About Limits, Continuity, Cauchy Sequences and Completeness in Metric Spaces

Herman Jaramillo

June 16, 2015

1 Introduction

A few lines only about this important concepts. I might change this introduction in the future. The context is that of metric spaces.

2 Limit

2 .1 Motivation

The concept of limit arrives from the need to fit data with better measurements. In the field of mathematics the irrational numbers were seen as limits of rational sequences. Fractions with infinite number of digit reprentations in the decimal expansion could motivate the definition of limit cleary.

For example, let us assume that we want to evaluate numerically 1/3. We use elementary division and find that

$$\frac{1}{3} = 0.33333\cdots$$

In the real world (of computers) we need to stop the division somewhere. If we want to improve precision we should admit an error or tolerance level that we can live with. The error could be noted as ϵ for the lowercase Greek letter e. This error can be measured with the absolute value, and we want to say that

$$\left|\frac{1}{3} - 0.333\right| < \epsilon.$$

The problem here is that we are trying to find the error between something that we know and something that we do not know. We do not know what exactly is 1/3 (in the world of decimals). But we can define a sequence that approaches this number. This sequence is the sequence (x_i) where

$$x_n = \sum_{i=1}^{n} 3 \times 10^{-i}$$

This sequence produces the numbers $0.3, 0.33, 0.333 \cdots$. One thing we know well is that two consecutive numbers of this sequence get very small as n goes large, that is

$$|x_{n-1} - x_n| = 3 \times 10^{-n}$$

Even more, given any two numbers $n > m$ which most of the time we want to be large, we have that

$$|x_n - x_m| = 3 \sum_{i=m}^{n} 10^{-i} = 3 \times 10^{-m} \sum_{i=1}^{n-m} 10^{-i}.$$

but we know that $\sum_{i=0}^{n-m} 10^{-i} < 2$. So we can say that

$$|x_n - x_m| = 6 \times 10^{-m}$$

This means that we have the error ϵ under control by just finding a big number N and let $m > N$ and $n > N$. This motivates the definition of a Cauchy sequence

Definition 2 .2 *Given a metric space (X, d), and a sequence (x_n) with $x_n \in X$, we say that the sequence is a* **Cauchy sequence** *if given $\epsilon > 0$, there is an $N \in \mathbb{N}$ such that for any $m > N$, and $n > N$, $d(x_n, x_m) < \epsilon$.*

This is just a definition and it says nothing about convergence of existing of a limit point. It is enough to check that the condition $d(x_{n+1} - x_n) < \epsilon$ is satisfied, since we can write (assuming $n < m$)

$$d(x_n, x_m) \leq d(x_n, x_{n+1}) + d(x_{n+1}, x_{n+2}) + \cdots + d(x_{m-1}, x_m);$$

by repeating the triangular inequality $m-n$ times. We can redefine the error as $\epsilon/(m-n)$, and so obtain $d(x_n, x_m) < \epsilon$ as desired. In general is much easier to check the inequality of a member of the sequence with its immediate neighbour.

We wish to have a point x such that the sequence (x_n) converges. That is, we want to say

Definition 2 .3 *Given a metric space* (X, d)*, we say that* x *is the* **limit of a sequence** (x_n) , with $x_n \in X$, if given any $\epsilon > 0$, there is a number $N \in \mathbb{N}$, such that for any $n > N$, $d(x_n, x) < \epsilon$. We write $\lim_{n\to\infty} x_n = x$.

Now, we are saying that there is a point x which is a limit point and that we would like to find. Let us provide an example. Choose X to be the interval $X = (0, 1)$ in the real numbers, and $x_n = 1/n$. Then we know that $\lim_{n\to\infty} x_n = 0$ (choose $N > 1/\epsilon$). However $0 \notin X$. So if our space is X we do not know the number 0. We say that the function does not converges in X. However if we add 0 to X then we have now $Y = [0, 1) = X \cup \{0\}$, and the function converges in Y. Adding points so that the limit exists is called completion.

The history of limits is quite interesting. The book Practical Anaysis in One Variable and particular the chapter 9 Sequences and Limits [1] from Donald Estep offers a good history on the matters discussed here.

