Point Set Topology

(25 January 2018, rev 26 December 2018)

Jim Stevenson

A non-mathematician friend asked me what was topology, in particular point set topology and algebraic topology. Here is my attempt at an explanation. Perhaps a good high-level summary of the idea is from the excellent 1972 book on the history of mathematics by Morris Kline ([1] pp.1158-9):

> A number of developments of the nineteenth century crystallized in a new branch of geometry, now called topology but long known as analysis situs. To put it loosely for the moment, topology is concerned with those properties of geometric figures that remain invariant when the figures are bent, stretched, shrunk, or deformed in any way that does not create new points or fuse existing points. The transformation presupposes, in other words, that there is a one-to-one correspondence between the points of the original figure and the points of the transformed figure, and that the transformation carries nearby points into nearby points. This latter property is called continuity, and the requirement is that the transformation and its inverse both be continuous. Such a transformation is called a homeomorphism. Topology is often loosely described as rubber-sheet geometry, because if the figures were made of rubber, it would be possible to deform many figures into homeomorphic figures. Thus a rubber band can be deformed into and is topologically the same as a circle or a square, but it is not topologically the same as a figure eight, because this would require the fusion of two points of the band. …
>
> Topology, as it is understood in this century, breaks down into two somewhat separate divisions: point set topology, which is concerned with geometrical figures regarded as collections of points with the entire collection often regarded as a space; and combinatorial or algebraic topology, which treats geometrical figures as aggregates of smaller building blocks, just as a wall is a collection of bricks.[1] Of course notions of point set topology are used in combinatorial topology, especially for very general geometric structures.

Before discussing algebraic topology I thought it prudent to begin with "general topology" (aka "point set topology"). I will try to lead into the subject via a historical path.

Historical Background

On another occasion I alluded to the crisis in mathematics caused by the advent of Fourier Series in the beginning of the 19th century and how it led to a deeper understanding of mathematics that included the origin of point set topology (see references [2] [3] [4]). But in fact the problems go back further to the very origin of calculus in the 17th century where the idea of infinite processes, including infinite sums, infinitesimals, and instantaneous rates of change, were challenged from the outset.

Zeno's Paradoxes

This unease actually traced back even further to practically the dawn of real mathematics with the Paradoxes of the Greek philosopher Zeno of Elea (c.450 BC):

[1] JOS: This is a bit obscure. Perhaps a better distinction is that point set topology is generally concerned with local behavior, while algebraic topology addresses global properties. And algebraic topology does this by attaching numbers and algebraic structures to the underlying spaces in such a way as to characterize specific properties, such as whether the spaces have holes, and how many holes.

Dichotomy Paradox

That which is in locomotion must arrive at the half-way stage before it arrives at the goal.

—as recounted by Aristotle, Physics VI:9, 239b10

If we say the distance to the goal is D, then the paradox is interpreted as saying first one has to travel half the distance (D/2), and then half of the remaining distance (that is, half of D/2 or D/4), and so on, ad infinitum. This is interpreted as forming the infinite sum

$$\frac{D}{2}+\frac{D}{4}+\frac{D}{8}+\frac{D}{16}+\ldots$$

and "everyone knows" adding in infinite number of positive quantities will increase without bound and so the infinite sum cannot possibly be the distance D. This obviously contradicts the physical fact that one can indeed cover the distance D and reach one's goal.

With one known exception this ancient Greek avoidance of infinite processes persisted for almost 2000 years and delayed the further advance of mathematics in this area until the arrival of the Middle Ages and philosophers like Nicholas of Cusa (1401–1464) who identified the infinite with God. Since God existed, so did infinity, and so off they went, ignoring the Greek warnings and developing the ideas of calculus with only the vaguest of intuitive ideas about infinite processes (in accordance with their religious beliefs).

The one Greek exception was Archimedes (c.287-212 BC), who used the idea of an infinite sum of inscribed triangles to yield the area of a circle (which we will consider below p.5). However, he avoided the direct notion of an actual infinite sum by employing the Eudoxus theory of Exhaustion. Nevertheless, his amazingly prescient insights remained ignored until Kepler (1571-1630) in the 17th century employed them almost 2000 years later in his discovery of the equal areas law.

Limits – Infinite Series

All of these infinite processes were eventually "handled" by the idea of "limits," which was codified in the 18th century by mathematicians such as Augustin-Louis Cauchy. To take a canonical example, consider Zeno's dichotomy paradox that engendered the infinite series

$$\tfrac{1}{2} + \tfrac{1}{4} + \tfrac{1}{8} + \tfrac{1}{16} + \ldots \qquad (1)$$

(where we have assumed the distance D = 1). This can be rewritten

$$(\tfrac{1}{2}) + (\tfrac{1}{2})^2 + (\tfrac{1}{2})^3 + (\tfrac{1}{2})^4 + \ldots$$

which is of the form

$$r + r^2 + r^3 + r^4 + \ldots$$

where $r = \tfrac{1}{2}$. If we consider a finite sum $S_n = r + r^2 + r^3 + r^4 + \ldots + r^n$, then the infinite sum is just what happens to the S_n as n = 1, 2, 3, …, that is, as n grows without bound, written $n \to \infty$. Now

$$S_n - rS_n = r - r^{n+1}$$

So

$$S_n \;=\; \frac{r(1-r^n)}{1-r}$$

Since $r < 1$, as $n \to \infty$, $r^n \to 0$ and $S_n \to r/(1-r)$. For $r = \tfrac{1}{2}$, this means $S_n \to 1$. That is, the partial sums S_n approach 1 "in the limit" as n grows arbitrarily large. So we *define* the infinite sum in (1) to be 1, the limit of the sequence S_n. (This agrees with the physical fact that we can actually cover the distance D = 1 after all.)

Traditionally the infinite sum beginning with 1 and with r < 1

$$1 + r + r^2 + r^3 + \dots + r^{n-1} + \dots = 1/(1 - r) \tag{2}$$

is called the **geometric series**.

If for an infinite series, the sequence of partial sums $S_n = a_1 + a_2 + \dots + a_n$ approaches a limit S as $n \to \infty$, we say the series **converges** and the sum is S. If it does not approach any limit (grows without bound or oscillates toward more than one value like 1 – 1 + 1 – 1 + …), we say the series **diverges** and we cannot assign a sum to it.

Divergent Harmonic Series

Notice that if $S_n \to S$, then the nth term $a_n = S_n - S_{n-1} \to 0$. But $a_n \to 0$ is a necessary and not sufficient condition for convergence, for consider the example of the **harmonic series**

$$1 + {}^1/_2 + {}^1/_3 + {}^1/_4 + {}^1/_5 + \dots + {}^1/_n + \dots \tag{3}$$

Then $a_n = 1/n \to 0$, as $n \to \infty$. Now consider the partial sums

$$S_n = 1 + {}^1/_2 + {}^1/_3 + {}^1/_4 + {}^1/_5 + \dots + {}^1/_n$$

Note the behavior of the following subsequence of partial sums

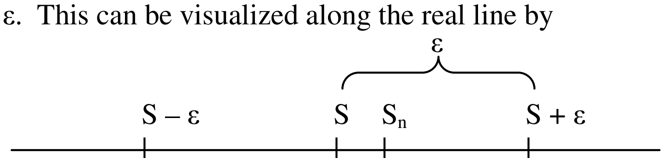
$S_2 = 1 + {}^1/_2$	$\geq 1 + {}^1/_2$	$= 1 + {}^1/_2$	$= 1{}^1/_2$
$S_4 = 1 + {}^1/_2 + {}^1/_3 + {}^1/_4$	$\geq 1 + {}^1/_2 + ({}^1/_4 + {}^1/_4)$	$= 1 + {}^1/_2 + {}^1/_2$	$= 2$
$S_8 = 1 + {}^1/_2 + {}^1/_3 + {}^1/_4 + \dots + {}^1/_8$	$\geq 1 + {}^1/_2 + \underbrace{({}^1/_4 + {}^1/_4)}_{2} + \underbrace{({}^1/_8 + \cdots + {}^1/_8)}_{4}$	$= 1 + {}^1/_2 + {}^1/_2 + {}^1/_2$	$= 2{}^1/_2$
…			
S_{2^k}	$\geq$		$= (k+2)/2$

This shows the subsequence S_{2^k} grows without bound. The whole sequence S_n cannot be converging to a finite limit if a subsequence is growing without bound. In fact, since for all n, $S_n < S_{n+1}$, all the other members of the sequence are being carried upwards by this increasing subsequence, so that the whole sequence grows without bound, that is, *diverges*.

The difference between the geometric series in (2) and the harmonic series in (3) is that the nth term r^n in (2) goes to zero much faster than 1/n in (3). Intuitively this means that eventually higher terms in the geometric series are negligible and can be ignored, but in the harmonic series the higher terms, though getting smaller, still contribute a non-negligible amount that accumulates to the point where it feeds back into the sum from the earlier terms.

