On convergence, continuity and compactness

Carmine-Emanuele Cella

June 23, 2023

Abstract

This short note collects several definitions and theorems (without proof) about convergence, continuity and compactness. It is intended to be used as a self-contained reference designed to give an overview of the topics from the ground up. The concepts are examined in the context of sequences, single and multi-variable functions and topological spaces.

1 Sets and functions

Definition 1. *A* set *is an unordered collection of distinct elements.*

If an element a belongs to the set A we write $a \in A$, otherwise we write $a \notin A$. Given two sets A and B, their *union* is written $A \cup B$ while their *intersection* is written $A \cap B$. If A is a included in or equal to B, we write $A \subseteq B$.

Definition 2 (Numbers)**.** *We define:*

- natural numbers *the set* $\mathbb{N} = \{1, 2, 3, \ldots, n\}$[1]
- integer numbers *the set* $\mathbb{Z} = \{0, \pm 1, \pm 2, \ldots, \pm n\}$*,*
- rational numbers *the set* $\mathbb{Q} = \{\frac{m}{n} : m, n \in \mathbb{Z}, n \neq 0\}$
- real numbers *the set* $\mathbb{R} = \{\mathbb{Q} \cup \mathbb{I}\}$ *where* $\mathbb{I}$ *is the set of all numbers that cannot be expressed with fractions.*

Numbers can be *ordered*: given two numbers a, b it is possible to decide which of the two is smaller: $a \leq b$ means that a is smaller or equal than b.

Definition 3. *A set* $A \subseteq \mathbb{R}$ *is* bounded above *if there exist a number* $b \in \mathbb{R}$ *such that* $a \leq b$ *for all* $a \in A$*. The number* b *is called* upper bound *for* A*. The* lower bound *is similarly defined.*

Definition 4. *The number* $s \in \mathbb{R}$ *is a* least upper bound *or* supremum *for* $A \subseteq \mathbb{R}$ *if:*

- s *is an upper bound for* A*;*
- *if* b *is an upper bound for* A*, then* $s \leq b$*.*

The *greatest lower bound* or *infimum* is similarly defined.

Axiom 1 (Completness of $\mathbb{R}$)**.** *Every nonempty set of real numbers that is bounded above has a least upper bound.*

Definition 5. *A* function $f : X \to Y$ *is the assignment of an element of the set* Y *(codomain) to each element of the set* X *(domain). The elements of* Y *that are associated by the function are called the* range *of* f*.*

A function $f : X \to Y$ is called *injective (one-to-one)* if elements of X have distinct images in Y. The function f is called *surjective (onto)* if for any $y \in Y$ there exists at least a $x \in X$ such as $y = f(x)$. If a function f is at the same time injective and surjective that it is called *bijective.* If there is a bijection from A onto B then A and B are said to have *equal cardinality*, and we write $A \sim B$.

Definition 6. *A set* S *is:*

- finite *if it empty or for some* $n \in \mathbb{N}$ *we have* $S \sim \{1, \ldots, \}$*.*
- infinite *if it is not finite.*
- denumerable *if* $S \sim \mathbb{N}$*.*
- countable *if it is finite or denumerable.*
- uncountable *if it is not countable.*

Theorem 1. $\mathbb{R}$ *is uncountable.*

An important function that will be used below is the *absolute value.*

[1] From now on we will assume an intuitive understanding of the symbols $0, 1, 2, 3, \ldots$ and of the symbols $\pm$ (plus or minus), $=$ (equal) and $\neq$ (not equal).

Definition 7 (Absolute value)**.** *The* absolute value *of a real number x is the function indicated as $|x|$ and defined as:*

$$|x| = \begin{cases} x & if \quad x \geq 0 \\ -x & if \quad x < 0. \end{cases} \tag{1}$$

2 Sequences

Definition 8. *A* sequence (a_n) *is a function whose domain is* $\mathbb{N}$.