We know that the differential and integral Calculus are based on top of the concept of limits. Newton understood the concept but could never get it in a formal presentation. More than one hundred years passed until Cauchy [2] came with the definition above.

The decimal representation of fractional numbers is attributed to Simon Stevin [3] We well know now that $1/2 = 0.5$ and fractions have a finite number of decimal digits, or an infinite periodic number of decimal digits. We could think about the rational numbers in these two different forms and say that in a way the rational numbers, such as $1/3$ in our example, are a completion of the rational numbers. That is, the periodic representations are also part of the rational numbers. However the history is more interesting that this. J. J. O'Connor and E.F. Robertson [4] offer a good introduction to the history of the real numbers. We extract a few interesting items related to the completion of the real line from the rational numbers.

---

[1] http://www.stat.colostate.edu/ estep/education/M618/notesonreals.pdf

[2] Note that Donal Estep gives credit for the $\epsilon - \delta$ definition of limit to Weirstrass, as well as for the introduction of the notation "lim".

[3] https://en.wikipedia.org/?title=Simon_Stevin

[4] http://www-history.mcs.st-and.ac.uk/HistTopics/Real_numbers_2.html

- Hankel (student of Weirstrass) in 1867 addressed the question whether there were other "number system" which had essentially the same rules as the real numbers.

- Two years after Hankel's announcement Méray published an article where he considered that Cauchy sequences of rational numbers converge to what he called "fictitions limit" (which we know today as irrational numbers).

- Heine, three years after Méray published a similar result. He defined sequences of rational numbers to be equivalent if they converge to the same (limit) real number. This created a partition of the real numbers in equivalence classes. Each equivalence class [5] is a real number and they complete the real line

- Cantor, in 1872 published the same idea of Heine.

2 .3.1 Example:

The use of Cauchy sequences is important because a Cauchy sequence contains a collection of elements that we should know, regardless their limit (which might be unknown to us). So when we compare elements, we are comparing objects that are known to us.

We provide three ways to compute $\sqrt{2}$ using rational sequences approximations. The first form illustrates directly how the Cauchy sequence appears in a simple algorithm. The bisection algorithm is easy to understand, easy to implement, but more costly than the Newton-Raphson method explained in the second example. The second example provides a recursive formula which is convenient from the computational point of view but it is not convenient from the theoretical (proof building) point of view. Still we proof that the sequence generated is Cauchy as a good exercise of a more complicated problems. The third example provides an analytic expansion which provides an easy path to proof that the sequence is a Cauchy sequence.

- The **bisection algorithm:** The idea for the bisection algorithm is easily explained with this example. Think of the function $f(x) := x^2 - 2$. We want to solve the equation $f(x) = 0$ for some $x \in [a, b]$ where $[a, b]$ is some interval such that $f(a) < 0$, and $f(b) > 0$. Since the

[5] If it the limit is rational, we can see it as a constant ($x_n = x$), with $x \in \mathbb{Q}$.

function is continuous there has to be a point in the interval such that $f(x) = 0$. We can then think about the interval $[1, 2]$. Since $f(1) < 0$ and $f(2) > 0$. The algorithm is as follows:

– Pick $x_0 = a$, $x_1 = b$, $x_2 = (x_0 + x_1)/2$ (the midpoint of the interval).

– Choose x_3 as follows:

$$x_3 = \begin{cases} \frac{x_0+x_2}{2} & \text{if } f(x_2) > 0 \\ \frac{x_1+x_2}{2} & \text{if } f(x_2) < 0 \end{cases}$$

– In general, for $n \geq 3$ choose

$$x_n = \begin{cases} \frac{x_{n-3}+x_{n-1}}{2} & \text{if } f(x_{n-1}) > 0 \\ \frac{x_{n-2}+x_{n-1}}{2} & \text{if } f(x_{n-1}) < 0 \end{cases}$$

A sequence of points created with this algorithm is

$$(1, 2, 1.5, 1.25, 1.375, 1.406, 1.4375 \cdots).$$

It takes up to 10 terms of the sequence to get closer than 1.41406265 to $\sqrt{2}$, after $x_2 = 1.5$. It takes only 3 iterations of the Newton's method to get there (this is shown in Figure 1.