Cauchy Definition of Limit

The way Cauchy codified this limit process in the 18^{th} century can be described as follows. S is called the **limit of an infinite sequence** of terms S_n if no matter how small a tolerance $\varepsilon > 0$ is chosen, one can find an integer N large enough such that for all n > N, S_n is within that tolerance ε of S, that is, $|S - S_n| < \varepsilon$. $|S - S_n|$ is the absolute value of the difference between S and S_n. That is, $-\varepsilon < S - S_n < \varepsilon$ or $S - \varepsilon < S_n < S + \varepsilon$. This can be visualized along the real line by



Another way of saying this is that eventually all the S_n will be in an open interval $(S - \varepsilon, S + \varepsilon)$ containing S. (An open interval (a, b) of the real line consists of all values x such that $a < x < b$. A closed interval [a, b] consists of all values x such that $a \le x \le b$.) So *Cauchy reduced the question of finding limits to solving sets of inequalities.*

Even though the religious Renaissance mathematicians might believe in an actualized infinity (Nicholas of Cusa said a circle *was* a polygon with an infinite number of sides), the concept of a limit avoids that idea, as did the method of Exhaustion. The best English equivalent would be the word "destination." It connotes an inevitable goal that has not yet been reached. It has an existence, but it only has meaning in relation to an ongoing process, not yet completed.

Geometric Series Example

To illustrate Cauchy's definition of a limit, consider again the geometric series in equation (2) and its partial sums $S_n = 1 + r + r^2 + r^3 + \ldots + r^{n-1}$. We claim $S_n \to S = 1/(1 - r)$, as $n \to \infty$. Now

$$S - S_n = \frac{1}{1-r} - \frac{1-r^n}{1-r} = \frac{r^n}{1-r}$$

We assumed $r < 1$ and implicitly that $r > 0$. Let's just assume $|r| < 1$, that is $-1 < r < 1$. Notice that we can assume $|r| > 0$, since if $r = 0$, then $S - S_n = 0$ for all n, and $S_n = S = 1$, and so 1 is trivially the limit. For any number a, if $0 < a < 1$, then $0 < a^2 < a$, and in fact $0 < a^{n+1} < a^n$. Hence, we can make a^n as small as we wish for sufficiently large n (there are some details here we ignore). Similarly, if we are given an arbitrarily small number $\varepsilon > 0$, we can find a suitably large N such that for all $n > N$, $|r|^n = |r^n| < \varepsilon\,|1 - r|$. Then

$$\left|S - S_n\right| = \left|\frac{r^n}{1-r}\right| < \frac{\varepsilon\left|1-r\right|}{\left|1-r\right|} = \varepsilon$$

And so S satisfies the limit definition.

Numerical Example. Suppose $r = {}^1/_2$ and $S_n = 1 + ({}^1/_2) + ({}^1/_2)^2 + ({}^1/_2)^3 + \ldots + ({}^1/_2)^{n-1}$. How big should N be so that for all $n \ge N$, $|S - S_n| < 0.0005$, that is, S_n approximates S (= 2) accurately to three decimal places? This means no matter how many terms are added after S_N, they will not contribute enough to cause rounding into the third decimal place.

So we want $|r^n| < \varepsilon\,|1 - r|$, that is,

$$({}^1/_2)^n < 5/10^4\ ({}^1/_2) \quad \text{or} \quad 10^4/5 < 2^{n-1} \quad \text{or} \quad 10^3 < 2^{n-2}$$

This means we want $n > 2 + 3/\log_{10}2 = 11.96$. So N = 12 should do the job. If we examine Table 1, we can see sure enough, S_{12} is within 0.0005 of S = 2.

So we now have a way to decide if an infinite sum can have a finite value assigned to it in a meaningful way, namely its partial sums must have a limit as n grows without bound. A similar limit idea can be given for defining areas and volumes.

Table 1

n	S_n	$S - S_n$
1	1.000000	1.000000
2	1.500000	0.500000
3	1.750000	0.250000
4	1.875000	0.125000
5	1.937500	0.062500
6	1.968750	0.031250
7	1.984375	0.015625
8	1.992188	0.007812
9	1.996094	0.003906
10	1.998047	0.001953
11	1.999023	0.000977
12	1.999512	0.000488

Limits – Area

Archimedes (c.287-212 BC).

Almost 2500 years ago Archimedes computed the area of a circle of radius r as πr^2 where π is the ratio of the circumference C of the circle to its diameter D = 2r. He did this by inscribing regular polygons of an increasing number of sides inside the circle and letting the number of sides increase without bound. The polygons were built starting with two triangles based on the diameter of the circle and inscribed in the corresponding semicircles. Subsequent polygons are generated by adding triangles based on the sides of the previous triangles and inscribed in the remaining space in the circle (see Figure 1). The nth polygon is the perimeter of the outer edges of the last set of triangles added in the nth step. If A_n represents the areas of the 2^n triangles added in the nth step (shown in a common color in Figure 1), then the area of the nth inscribed regular polygon P_n is

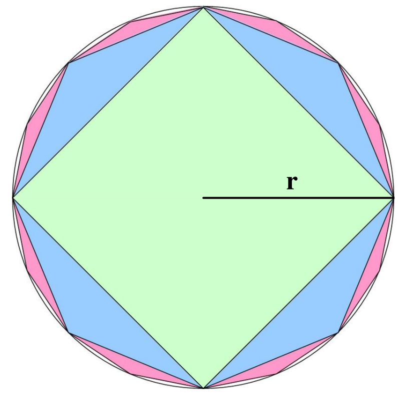


Figure 1 Archimedes Circle Computation

$$P_n = A_1 + A_2 + A_3 + \ldots + A_n \tag{4}$$

Rather than attempt an infinite sum, Archimedes argued using the Exhaustion Principle. He asserted that the sum was the area T of a triangle with altitude the radius r and base the circumference C of the circle, that is, $T = {}^1/_2\, r\, C = {}^1/_2\, r\, (\pi D) = \pi\, r^2$. He then argued by contradiction, that is, he supposed the area A of the circle was greater than T, T < A, and claimed that he could take a polygon with a sufficiently large number of sides n so that P_n is so close to A that $T < P_n < A$. But then he showed that all the areas of the inscribed polygons had to be *less* than the area $T = \pi r^2$—and so a contradiction. He then arrived at a similar contradiction by supposing A < T and using a series of circumscribing polygons that shrank to the area of the circle inscribed in all these polygons. Thus the only non-contradictory answer was that A had to equal T.

Numerical Example. I performed a computer implementation of equation (4) assuming a unit radius (r = 1). This means the area of the circle should just be π = 3.141592654... (to 9 decimal places). Table 2 shows the results. After 10 steps, the area of polygon P_{10} is already good to 5 decimal places (the error $\pi\, r^2 - P_{10} < 0.000005$). This is pretty fast convergence.

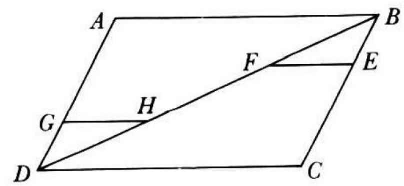
Table 2 Output for r = 1

Step n	Area Increment A_n	Area P_n	Area Diff πr^2 - P_n	Increment Ratio A_{n-1}/A_n
1	2.0000000	2.000000	1.1415927	
2	0.8284271	2.828427	0.3131655	0.4142100
3	0.2330403	3.061467	0.0801252	0.2813000
4	0.0599777	3.121445	0.0201475	0.2573700
5	0.0151033	3.136548	0.0050442	0.2518200
6	0.0037827	3.140331	0.0012615	0.2504500
7	0.0009461	3.141277	0.0003154	0.2501100
8	0.0002366	3.141514	0.0000789	0.2500300
9	0.0000591	3.141573	0.0000197	0.2500100
10	0.0000148	3.141588	0.0000049	0.2500000
11	0.0000037	3.141591	0.0000012	0.2500000

Out of curiosity I also computed the successive ratios of the incremental areas A_{n-1}/A_n. Very quickly the ratios tended toward $^1/_4$. So the sum of incremental areas approximated a geometric series with ratio $^1/_4$. Why this is of interest is that Archimedes carried out a similar construction of a series of inscribed triangular areas to find the area bounded by a parabola and line crossing its axis. In that case, he proved directly that the ratios of successive areas were all exactly equal to $^1/_4$, and then he computed the sums of the resulting geometric series! (It is hard to believe his achievements lay fallow for almost 2000 years.)

