R} \to \mathbb{R}$ is only continuous at the origin. First, for any $\varepsilon > 0$ we may choose any $\delta > 0$, such that $\delta < \varepsilon$; then for any $x \in B(0, \delta)$ we have

$$|x\mathcal{D}(x) - 0| \leq |x| < \delta < \varepsilon \, .$$

Hence $x\mathcal{D}(x)$ is by definition continuous at $x = 0$. Intuitively, by multiplying the Dirichlet function $\mathcal{D}$ with x we have "pacified" its jumping at the origin. The reader is invited to prove that the function $x\mathcal{D}(x)$ is indeed discontinuous at any $x \neq 0$.

The function may not have well-defined one-sided limits, although itself may be well-defined at that point. For example, function $\phi : \mathbb{R} \to \mathbb{R}$, defined as

$$\phi(x) = \begin{cases} \sin(1/x) \, , & x > 0 \\ 0 \, , & x \leq 0 \end{cases}$$

doesn't have a well-defined limit from the right at $x = 0$, although $\phi(0) = 0$. Another similar example, the function $\psi : \mathbb{R} \to \mathbb{R}$, defined as

$$\psi(x) = \begin{cases} \sin(1/x) \, , & x \neq 0 \\ 0 \, , & x = 0 \end{cases}$$

doesn't have a well-defined limit from the right, nor the limit from the left at $x = 0$, although $\psi(0) = 0$.