In the first step, after the first two points where created the interval is reduce to 1/2 of the original interval. Then after every step it is reduced each time by 1/2. So after n points are selected we can say that $|x_{n+1} - x_n| \leq 2^{-n}$. Clearly this is saying that the the series is Cauchy, So if the error is $\epsilon = 2^{-n}$, we can find choose any $N > |\log_2(\epsilon)|$, and this will guarantee that our error ϵ is a good bound for the level of accuracy required.

- The **Newton's method:** This method, also called the Newton-Raphson method [6] [7] is designed to compute roots of a function based on the iteration

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \tag{2 .1}$$

[6] https://en.wikipedia.org/wiki/Newton%27s_method

[7] An easy way to interprete this formula is to think that $f(x) = 0$, (the root we want to find) and, and $x_{n+1} = x$ (we stop close enough to the root) write the equation as

$$x_{n+1} = x_n + \frac{f(x) - f(x_n)}{f'(x_n)}$$

provided the derivative exists and is non-zero in the search interval. We know $\sqrt{2}$ is irrational. We can use Newton's method to solve the equation $f(x) := x^2 - 2 = 0$. That is,

$$x_{n+1} = x_n - \frac{x_n^2 - 2}{2\,x_n} = \frac{x_n}{2} + \frac{1}{x_n} \tag{2 .2}$$

which with the initial guess of $x_0 = 1$, produces the sequence $(1, 3/2, 17/12, \cdots)$. Or in decimals $(1, 1.5, 1.4166\ldots, \cdots)$.

Actually we can rewrite equation 2 .2 in two different ways and interpret it.

$$x_{n+1} = \frac{x_n^2 + 2}{2x_n} = \frac{1}{2}\left(x_n + \frac{2}{x_n}\right).$$

The first interpretation, for the expression in the middle, is that for $x_n \approx \sqrt{2}$, we have in the numerator twice the square of the denominator. So to recover the root, we should divide by 2 and then by x_n to eliminate the power of two in the numerator. This is precisely what is written in the equation. The second interpretation is that each term is the average (the 1/2 in the last expression) of the previous iteration and the inverse of the first iteration weighted by 2, which is the squared value. So if the first term x_n overestimates the value, the second balance this value by having it in the denominator, then the average is taken. In general change 2 by $a > 0$ and this works the same for any positive root. This method is called the Babylonian or Heron's method [8].

We would like to check that the sequence is Cauchy. Note that since

---

from which

$$f'(x_n) = \frac{f(x) - f(x_n)}{(x - x_n)}$$

which is the approximation for the derivative at point x_n. Another way (and the most commonly interpretation) is to write this as

$$f(x) = f(x_n) + f'(x_n)(x - x_n).$$

which geometrically is the segment tangent at f at $x = x_n$. This is illustrated in Figure 1

[8] https://en.wikipedia.org/wiki/Methods_of_computing_square_roots

the initial value is $x_0 = 1$, then from

$$x_{n+1} = \frac{1}{x_n} + \frac{x_n}{2},$$

all the values of the sequence are bounded below by 1. That is $x_n \geq 1$, $\forall n \in \mathbb{N}$. We search for an upper bound. For this we show that the sequence (x_n) is decreasing, after the first evaluation and so we can start at the first value $x_1 = 3/2$.

$$x_{n+1}^2 = \frac{1}{4}\left(x_n + \frac{2}{x_n}\right)^2 = \frac{1}{4}\left(x_n - \frac{2}{x_n}\right)^2 + 2 \geq 2, \qquad (2\ .3)$$

that is, for $n = 0, 1, 2, \cdots$ we have $2 - x_{n+1}^2 < 0$, and then

$$x_{n+1} - x_n = \frac{2 - x_n^2}{2x_n} \leq 0 \quad n \geq 1. \qquad (2\ .4)$$

Then, after passing the first element $x_0 = 1$, (x_n) is decreasing and we can locate all (x_n) in the interval $[1, 3/2]$

If after some n, $x_n = 2$, then any new term on the iteration will be 2 and that is the limit. Let us then assume that we do not get in any finite step $x_n = 2$, so the sequence is strictly decreasing and we have an infinite number of values in the finite interval $[1, 3/2]$ which by the Bolzano-Weirstrass theorem has a limit point.