17^{th} Century (The Scientific Revolution)

With the embrace of the infinite after the Middle Ages and the Renaissance, mathematicians in the 17^{th} century could contemplate infinite processes directly. In the effort to compute general areas, an early idea was the notion of "indivisibles" as promulgated by Cavalieri ([1] pp.349-350):

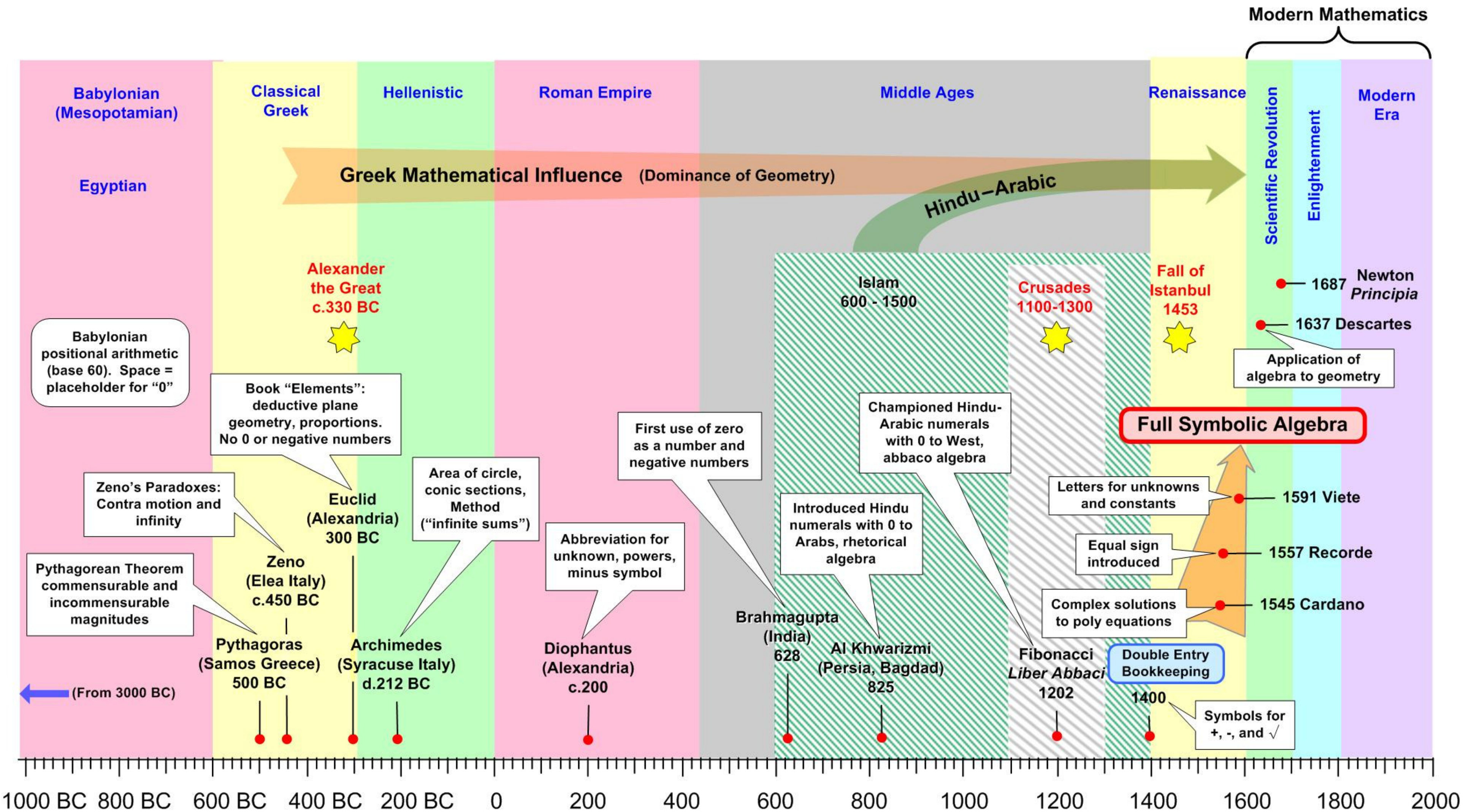
> Bonaventura Cavalieri (1598-1647), a pupil of Galileo and professor in a lyceum in Bologna, was influenced by Kepler and Galileo and urged by the latter to look into problems of the calculus. Cavalieri developed the thoughts of Galileo and others on indivisibles into a geometrical method and published a work on the subject, *Geometria Indivisibitibus Continuorum Notta quadam Ratione Promota* (Geometry Advanced by a thus far Unknown Method, Indivisibles of Continua, 1635). He regards an area as made up of an indefinite number of equidistant parallel line segments and a volume as composed of an indefinite number of parallel plane areas; these elements he calls the indivisibles of area and volume, respectively. Cavalieri recognizes that the number of indivisibles making up an area or volume must be indefinitely large but does not try to elaborate on this. Roughly speaking, the indivisibilitists held, as Cavalieri put it in his *Exercitationes Geometricae Sex* (1647), that a line is made up of points as a string is of beads; a plane is made up of lines as a cloth is of threads; and a solid is made up of plane areas as a book is made up of pages. However, they allowed for an infinite number of the constituent elements.
>
> Cavalieri's method or principle is illustrated by the following proposition, which of course can be proved in other ways. To show that the parallelogram ABCD (Fig. 17.7) has twice the area of either triangle ABD or BCD, he argued that when GD = BE, then GH = FE. Hence triangles ABD and BCD are made up of an equal number of equal lines, such as GH and EF, and therefore must have equal areas. …
>
> 
>
> **Figure 17.7**
>
> Cavalieri's indivisibles were criticized by contemporaries, and Cavalieri attempted to answer them; but he had no rigorous justification. At times he claimed his method was just a pragmatic device to avoid the method of exhaustion. Despite criticism of the method, it was intensively employed by many mathematicians. Others, such as Fermat, Pascal, and Roberval, used the method and even the language, sum of ordinates, but thought of *area as a sum of infinitely small rectangles rather than as a sum of lines*. [JOS: my emphasis]

This last statement refers to "infinitesimals," which we will discuss in a moment (below p.8). But first we need to understand some other mathematical developments that came to fruition in the 17^{th} century that played a role in the story of infinite processes.

Symbolic Algebra – Coordinate Systems – Functions

Symbolic Algebra (see Figure 2)**.** Since Babylonian and Egyptian times, cultures had evolved symbols (numerals) to represent numbers, but nothing more. There were no symbols for arithmetic operations such as addition, subtraction, multiplication, division, raising to a power, or taking roots, or for equality. What in modern times would be called algebraic word problems were all there was—rhetorical statements of problems with no translation into symbolic equations. Solutions were also verbal or geometric constructions.

For centuries mathematicians were aware of problems involving the solution of quadratic and cubic equations, but they were solved by geometric constructions. Ever since Pythagoras (c.500 BC) is credited with discovering that the hypotenuse of a right triangle with unit sides is irrational ($\sqrt{2}$), the Greeks limited their explicit use of numbers to whole numbers and ratios of whole numbers (via proportions), that is, rational numbers. Irrational numbers were only represented as lengths of lines, or areas or volumes and so manipulated through geometric constructions, usually with compass and ruler alone. A solution to a word problem involving quadratic equations, for example, was found as a length of a line in a geometric figure that was the result of a construction.



Mazur, *Enlightening Symbols* (2014); Kline, *Mathematical Thought from Ancient to Modern Times* (1972); Heeffer, "On the Nature and Origin of Algebraic Symbolism" (2009)

Figure 2 Timeline of Symbolic Algebra Development

But with the pressures of mercantilism in the Late Middle Ages and Renaissance and the arrival of the Hindu-Arabic numerals, including 0, calculators wanted more than just a mechanical means of computation, such as an abacus. And so algebraic symbols and notation arose during this period, coming to fruition at the beginning of the 17th century. For example, Viète (c. 1591) is credited with employing the (capitalized) letters toward the beginning of the alphabet (actually, consonants) to represent constants (e.g., B, C, D, etc.), and the letters toward the end of the alphabet (actually, vowels) to represent unknowns and variables (e.g., A, E, I, O, U). (But still all numbers were assumed positive, to reflect their physical origin.)

Coordinate system. French mathematicians Pierre de Fermat (1607-1665) and Rene Descartes (1596-1650) are credited with using the variables x, y in algebraic expressions to represent geometric figures, such as parabolas, hyperbolas, etc. How this began was similar to an aspect of the discussion above in Cavalieri's proof that the areas of the two triangles resulting from a diagonal of a parallelogram are equal (Figure 17.7). Cavalieri associated the varying line GH with the varying line DG. Fermat and Descartes would label the length of the line GH as y and the length of the line DG as x and thus have a way of designating the point H on the line by the pair of number x, y, called coordinates. Eventually, figures were oriented so that the x-coordinate was along a horizontal line (x-axis) and the y-coordinate (ordinate) was along a vertical line (y-axis). Then every point on the curve could be represented by a pair of coordinates (x, y). Descartes explored this arrangement quite extensively and showed how procedural manipulations of algebraic symbols could find solutions to problems that paralleled the geometric methods of constructing the solution, only much more easily. We give him credit for this approach by calling the coordinate system, Cartesian, and the subject matter, analytic geometry.