Definition 9 (Convergence)**.** *A sequence (a_n) converges to a real number a if, for every $\epsilon > 0$, there exists an $n \in N$ such that for $n \leq N$ it follows that $|a_n - n| < \epsilon$.*

The number a is called the *limit* of the sequence and it is usually denotated either by $\lim a_n = a$ or by $(a_n) \to a$.[2]

Theorem 2. *Every convergent sequence is bounded.*

Definition 10. *Let (an) be a sequence of real numbers, and let $n_1 < n_2 < n_3 < n_4 < n_5 < \ldots$ be an increasing sequence of natural numbers. Then the sequence $(a_{n_1}, a_{n_2}, a_{n_3}, a_{n_4}, a_{n_5}, \ldots)$ is called a* subsequence *of (a_n) and is denoted by (a_{n_k}), where $k \in N$ indexes the subsequence.*

Theorem 3. *Subsequences of a convergent sequence converge to the same limit as the original sequence.*

Theorem 4 (Bolzano-Weierstrass)**.** *Every bounded sequence contains a convergent subsequence.*

3 Real single-variable functions

In this section, all the functions will be of the type $f : \mathbb{R} \to \mathbb{R}$.

3.1 Limits and continuity

Definition 11. *The function $f(x)$ has* limit *L if, for any $\epsilon > 0$ there exists $\delta > 0$ such that for all x with $0 < |x - a| < \delta$ we have $|f(x) - L| < \epsilon$.*

We then write:

$$L = \lim_{x \to a} f(x). \tag{2}$$

Intuitively, the function $f(x)$ has a limit L at a point a if, for numbers x near a, the value of the function approaches the number L.

Definition 12. *The function $f(x)$ is said to be* continuous *at a if its limit for $x \to a$ is equal to $f(a)$:*

$$\lim_{x \to a} f(x) = f(a). \tag{3}$$

In other words, for any $\epsilon > 0$, there exists $\delta > 0$ such that for all x with $0 < |x - a| < \delta$ we have $|f(x) - f(a)| < \epsilon$.

3.2 Differentiation and integration

Definition 13. *The function $f(x)$ is said to be* differentiable *if the limit:*

$$L = \lim_{x \to a} \frac{f(x) - f(a)}{x - a} \tag{4}$$

exists.

This limit is called the *derivative* of f and is denoted by $f(a)'$ or $\frac{df}{dx}(a)$. Intuitively, the derivative of a function gives the slope of the tangent line to the curve $y = f(x)$.

Let $f(x)$ be a real-valued function whose domain is the closed interval $[a, b]$. The partitions Δt of the interval $[a, b]$ are given by the rectangles whose area is:

$$\Delta t = \frac{b - a}{n} \tag{5}$$

for each positive integer n. As such, $t_0 = a, t_1 = t_0 + \Delta t, \ldots, t_n = t_{n-1} + \Delta t = b$. Let l_k and u_k two points in the range $[t_{k-1}, t_k]$ such that: $f(l_k) \leq f(t)$ and $f(u_k) \geq f(t)$.

[2] The symbol $\to$ can be read as *tends to* or *approaches.*

Definition 14. *The* lower sum $L(f, n)$ *of* $f(x)$ *is given by:*

$$L(f, n) = \sum_{k=1}^{N} f(l_k)\Delta t. \tag{6}$$

Similarly, the upper sum $U(f, n)$ *of* $f(x)$ *is given by:*

$$U(f, n) = \sum_{k=1}^{N} f(u_k)\Delta t. \tag{7}$$

In other words, $L(f, n)$ is the sum of the rectangles below the curve $y = f(x)$, while $U(f, n)$ is the sum of the rectangles above the curve.

Definition 15. *The function* $f(x)$ *on the closed interval* $[a, b]$ *is said to be* integrable *if the following two limits exist and are equal:*

$$\lim_{n\to\infty} L(f, n) = \lim_{n\to\infty} U(f, n). \tag{8}$$

This limit is called *integral* of $f(x)$ and is denoted by:

$$\int_a^b f(x)dx. \tag{9}$$

Intuitively, the integral of $f(x)$ represents the area under the curve $y = f(x)$ above the x-axis.

Theorem 5 (Fundamental theorem of calculus)**.** *Let* $f(x)$ *be a continuous function defined on the closed interval* $[a, b]$ *and let*

$$F(x) = \int_a^x f(t)dt \tag{10}$$

be its integral. The function $F(x)$ *is differentiable and:*

$$\frac{dF(x)}{dx} = \frac{d\int_a^x f(t)}{dx} = f(x). \tag{11}$$

If $G(x)$ *is a differentiable function defined on the closed interval* $[a, b]$ *and its derivative is* $f(x)$*, then:*

$$\int_a^b f(x)dx = G(b) - G(a). \tag{12}$$

3.3 Convergence

Definition 16. *A sequence of functions* $f_n : [a, b] \to \mathbb{R}$ converges pointwise *to a function* $f(x) : [a, b] \to \mathbb{R}$ *if for all* $\alpha \in [a, b]$ *and given* $\epsilon > 0$ *there is a positive integer* N *such that for all* $n \geq N$*, we have* $|f(\alpha) - f_n(\alpha)| < \epsilon$*.*