We further show this without recurring to the Bolzano-Weirstrass theorem. From equation 2 .3

$$x_{n+1}^2 - 2 = \frac{(x_n^2 - 2)^2}{4x_n^2} \quad \text{hence} \quad 0 \leq\ x_{n+1}^2 - 2 \leq \frac{1}{4}(x_n^2 - 2)^2.$$

In the last equation, we replace

$$x_n^2 - 2 \leq \frac{1}{4}(x_{n-1}^2 - 2)^2$$

and so

$$(x_n^2 - 2)^2 \leq \left(\frac{1}{4}\right)^2 (x_{n-1}^2 - 2)^4$$

and keep doing recursively until

$$0 \leq x_{n+1}^2 - 2 \leq \left(\frac{1}{4}\right)^{n+2} (x_0^2 - 2)^{2(n+2)} = \left(\frac{1}{4}\right)^{n+2}$$

since $x_0 = 1$. Now from equation 2 .4,

$$|x_{n+1} - x_n| = \left|\frac{2 - x_n^2}{2x_n}\right| \leq |2 - x_n^2| \leq \left(\frac{1}{4}\right)^{n+2}$$

From which (x_n) is a Cauchy sequence. Figure 1 illustrates the Newton's method for this example. Figure 2 shows the points selected my the Newton's method.

- **The Taylor Series approach:** A way to get a rational sequence that converges to $\sqrt{2}$, and such that we can check the Cauchy sequence criteria easily is through a Taylor series expansion. That is, let us assume $f(x) = \sqrt{1+x}$, and let us expand this series at $x = 1$. Not that $f(1) = \sqrt{2}$. That is, from the binomial expansion

$$(1+x)^{1/2} = 1 + \frac{x}{2} - \frac{x^2}{2.4} + \frac{1.3x^3}{2.4.6} \cdots (-1)^{k+1} \frac{(2k-3)!!}{(2k)!!} + \cdots$$

where the double factorial $n!!$ is defined as

$$n!! = n(n-2)(n-4)\cdots 1$$

where the "1" at the end is either 1! for n odd or 0!, for n even. In the particular case of $x = 1$, we have

$$\sqrt{2} = 1 + \frac{1}{2} + \sum_{k=2}^{\infty} (-1)^{k+1} \frac{(2k-3)!!}{(2k)!!}.$$

We show that the sequence

$$x_n = \sum_{k=0}^{n} (-1)^{k+1} \frac{(2k-3)!!}{(2k)!!}.$$

is Cauchy. Here

$$|x_{n+1} - x_n| = \left|\frac{[2(n+1)-3)]!!}{2(n+1)!!}\right| = \left|\frac{(2n-1)!!}{(2n+2)!!}\right| \leq \left|\frac{(2n-1)!!}{(2n+1)!!}\right|$$