Functions

The dependency of the variable y on the variable x that we noted above was recognized by others as well, such as Galileo (1564-1642). This was the origin of the function concept. We say **y is a function of x** if to each value of x (independent variable) we can assign by some rule a *unique* value for y (dependent variable). Initially the rule was thought of as a formula or equation relating the variables, such as $y = 2x^2$ for a parabola, $y = 1/x$ for a hyperbola, or Galileo's formula $s = {}^1/_2\, gt^2$ for a body falling a distance s during a time t where g is the constant of acceleration. Later the functional relationship was written more abstractly as y = f(x) where f denoted the function. The curve of points (x, y) that corresponds to the functional relationship is called the **graph** of the function (Figure 3).

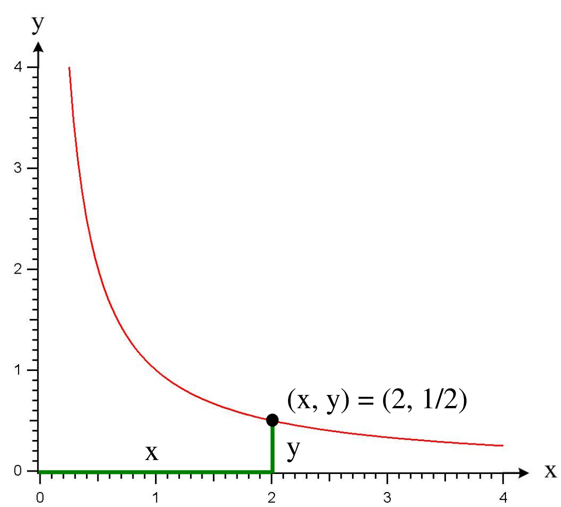
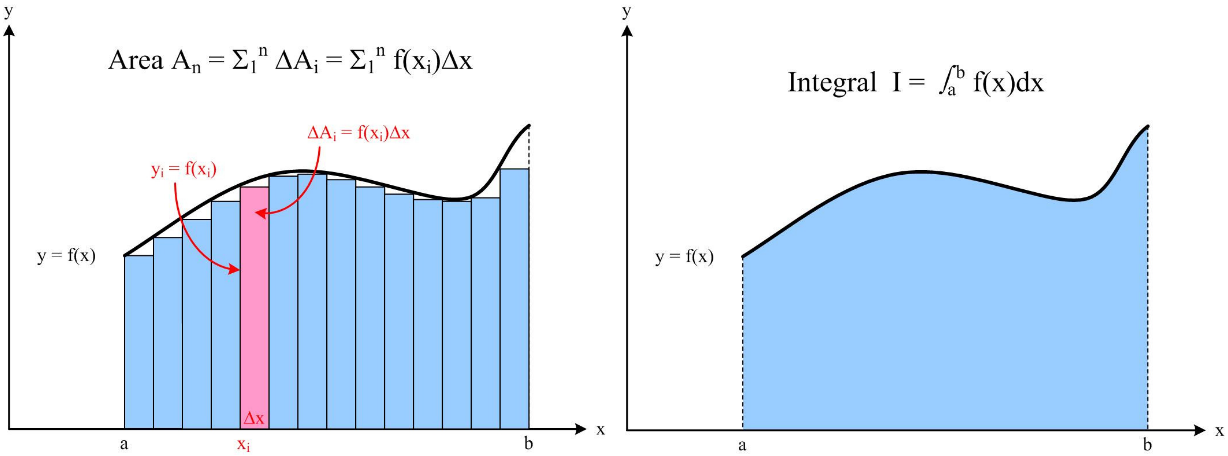


Figure 3 Graph of y = 1/x

Limits – Integration

Now we return to the challenge of trying to find areas inside regions bounded by curves. Consider the problem of trying to find the area under a curve (graph of y = f(x)) between the vertical lines x = a and x = b and the x-axis (see Figure 4(a)). Instead of using lines as Cavalieri did, we shall consider a set of n narrow rectangles inscribed in the region of vanishingly small width Δx ("Δ" represents "small change in") and height f(x) where the rectangle just touches the graph y = f(x). We then sum these rectangles to get an approximation to the area under the curve:



(a) (b)

Figure 4

$$\text{Area } A_n = \Sigma_{i=1}^n \Delta A_i = \Sigma_{i=1}^n f(x_i)\Delta x$$

As we shrink the width of the rectangles and thus increase the number of rectangles ($\Delta x \to 0$ and $n \to \infty$), the approximating sum of rectangles approaches the actual area in the limit. We call this limit the **integral of f(x) from a to b** and denote it $\int_a^b f(x)dx$ (see Figure 4(b)). So we have

$$\Sigma_{i=1}^n f(x_i)\Delta x \quad \to \quad \int_a^b f(x)dx$$
$$\text{as } \Delta x \to 0,\ n \to \infty$$

(The capital Greek letter for s, sigma Σ, is used to denote a sum, possibly infinite, of discrete entities. Leibniz, Newton's independent cofounder of calculus, used the notation on the right for the integral where the Greek sigma becomes an elongated s, $\int$. And Leibniz also converted the Δx to dx, which he called a differential and imagined it as an infinitesimal entity, that is, an infinitely small value that was not zero!—the more recent embodiment of Cavalieri's indivisible.)

There is an amazing and wonderful way to calculate integrals called the Fundamental Theorem of Calculus that we don't have time to go into now, unfortunately.

As an illustration, a very important function can be represented by the area under a curve. This is the natural logarithm of x, denoted ln(x). It is the logarithm to the base e (Euler's constant)[1] and is given by the area under the hyperbola y = 1/t from 1 to x. That is, if we allow the right hand endpoint of the interval of integration to vary, we have a rule that assigns to each value of x an area from 1 to x, and this number is ln(x). With a little bit of thought, we can see that

$$\ln(xy) = \ln(x) + \ln(y),$$

which is the essential logarithm property.

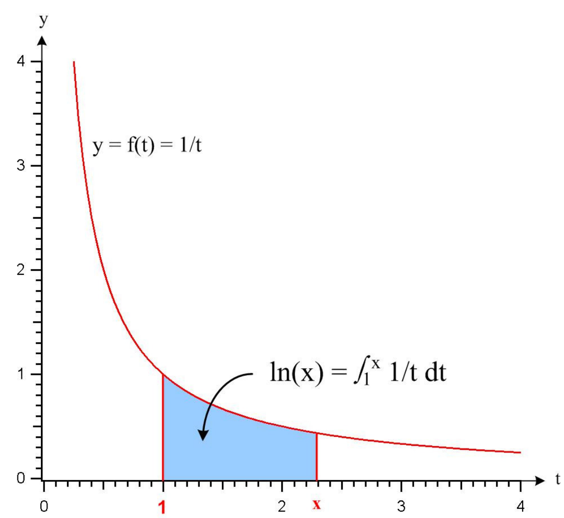


Figure 5 Natural Logarithm of x: ln(x)

---

[1] JOS: In other words, $y = \ln x = \log_e x$ also means $e^y = x$. Euler's constant e is the limit of the expression $(1 + {}^1/_n)^n$ as $n \to \infty$, which shows up in computations of compound interest and in many other places.

The question soon arises as to what types of functions have integrals (allow the process of approximating sums of rectangles to have a limit). Suppose we consider the function

$$f(x) = \begin{cases} 1/x & 0 < x \leq 1 \\ 0 & x = 0 \end{cases}$$

which is our hyperbola for $x > 0$ where we have added the value 0 at $x = 0$. If we divide the closed unit interval $[0, 1]$ $(0 \leq x \leq 1)$ into n equally spaced intervals $(\Delta x = 1/n)$, then $\Sigma_1^n\, f(x)\, \Delta x = \Sigma_1^n\, f(1/n)\ 1/n = n\ f(1/n)\ 1/n = f(1/n) = 1/(1/n) = n$, which grows without bound. The limit and therefore the integral does not exist.

One characteristic of this example is that we cannot draw its graph from $x = 1$ on the right to $x = 0$ on the left without lifting our pencil and jumping to (0, 0). A function with the property of being able to draw its graph without lifting the pencil is said to be **continuous**.

It turns out that *if a function is continuous over a closed interval [a, b], then it has an integral there*. But even some discontinuous functions can have integrals, so long as they are bounded. It was the study of the nature of these points of discontinuity that was one of the motivations for point set topology.