Intuitively, a sequence of functions $f_n(x)$ will converge pointwise to a function $f(x)$ if, given any a, eventually (for huge n) the numbers $f_n(a)$ become close to $f(a)$. Note the pointwise limit of continuous functions need not to be continuous.

Definition 17. *A sequence of functions* $f_n : [a, b] \to \mathbb{R}$ *will* converge uniformly *to a function* $f : [a, b] \to \mathbb{R}$ *if given any* $\epsilon > 0$*, there is a positive integer* N *such that for all* $n \geq N$*, we have* $|f(x) - f_n(x)| < \epsilon$ *for all points* x*.*

Intuitively, if there is a *tube* of width 2ϵ centered around $f(x)$, all the functions $f_n(x)$ will eventually fit inside it.

Theorem 6 (Uniform limit)**.** *Let* $f_n : [a, b] \to \mathbb{R}$ *be a sequence of continuous functions converging uniformly to a function* $f(x)$*. Then* $f(x)$ *will be continuous.*

4 Real vector-valued functions

Definition 18. *A function* $f : \mathbb{R}^n \to \mathbb{R}^m$ *is called* vector-valued *since for any vector* $x \in \mathbb{R}^n$*, the value of* $f(x)$ *is a vector in* $\mathbb{R}^m$*.*

In this section, all the functions will be of the type $f : \mathbb{R}^n \to \mathbb{R}^m$.

4.1 Limits and continuity

Definition 19. *The function* $f : \mathbb{R}^n \to \mathbb{R}^m$ *has limit* $L = (L_1, \ldots, L_m) \in \mathbb{R}^m$ *at the point* $a = (a_1, \cdots, a_n) \in \mathbb{R}^n$ *if given any* $\epsilon > 0$*, there is some* $\delta > 0$ *such that for all* $x \in \mathbb{R}^n$*, if* $0 < |x - a| < \delta$*, we have* $|f(x) - L| < \epsilon$*.*

This limit is denoted by $\lim_{x\to a} f(x) = L$.

Definition 20. *The function* $f : \mathbb{R}^n \to \mathbb{R}^m$ *is said to be* continuous *at a point* $a \in \mathbb{R}^n$ *if* $\lim_{x\to a} f(x) = f(a)$*.*

Both the definitions of limit and continuity rely on the existence of a distance (norm).

4.2 Derivative and Jacobians

Definition 21. *A function* $f : \mathbb{R}^n \to \mathbb{R}^m$ *is said to be* differentiable *at a point* $a \in \mathbb{R}^n$ *if there is an* $m \times n$ *maxtrix* $A : \mathbb{R}^n \to \mathbb{R}^m$ *such that:*

$$\lim_{x \to a} \frac{|f(x) - f(a) - A \cdot (x - a)}{|x - a|} = 0. \tag{13}$$

If this limit exists, the matrix A is denoted by $Df(a)$ and is called the *Jacobian.* This definition agrees with the usual definition of derivative for a function $f : \mathbb{R} \to \mathbb{R}$.

Theorem 7. *Let the function* $f : \mathbb{R}^n \to \mathbb{R}^m$ *be made by* m *differentiable functions* $f_1(x_1, \ldots, x_n), \ldots, f_m(x_1, \ldots, x_n)$ *so that:*

$$f(x_1, \ldots, x_n) = \begin{bmatrix} f_1(x_1, \ldots, x_n) \\ \vdots \\ f_m(x_1, \ldots, x_n) \end{bmatrix}. \tag{14}$$

Then f *is differentiable and the Jacobian is:*

$$Df(x) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix} \tag{15}$$

where $\frac{\partial f_m}{\partial x_n}$ is the partial derivative of f along dimension m.