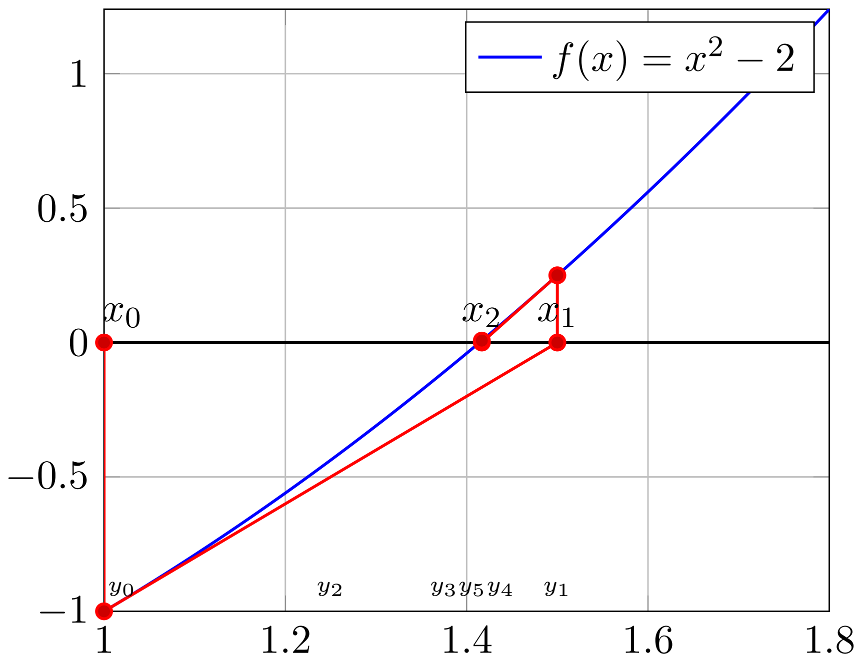

Figure 1: The Newton-Raphson method on $y = x^2 - 2$. The method converges so fast that after two iterations we could not distinguish more roots at the resolution of the plot. Here is the description of the path (red) of search: Start at $x_0 = 1$, find $f(x_0) = -1$, trace a tangent to f until hits the x-axis at $x_1 = 1.5$. Find $f(x_1) = 0.25$, trace a tangent from x_1 to the curve f until hitting the x axis at $x_2 = 1.4167$, then find $f(x_2) = 0.00703889$. This is so small that after this points in the plot would look as one. It is clear that the Newton-Raphson method converges faster than the bisection method. The bisection method choose the middle of the current interval each time regardless the shape of the function f. The Newton-Raphson method, uses tangents along the curve for f to get faster to the solution. It is a non linear sequence of linear events. In the bottom of the figure the points $y_0, y_1, \cdots, y_5$ show the x-coordinates of the bisection algorithm.

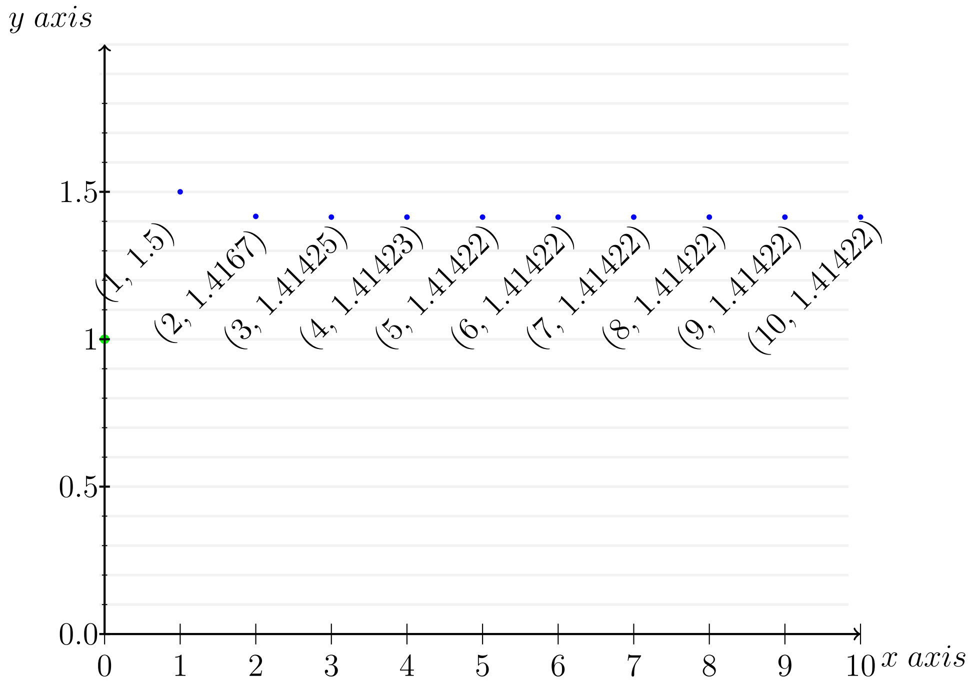

Figure 2: Computation of $\sqrt{2}$. After the 4^{th} iteration the function converges within 5 decimal places.

That is,

$$|x_{n+1} - x_n| \leq \frac{1}{2n+1} \to 0 \quad \text{when} \quad n \to \infty.$$

> We do not compare how fast the Taylor series (binomial expansion) approximation is compared to Newton-Raphson. We leave this to the reader.