Limits – Continuous Functions

Again Cauchy codified the meaning of a continuous function which we turn to now. It is no surprise that it is based on the limit idea. A function $y = f(x)$ is said to be continuous at a value $x = a$, if it is defined there ($f(a)$ exists), $f(x)$ approaches a limit L as x approaches a, and that limit $L = f(a)$. That is, $f(x) \to f(a)$ as $x \to a$. Using the limit terminology, this means no matter how small a tolerance $\varepsilon > 0$ we place around $f(a)$ we can find a tolerance $\delta > 0$ around a such that whenever $|\,x - a\,| < \delta$, we have $|\,f(x) - f(a)\,| < \varepsilon$. (By the way, it is perhaps unfortunate that traditionally the tolerances are represented by the Greek letters epsilon ε and delta δ.) In terms of open intervals, we can say that given any small open interval $(f(a) - \varepsilon,\ f(a) + \varepsilon)$ around $f(a)$ we can find a small open interval $(a - \delta,\ a + \delta)$ around a such that $f\,[(a - \delta,\ a + \delta)] \subset (f(a) - \varepsilon,\ f(a) + \varepsilon)$. (For sets A and B, the expression $f(A) \subset B$ means for every $x \in A$, $f(x) \in B$, where $\in$ means "in" or "is an element of".) See Figure 6.

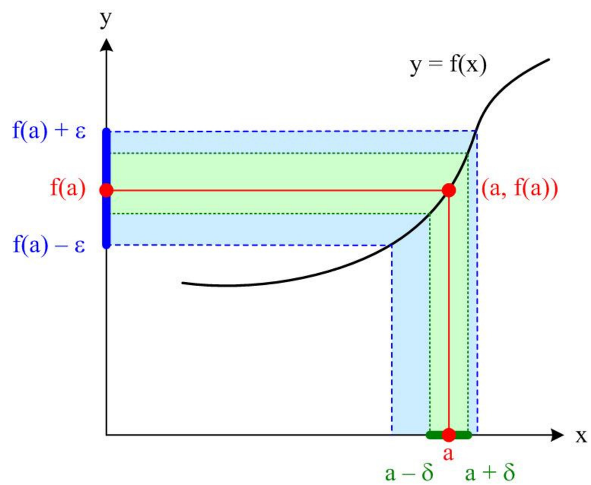


Figure 6 Continuous Function at a Point a

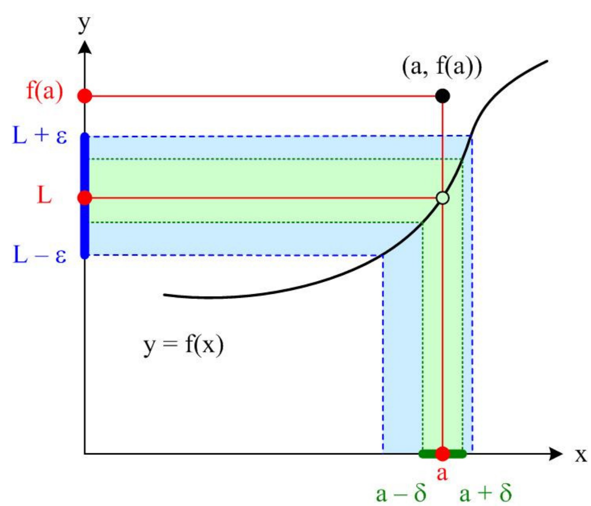


Figure 7 Discontinuous at x = a

Examples of Discontinuous Functions

Right off we can cite the hyperbola $y = 1/x$ as discontinuous at $x = 0$, since the limit as $x \to 0$ does not exit. Similarly the potential sum of the geometric series $1/(1 - x)$ is discontinuous at $x = 1$ because again the limit as $x \to 1$ does not exit. Figure 7 shows a generic discontinuous case where even though the limit L exists and the function is defined at $x = a$, $f(a) \neq L$.

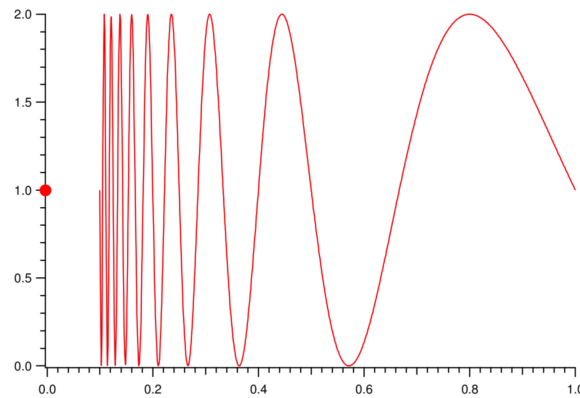


Figure 8 $y = 1 + \sin(2\pi/x)$

Figure 8 shows the graph of the function $f(x) = 1 + \sin(2\pi/x)$ for $0 < x \leq 1$ and $f(0) = 1$. $f(x)$ is defined and bounded on the closed interval $[0, 1]$ but it has no limit as $x \to 0$ since it oscillates between 0 and 2 infinitely often (the graph could not be shown all the way to 0).

Finally there is the canonical example of a function that is continuous *nowhere*.

$$f(x) = \begin{cases} 1 & x \text{ rational} \\ 0 & x \text{ irrational} \end{cases}$$

$f(x)$ is defined everywhere on the real line, but is continuous at no point. This is because of the property of the real line that between every two rationals there is an irrational and between every two irrationals there is a rational. So if a is an irrational point so that $f(a) = 0$, then as $x \to a$, there will always be a rational point z between x and a so that $f(z) = 1$. Therefore $f(x) \not\to 0 = f(a)$. Similarly if a is a rational point so that $f(a) = 1$. As $x \to a$, there will always be an irrational point z between x and a so that $f(z) = 0$. Therefore $f(x) \not\to 1 = f(a)$.

Limits – Sequences of Points

Decimal expansions.

Recall how we can obtain a decimal expansion for $\sqrt{2}$.

$1^2 = 1 < 2 < 4 = 2^2$

$\Rightarrow 1 < \sqrt{2} < 2$

$1.4^2 = 1.96 < 2 < 2.25 = 1.5^2$

$\Rightarrow 1.4 < \sqrt{2} < 1.5$

$1.41^2 = 1.9881 < 2 < 2.0164 = 1.42^2$

$\Rightarrow 1.41 < \sqrt{2} < 1.42$

$1.414^2 = 1.999396 < 2 < 2.002225 = 1.415^2$

$\Rightarrow 1.414 < \sqrt{2} < 1.415$

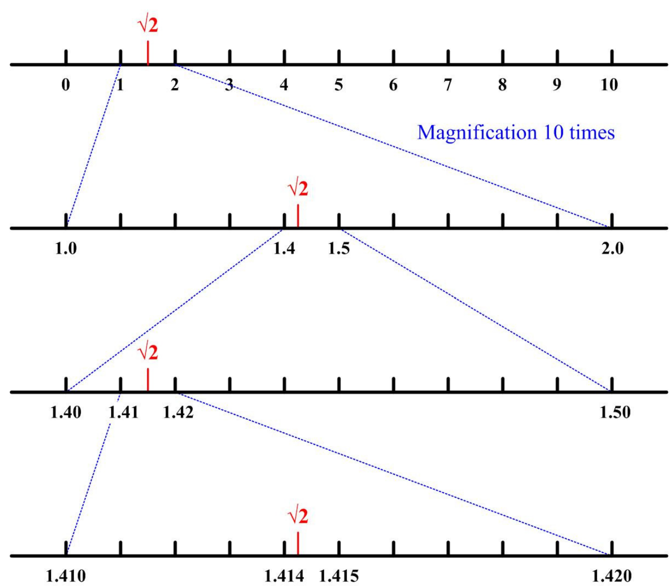


Figure 9 Decimal expansion for $\sqrt{2}$

Figure 9 shows geometrically what is happening. First we find that $\sqrt{2}$ lies between 1 and 2. We divide the interval between 1 and 2 into 10 parts each of width 1/10. We find $\sqrt{2}$ lies between 1.4 and 1.5, so we divide that interval again into 10 parts, now of width 1/100. We keep going in this way keeping $\sqrt{2}$ in an interval 1/10 the length of the previous. Table 3 shows that the left-hand endpoints of the intervals approximate $\sqrt{2}$ better and better, and these successive endpoints represent the decimal expansion of $\sqrt{2}$. Figure 10 shows this graphically as well.