Theorem 8 (Chain rule)**.** *Let* $f : \mathbb{R}^n \to \mathbb{R}^m$ *and* $g : \mathbb{R}^m \to \mathbb{R}^l$ *be differentiable functions. Then the composition function*

$$g \circ f : \mathbb{R}^n \to \mathbb{R}^l \tag{16}$$

is also differentiable with derivative given by: if $f(a) = b$*, then*

$$D(g \circ f)(a) = D(g)(b) \cdot D(f)(a). \tag{17}$$

In other words, to find the derivative of the composition $g \circ f$, we need to multiply the Jacobian matrix for g times the Jacobian matrix for f.

The Jacobian $Df(a)$ can be thought of as a linear map from $\mathbb{R}^n \to \mathbb{R}^m$ and $f(a)$ as a translation. Thus the vector $y = f(x)$ can be approximated by:

$$y \approx f(a) + Df(a) \cdot (x - a). \tag{18}$$

4.3 Inverse functions

In eqution 18 we have seen that vector-valued functions can be approximated a matrix, the Jacobian. In general, the connection between the properties of matrices and the properties of vector-valued functions is very important. For example, if the Jacobian is invertible, then the original vector-valued function is also have an inverse, at least locally.

Definition 22 (Open neighborhood)**.** *By an* open neighborhood U *of a point* $a \in \mathbb{R}^n$*, we mean that given any* $a \in U$*, there is a* $\epsilon > 0$ *such that:*

$$\{x : |x - a| < \epsilon\} \in U. \tag{19}$$

Theorem 9 (Inverse function)**.** *Let* $f : \mathbb{R}^n \to \mathbb{R}^m$ *be a vector-valued continuously differentiable function, for which* $\det Df(a) \neq 0$*, at some point* $a \in \mathbb{R}^n$*. Then there is an open neighborhood* U *of* $a \in \mathbb{R}^n$ *and and an open neighborhood* V *of* $f(a) \in \mathbb{R}^n$ *such that* $f : U \to V$ *is one to one, onto and has a differentiable inverse* $g : V \to U$*.*

In other words, $g \circ f : U \to U$ is the identity and $f \circ g : V \to V$ is the identity.

For the case of $f : \mathbb{R} \to \mathbb{R}$, it is easy to have an intuitive intepretation of the *Inverse function* theorem. A derivative can be thought of as a the slope of the tangent line to a f. If the slope is zero, then the tangent is parallel to the x-axis (in other words, it is horizontal). Such a line is a constant and can be the derivative of an infinite number of functions, so f is not invertible.

5 Topological interpretation

Many of the definitions and results presented so far, can be reinterpreted in a more general way with the language of topology. The advantage of this interpretation is that a specific problem can be harder to solve than some abstract generalisation of it.

5.1 Topological spaces

Definition 23 (Topological space)**.** *Let X be a set of points. A collection of subsets $U = \{U_\alpha\}$ forms a* topology *on X if:*

1. *Any arbitrary union of the U_α is another set in the collection U.*
2. *The intersection of any finite number of sets U_α in the collection U is another set in U.*
3. *Both the empty set $\emptyset$ and the whole space X must be in U.*

The pair (X, U) is called a topological space.

Definition 24 (Open and closed sets)**.** *The sets U_α in the collection U are called* open sets*. A set C is* closed *if its complement $X - C$ is open.*

Definition 25. *Let A be a subset of a topological space X. The* induced topology *on A is described by letting the open sets on A be all the sets of the form $U \cap A$, where U is an open set in X.*

Definition 26. *A collection $\Sigma = \{U_\alpha\}$ of open sets is called an* open cover *of a subset A if A is contained in the union of the U_α.*

Definition 27 (Compactness)**.** *The subset A of a topological space X is* compact *if given any open cover of A, there is a finite subcover.*

This means that if $\Sigma = \{U_\alpha\}$ is an open cover of A in X, then A is compact if it is included in a finite union of n elements of U_α:

$$A \subset (U_1 \cup U_2 \cup \ldots \cup U_n). \tag{20}$$

Definition 28. *A topological space X is* Hausdorff *if given any two points $x_1, x_2 \in X$, there are two open sets U_1 and U_2, with $x_1 \in U_1$ and $x_2 \in U_2$, whose intersection is empty.*

In other terms, X is Hausdorff if points can be separated from each other by disjoint open sets.