The ideas exposed above, are part of the history of the completion of the real line. This should motivate the following definition

Definition 2 .4 *A metric space* (X, d) *is* **complete** *if any Cauchy sequence* (x_n)*, with* $x_n \in X$*, converges in* X*. That is if* $\lim_{n\to\infty} x_n = x$*, with* $x \in X$*.*

The concept of complete spaces is very important in analysis and topology. In fact the following definition introduces Banach Spaces [9]

[9] https://en.wikipedia.org/?title=Banach_space

Definition 2 .5 *A* **Banach Space** *is a complete normed vector space.*

We show now that if a space is complete then convergence of a sequence is equivalence to convergence of a Cauchy sequence.

Definition 2 .6 *Given a function $f : X \to Y$, where (X, d_X), and (Y, d_Y) are metric spaces, we say that*

$$\lim_{x \to x_0} f(x) = y,$$

if given $\epsilon > 0$, there exists a $\delta > 0$, such that if $d_X(x, x_0) < \epsilon$, then $d_Y(f(x), y) < \delta$. Here y is the **limit of the function** *f as x approaches x_0.*

We now define boundness

Definition 2 .7 *A set $A \subset X$ is said* **bounded** *if the set of all distances*

$$D = \{d(x, y) : x, y \in X\}$$

is bounded in the real numbers. The sup$(D) = \delta(A)$ is known as the **diameter** *of the set. A sequence that is bounded as a set is defined as a* **bounded sequence** *.*

We show the boundedness, limit Theorem. That is,

Theorem 2 .8 *Let (X, d) be a metric space then:*

- *A convergent sequence in X is bounded, and its limit is unique.*
- *If $x_n \to x$ and $y_n \to y$, then $d(x_n, y_n) \to d(x, y)$.*

Proof: Let (X, d) be a metric space.

- – Let us assume that a sequence (x_n) is converging to a point x. That is, there is an $N \in \mathbb{N}$ such that for each $n > N, d(x_n, x) < \epsilon$. Hence $\forall n > N$, we have that $x_n \in B(x, \epsilon)$. We want to squeeze $x_1, x_2, \cdots, x_{N-1}$ in that ball, or make it larger, but finite, if necessary. That is, pick

  $$r = \max\{\max_{i=1}^{N-1}\{|x - x_i|\}\ ,\ \epsilon\}$$

  then, by construction, the ball $x_n \in B(x, r)$ for $n = 1, 2, \cdots, \infty$, so the sequence (x_n) is bounded.

- Let us now show that if the limit exists, is unique. By contradiction let us assume that there are two limits α and β. Such that $\alpha \neq \beta$. Then from the definition of limit there are numbers N_1 and N_2 such that

  $$d(x_n, \alpha) < \frac{\epsilon}{2} \qquad n > N_1$$
  $$d(x_n, \alpha) < \frac{\epsilon}{2} \qquad n > N_2$$

  So for any $n > \max\{N_1, N_2\}$ we should have that (using the triangular inequality)

  $$d(\alpha, \beta) \leq d(x_n, \alpha) + d(x_n, \beta) < \epsilon$$

  and since ϵ is as small as we want, $d(\alpha, \beta) = 0$ and so $\alpha = \beta$. That is the limit is unique.

- On the second item. If $x_n \to x$ there is a number $N_1 \in \mathbb{N}$ such that for $n > N_1$, $d(x, x_n) < \epsilon/2$, likewise there is a number $N_2 \in \mathbb{N}$ such that for any $n > N_2$, $d(y, y_n) < \epsilon/2$. Then from the triangular inequality and for any $n > \max\{N_1, N_2\}$

  $$d(x, y) \leq d(x, x_n) + d(x_n, y)$$
  $$d(x_n, y) \leq d(x_n, y_n) + d(y_n, y)$$

  so

  $$d(x, y) \leq d(x, x_n) + d(x_n, y_n) + d(y_n, y).$$

  Then

  $$d(x, y) - d(x_n, y_n) \leq d(x, x_n) + d(y_n, y).$$

  If in the previous inequality we swap x_n with x and y_n with y and since the metric is symmetric we find:

  $$d(x_n, y_n) - d(x, y) \leq d(x, x_n) + d(y_n, y).$$

  Hence

  $$|d(x_n, y_n) - d(x, y)| \leq d(x, x_n) + d(y_n, y) < \epsilon$$

  Then we say that $d(x_n, y_n) \to d(x, y)$.