Table 3

$\sqrt{2} - 1.4$	$< 1.5 - 1.4$	$= .1$	$= 1/10$
$\sqrt{2} - 1.41$	$< 1.42 - 1.41$	$= .01$	$= 1/10^2$
$\sqrt{2} - 1.414$	$< 1.415 - 1.414$	$= .001$	$= 1/10^3$
$\sqrt{2} - 1.4142$	$< 1.4143 - 1.4142$	$= .0001$	$= 1/10^4$

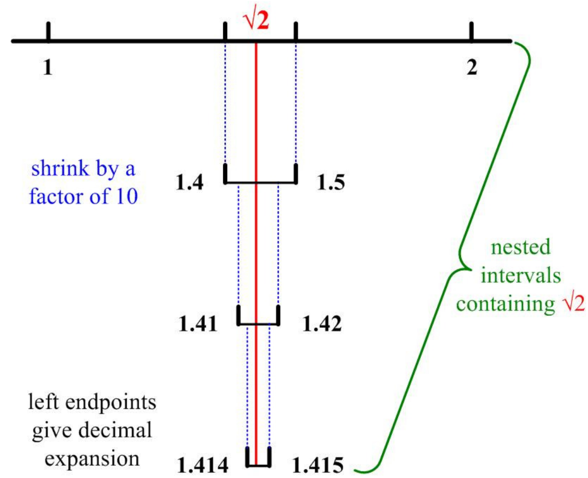


Figure 10 Nested intervals all containing $\sqrt{2}$

In fact $\sqrt{2}$ is the only point which belongs to all the nested intervals whose left endpoints yield the decimal expansion.

Thus for each point on the line we get a unique sequence of nested intervals containing only that point in common, and the left endpoints of the intervals yield a decimal expansion for the point (repeating or non-repeating). Conversely, for any decimal expansion, there is a corresponding sequence of nested intervals taking their left endpoints from the expansion, each having one tenth the length of its predecessor, and all having only one point in common.

Thus we get a correspondence between all the points on the line and decimal expansions, where repeating decimals correspond exactly to the points representing rational numbers and non-repeating expansions correspond to irrational numbers. And so every real number corresponds to a unique point on the number line.

Example: 0.9999… "=" 1.

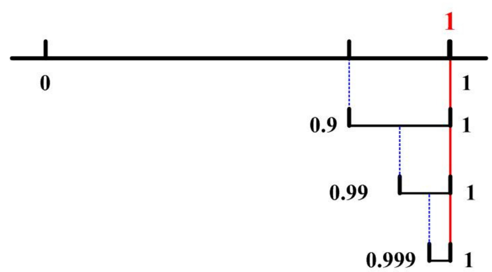


Figure 11 0.9999… = 1

The nested interval idea provides another way of seeing the basis for the seemingly mysterious statement that the number 1 has an alternative decimal expansion as 0.999999…. (Figure 11). We see that the 9s in the decimal expansion mean we are looking at the right-most interval within the closed interval [0, 1] and that means the number 1 is always included as the right endpoint of each subinterval. Clearly it is the only number that will be included in all the nested intervals, and so 0.999… represents its decimal expansion. But so does 1.0000…. So we have an instance where our nested interval definition provides two possible decimal expansions for the same number. Therefore they must be equivalent. (We didn't explicitly discuss what happens when the point of interest turns out to land on an endpoint of our subintervals.)

Limit Points

A point on the number line corresponding to an irrational number has the interesting characteristic of being the limit of a sequence of points corresponding to rational numbers (left endpoints of intervals). A point x on the number line is called a **limit point of a set A**, if every open interval containing x also contains a point from A. For example, we saw that for every open interval $(\sqrt{2} - 1/10^n, \sqrt{2} - 1/10^n)$ of $\sqrt{2}$, we could find a smaller (by a tenth) closed interval $[\sqrt{2} - 1/10^{n+1}, \sqrt{2} - 1/10^{n+1}] \subset (\sqrt{2} - 1/10^n, \sqrt{2} - 1/10^n)$ so that the left endpoint (term in the rational decimal expansion) is contained in that open interval. Therefore, the irrationals correspond to limit points of the rationals.

Any set that contains all its limit points is called **closed**. Notice that a closed interval [a, b] is a closed set using this definition. For if c is any point not in [a, b], say c < a (Figure 12), then choose

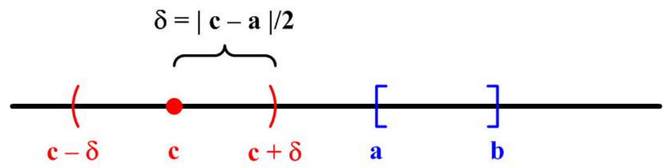


Figure 12   Proof [a, b] is a closed set

$\delta = |c - a|/2$. The open interval $(c - \delta, c + \delta)$ contains no points of [a, b] and so c cannot be a limit point of [a, b]. (A similar argument holds if $b < c$.) That means [a, b] contains all its limits points, and so is closed. (Notice that the endpoints, a and b, are limit points of [a, b].)

Limits – Sequences of Functions

Power Series

We return to a discussion of functions. We showed above (Geometric Series Example, p.4), replacing r by x, that $1/(1 - x) = 1 + x + x^2 + x^3 + \ldots$ so long as $|x| < 1$. This infinite series is of the form $S(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \ldots$ and is called a **power series**. Its partial sums are polynomials $S_n(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \ldots + a_{n-1} x^{n-1}$. Then the geometric series is a power series where all the coefficients $a_k = 1$.

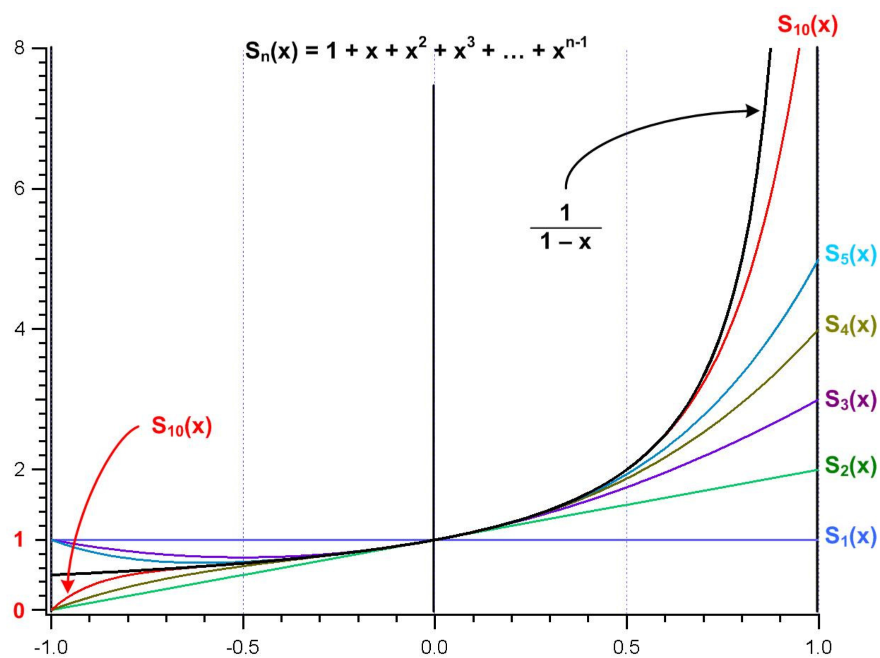


Figure 13   Partial sum functions converging to geometric series on (-1, 1)

Example – Geometric Series. Figure 13 shows the graph of $1/(1 - x)$ with some of the partial sums. Several things are of interest. We can see that the partial sum functions approximate $1/(1 - x)$ better and better throughout the open interval (-1, 1). But strange things happen at the endpoints. Of course, at 1 each partial sum $S_n(1) = n$, a finite number, which cannot approximate the unbounded $1/(1 - x)$ there. Still the $S_n(1)$ are growing without bound and trying to catch up to $1/(1 - x)$, as it were. The left hand endpoint -1 is the really interesting feature. The function $1/(1 - x)$ is actually defined there and is 1/2. But the partial sums $S_n(1) = {}^1/_2 + (-1)^{n-1}\,{}^1/_2$ oscillate between 1 and 0, that is, the geometric series looks like $1 - 1 + 1 - 1 + \ldots$ at $x = -1$, which diverges. Nevertheless, the partial sum functions $S_n(x)$ do get closer to ${}^1/_2$ as $x \to -1$, but turn away at the last moment to reach either 1 or 0 at the endpoint, depending on whether n is odd or even.