Definition 29 (Continuity)**.** *A function $f : X \to Y$ is* continuous*, where X and Y are two topological spaces, if given any open set U in Y, then the inverse image $f^{-1}(U)$ in X must be open.*

Definition 30. *A topological space X is* connected *if it is not possible to find two open sets U and V in X with $X = U \cup V$ and $U \cap V = \emptyset$.*

5.1.1 Bases for a topology

Topological spaces can have a *basis*; the usual interpretation of this word refers to a list of vectors in a vector space that generates uniquely the entire vector space. In a topology, a basis is a collection of open sets that generate the entire topology.

Definition 31. *Let X be a topological space. A collection of open sets forms a* basis *for the topology if every open set in X is the (possibly infinite) union of sets from the collection.*

Definition 32. *A topological space is* second countable *if it has a basis with a countable number of elements.*

5.2 Metric spaces

Any set that has an associatoed notion of distance (*metric*) automatically has a topology.

Definition 33. *A* metric *on a set X is a function:*

$$\rho : X \times X \to \mathbb{R} \tag{21}$$

such that for all points $x, y, z \in X$ we have:

- *$\rho(x, y) \geq 0$ and $\rho(x, y) = 0$ if and only if $x = y$.*
- *$\rho(x, y) = \rho(y, x)$.*
- *(triangle inequality) $\rho(x, z) \leq \rho(x, y) + \rho(y, z)$.*

Definition 34 (Metric space)**.** *The set X with its metric ρ is called a* metric space *and is denotated by (X, ρ).*

Definition 35. *A set U in X is open if for all points $a \in U$, there is some real number $\epsilon > 0$ such that*

$$\{x : |x - a| < \epsilon\} \tag{22}$$

is contained in U.

5.3 Standard topology on $\mathbb{R}^n$

The set of real numbers $\mathbb{R}$ has a natural (Euclidian) notion of distance:

$$|a-b| = \sqrt{(a_1-b_1)^2 + \ldots + (a_n-b_b)^2)} \tag{23}$$

where $a, b \in \mathbb{R}^n$.

Using this distance, it is possible to create equivalent definitions of open/closed sets and continuity for $\mathbb{R}^n$, to create the so called *standard topology on* $\mathbb{R}^n$.

Definition 36. *A set U in $\mathbb{R}^n$ will be* open *if given any $a \in \mathbb{R}^n$, there is a real number $\epsilon > 0$ such that:*

$$\{x : |x-a| < \epsilon\} \tag{24}$$

is contained in U.

From this definition it follows that sets of the form $(a, b) = \{x : a < x < b\}$ are open, while sets of the form $[a, b] = \{x : a \leq x \leq b\}$ are closed.

Definition 37 (Standard topology)**.** *The collection of all open sets $(a, b) = \{x : a < x < b\}$ with $a, b \in \mathbb{R}^n$ is called the* standard topology *on $\mathbb{R}^n$.*

Theorem 10. *The standard topology on $\mathbb{R}^n$ is Hausdorff.*

Theorem 11 (Continuity)**.** *Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be a function. For all $a \in \mathbb{R}^n$:*

$$\lim_{x \to a} f(x) = f(a) \tag{25}$$

if and only if for any open set $U \in \mathbb{R}^m$, the inverse image $f^{-1}(U)$ is open in $\mathbb{R}^n$.

Definition 38. *A subset A is* bounded *in $\mathbb{R}^n$ if there is some fixed real number r such that for all $x \in A$,*

$$|x| < r. \tag{26}$$

Examples The interval (a, b) is bounded but not closed; the interval $[a, \infty)$ is closed but not bounded.

For the standard topology on $\mathbb{R}^n$, compactness is equivalent to the intuitive idea of being both closed and bounded. This equivalence is the goal of the following theorem.

Theorem 12 (Heine-Borel)**.** *A subset A of $\mathbb{R}^n$ is* compact *if and only if it is closed and bounded.*

Theorem 13. *On the real line $\mathbb{R}$, a closed interval $[a, b]$ is compact.*

Theorem 14. *A subset A in $\mathbb{R}^n$ is compact if every infinite sequence (x_n) of points in A has a subsequence converging to a point in A.*

In other words, if (x_n) is a collection of points in A, there must be a point $p \in A$ and a subsequence x_{n_k} with $\lim_{k\to\infty} x_{n_k} = p$.

References

[1] T. Garrity, *All the math you missed (but need to know for graduate school)*, Cambridge, 2002.

[2] S. Abbott, *Understanding analysis*, Springer, second edition, 2015.

[3] C. Pugh, *Real mathematical analysis*, Springer, second edition, 2015.