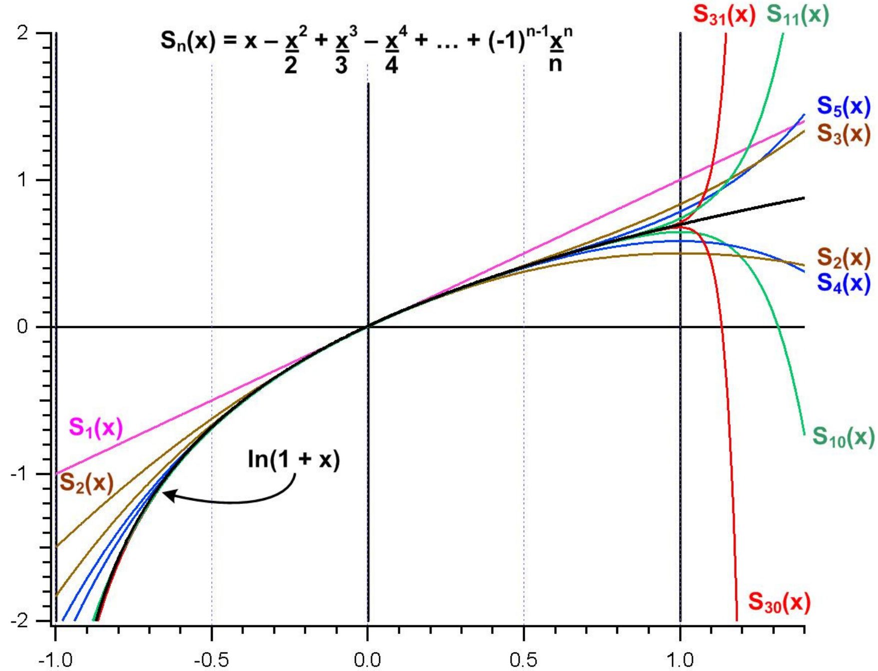


Figure 14 Partial sum functions converging to ln(1 + x) series on (-1, 1]

Example – Logarithmic Series. It turns out that ln(1 + x) has a power series expansion

$$\ln(1 + x) = x - x^2/2 + x^3/3 - x^6/6 + \ldots + (-1)^{n-1} x^n/n + \ldots$$

which also converges for $|x| < 1$. In fact it also converges at its right endpoint x = 1. That series is an alternating series 1 – 1/2 + 1/3 – 1/4 + … which does converge. At the left endpoint, x = -1, we have the negative of the harmonic series, –1 – 1/2 – 1/3 –1/4 – …, which we already know diverges. Figure 14 shows the graph of ln(1 + x) and some partial sum functions $S_n(x)$. What is really strange from the plots is the behavior of the partial sum functions outside the interval (-1, 1] on the right. They seem to abruptly turn off in all directions.

There is something curious about the interval of convergence for the power series. We have seen both examples are symmetric about 0. If we consider a more general power series at a point x_0 different from 0, it takes the form

$$S(x - x_0) = a_0 + a_1 (x - x_0) + a_2 (x - x_0)^2 + a_3 (x - x_0)^3 + \ldots$$

Our previous expression was just the power series when $x_0 = 0$. Note $S(0) = a_0$, so S(x) or $S(x - x_0)$ converges for at least one point, namely, $x = x_0$. Now we have the important result, as stated in *Wikipedia*: "If x_0 is not the only convergent point, then there is always a number r with $0 < r \le \infty$ such that the series converges whenever $|x - x_0| < r$ and diverges whenever $|x - x_0| > r$. The number r is called the **radius of convergence** of the power series." ("$r = \infty$" means the series converges everywhere along the real line.) So a power series will *always* converge in a symmetric interval about x_0. Therefore both of our example series have a radius of convergence = 1.

An amazing result is that all the basic functions of interest in calculus have power series expansions :

$e^x \quad = 1 + x + x^2/2! + x^3/3! + \ldots + x^n/n! + \ldots$

$\sin x \quad = x - x^3/3! + x^5/5! - x^7/7! + \ldots + (-1)^n x^{2n+1}/(2n+1)! + \ldots$

$$\cos x \quad = 1 - x^2/2! + x^4/4! - x^6/6! + \ldots + (-1)^n x^{2n}/(2n)! + \ldots$$

$$\ln(1 + x) = x - x^2/2 + x^3/3 - x^6/6 + \ldots + (-1)^n x^n/n + \ldots$$

where n = 0, 1, 2, … and n! = n(n – 1)(n – 2) … 3•2•1. For example, 5! = 5•4•3•2•1 = 120. To satisfy the schemata, we define 0! = 1.

So we can consider all the basic functions of interest in calculus as being "infinite" polynomials. We can add, subtract, multiply, and divide power series like polynomials. But more powerfully we can integrate (and differentiate, which we have not discussed) like polynomials, which is easy. This is because

$$S_n(x) \to S(x) \;\Rightarrow\; \int S_n(x)dx \to \int S(x)dx \tag{4}$$

This is effectively the approach Newton took in his calculations with the early calculus.

For example, using the geometric power series (slightly modified with + x)

$$1/(1 + x) = 1 - x + x^2 - x^3 + \ldots$$

we have

$$\ln(1 + x) = \int 1/(1 + x)\, dx = \int (1 - x + x^2 - x^3 + \ldots)dx = x - x^2/2 + x^3/3 - x^6/6 + \ldots$$

We have ignored whether all these operations are legal, that is, whether all the various limits, such as in equation (4), exist. As mathematicians began to worry about this, they discovered a condition that resolved the problem, to be discussed next.

Uniform convergence

When we discussed limits for sequences before, we were restricted to numbers, not functions. We said a sequence of numbers S_n had a numerical limit S, if no matter how small an $\varepsilon > 0$ we chose, there would be an N so large that for all n > N, $| S - S_n | < \varepsilon$. This idea still works for functions in the following way. For each value of x in the interval I of convergence, the $S_n(x)$ form a sequence of numbers approaching S(x) in the limit. So then we would write given $\varepsilon > 0$, there would be an N so large that for all n > N, $| S(x) - S_n(x) | < \varepsilon$. The key here is that in general N depends on x. That is, for a fixed $\varepsilon > 0$, we might have to choose a different large N for each choice of x, since there would be a different resulting sequence of numbers. On the other hand, it might happen that one N would work for all the x in the interval I of interest. That is, for all n > N, and for all x in the interval I, $| S(x) - S_n(x) | < \varepsilon$. Equivalently, this would be for all n > N,

$$S(x) - \varepsilon < S_n(x) < S(x) + \varepsilon$$

for all $x \in I$. This latter type of convergence is called **uniform convergence** of a sequence of functions. The situation is illustrated in Figure 15. An equivalent way of expressing this is to say given any $\varepsilon > 0$, there is an N so large that for all n > N, $\max_{x \in I} | S(x) - S_n(x) | < \varepsilon$.

It can be shown that if r is the radius of convergence of a power series around $x = x_0$ with partial sum functions $S_n(x)$, then the S_n converge uniformly to the sum S on any closed interval $[a, b] \subset (x_0 - r,\ x_0 + r)$. And so $S_n \to$ S uniformly $\Rightarrow \int_a^b S_n \to \int_a^b S$ (equation (4)).

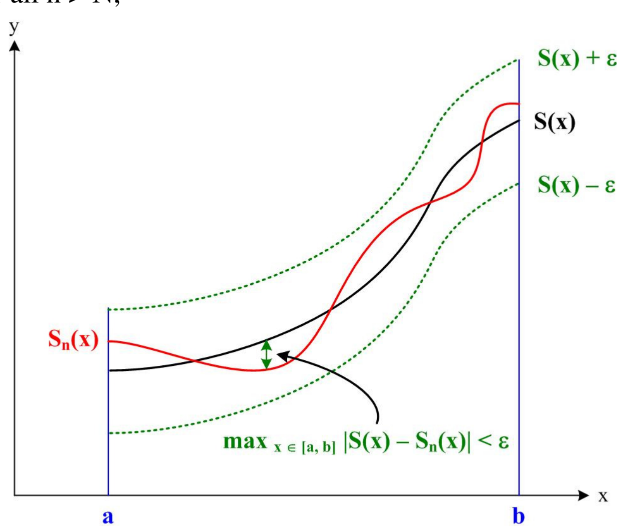


Figure 15 Uniform convergence

Metric Spaces

We have almost reached our goal of describing point set topology. All the things we have discussed culminated in the beginning of the 20^{th} century with a beautiful abstraction that was enormously fruitful.

Suppose we designate the set of continuous functions on a closed interval [a, b] by C[a, b]. Then as we mentioned above (p.10) every function in C[a, b] has an integral. Therefore the integral defines a "function" (transformation) F from the set C[a, b] to the real numbers $\mathbb{R}$, written

$$F : f \in C[a, b] \to \int_a^b f \in \mathbb{R} \tag{5}$$

Now we need to make explicit what we mean by $f_n \to f \;\Rightarrow\; F(f_n) \to F(f)$ (equation (4)) which looks very much like a continuity criterion.

First notice that the expression $| x - y |$ where x and y are two real numbers measures the distance between them along the real line. We might alternatively write this d(x, y) . Then this distance $d(x, y) \in \mathbb{R}$ has the following properties:

(1) $d(x, y) = d(y, x) \geq 0$

(2) $d(x, y) = 0$ if and only if $x = y$

(3) $d(x, y) \leq d(x, z) + d(z, y)$ for any other real number z (the triangle inequality)

In all the arguments above where we used the absolute value $| x - y |$, we could have used d(x, y). According to Kline ([1] p.1079) this general, axiomatic definition of distance was presented by the French mathematician Maurice Fréchet in the beginning of the 20^{th} century and is called a **metric**. Any space on which a metric is defined is called a **metric space**. All the notions of limits of sequences and variables, including continuity, that we discussed above carry over into metric spaces (just change $| x - y |$ to d(x, y)).

For example, the distance between points on the real number line $\mathbb{R}$ can be generalized to distance between points in the plane $\mathbb{R}^2$ or in 3-dimensional space $\mathbb{R}^3$. Suppose P and Q are two points in the plane $\mathbb{R}^2$ with coordinates (x_1, y_1) and (x_2, y_2), respectively. Then define the distance between them d(P, Q) using the Pythagorean Theorem by

$$d(P, Q) = [\, | x_1 - x_2 |^2 + | y_1 - y_2 |^2 \,]^{1/2}$$

(See Figure 16) It is not hard to show d(P, Q) satisfies the properties (1) – (3) for a metric on $\mathbb{R}^2$. d(P, Q) is called the **Euclidean metric**. Notice that if P and Q are restricted to the real line $\mathbb{R}$, then $d(P, Q) = | x_1 - x_2 |$, the original absolute value metric.

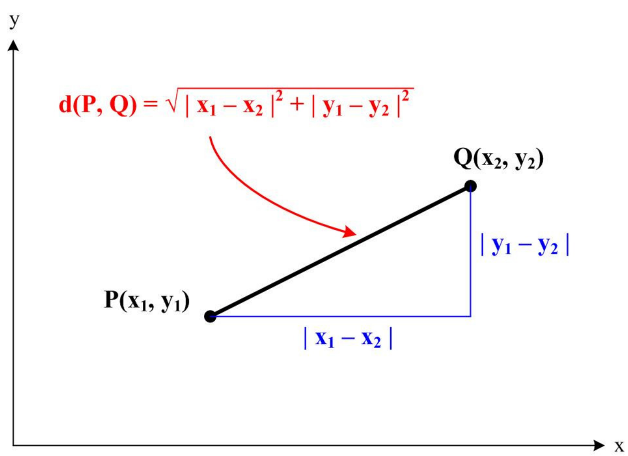


Figure 16 Euclidean metric in $\mathbb{R}^2$

Fréchet defined a d(f, g) on C[a, b] by

$$d(f, g) = \max_{x \in [a, b]} | f(x) - g(x) |$$

He then showed d(f, g) satisfied all the properties (1) – (3) for a metric on C[a, b]. This is sometimes called the **max metric** or **uniform metric** and it captures the idea of uniform convergence. From the discussion above we saw that with uniform convergence we can essentially ignore the independent variables $x \in [a, b]$. We can now write $f_n \to f$ to mean that for every $\varepsilon > 0$, we can find an N such that for all $n > N$, $d(f, f_n) < \varepsilon$. Then the real-valued transformation F can be shown to be continuous using the max metric. That is, F is continuous at f if for every $\varepsilon > 0$, we can find a $\delta > 0$ such that if $d(g, f) < \delta$, then $d(F(g), F(f)) < \varepsilon$, where the second metric d(,) is just the absolute value metric on the reals.

It is possible to define other metrics on C[a, b]. For example,

$$d(f, g)_2 = \left[\int_a^b | f(x) - g(x) |^2 \, dx \right]^{½}$$

is called the **L_2 metric** and is the function analog of the Euclidean metric. The convergence of Fourier series is defined with the L_2 metric.

We can also couch these ideas in terms of neighborhoods. Define a **neighborhood** N(p, r) of a point p in a metric space to be the set of points q in the space such that $d(p, q) < r$, for some $r > 0$. Then a function $f: X \to Y$ from one metric space X to another Y is continuous at p in X if for every neighborhood $N(f(p), \varepsilon)$ of f(p) in Y we can find a neighborhood $N(p, \delta)$ of p in X such that $f(N(p, \delta)) \subset N(f(p), \varepsilon)$.

Fréchet reduced the complicated idea of a function to a point in some metric space. We can talk about limit points and closed sets in the function space C[a, b]. Fréchet's abstraction launched one of the most fruitful mathematical pursuits in the 20th century called functional analysis.

Topological Spaces

From Kline's history ([1] pp.1159-60):

> The origins of point set topology have already been related (Chap.46, sec. 2). Fréchet in 1906, stimulated by the desire to unify Cantor's theory of point sets and the treatment of functions as points of a space, which had become common in the calculus of variations, launched the study of abstract spaces. The rise of functional analysis with the introduction of Hilbert and Banach spaces gave additional importance to the study of point sets as spaces. The properties that proved to be relevant for functional analysis are topological largely because limits of sequences are important. Further, the operators of functional analysis are transformations that carry one space into another.
>
> As Fréchet pointed out, the binding property need not be the Euclidean distance function. He introduced (Chap. 46, sec. 2) several different concepts that can be used to specify when a point is a limit point of a sequence of points. In particular he generalized the notion of distance by introducing the class of metric spaces. In a metric space, which can be two-dimensional Euclidean space, one speaks of the neighborhood of a point and means all those points whose distance from the point is less than some quantity ε, say. Such neighborhoods are circular. One could use square neighborhoods as well. However, it is also possible to suppose that the neighborhoods, certain subsets of a given set of points, are specified in some way, even without the introduction of a metric. Such spaces are said to have a neighborhood topology. This notion is a generalization of a metric space. Felix Hausdorff (1868-1942), in his *Grundzüge der Mengenlehre* (Essentials of Set Theory, 1914), used the notion of a neighborhood (which Hilbert had already used in 1902 in a special axiomatic approach to Euclidean plane geometry) and built up a definitive theory of abstract spaces on this notion.

And so we arrive at the end of the road at our goal of general or point set topology. I thought I would close with the abstract definition for topology given by *Wikipedia*. Hopefully the terms used in the definitions will be less opaque and the reason for the abstraction more understandable. One of the great accomplishments of 20th century mathematics was the discovery of how ideas and constructs in one area could be adapted and generalized to provide unusual insights in another area. Functional Analysis based on point set topology is one of the supreme examples of this accomplishment.

Wikipedia Definitions

(*Wikipedia* topology:) In mathematics, **topology** (from the Greek τόπος, place, and λόγος, study) is concerned with the properties of space that are preserved under continuous deformations, such as stretching, crumpling and bending, but not tearing or gluing.

(*Wikipedia* general topology:) In mathematics, **general topology** is the branch of topology that deals with the basic set-theoretic definitions and constructions used in topology. It is the foundation of most other branches of topology, including differential topology, geometric topology, and algebraic topology. Another name for general topology is **point set topology**.

The fundamental concepts in point set topology are continuity, compactness, and connectedness:

- Continuous functions, intuitively, take nearby points to nearby points.
- Compact sets are those that can be covered by finitely many sets of arbitrarily small size.
- Connected sets are sets that cannot be divided into two pieces that are far apart.

The words 'nearby', 'arbitrarily small', and 'far apart' can all be made precise by using [neighborhoods].[2] If we change the definition of '[neighborhood]', we change what continuous functions, compact sets, and connected sets are. Each choice of definition for '[neighborhood]' is called a topology. A set with a topology is called a topological space.

(*Wikipedia* topological space:) In topology and related branches of mathematics, a **topological space** may be defined as a set of points, along with a set of neighbourhoods for each point, satisfying a set of axioms relating points and neighbourhoods. The definition of a topological space relies only upon set theory and is the most general notion of a mathematical space that allows for the definition of concepts such as continuity, connectedness, and convergence. Other spaces, such as manifolds and metric spaces, are specializations of topological spaces with extra structures or constraints. Being so general, topological spaces are a central unifying notion and appear in virtually every branch of modern mathematics. The branch of mathematics that studies topological spaces in their own right is called point set topology or general topology.

References

1. Kline, Morris, *Mathematical Thought from Ancient to Modern Times*, Oxford University Press, New York, 1972 (https://archive.org/details/MathematicalThoughtFromAncientToModernTimes, retrieved 5/26/2017).

   This 1200 page book is probably the best history of mathematics available. It is quite detailed, comprehensive, and up-to-date, at least to the mid-20th century. It generally does a good job of explaining the origin and evolution of mathematical ideas, though it assumes a basic familiarity with algebra and elementary calculus (infinite series, differentiation, and integration). Shorter books may provide a better overview for the novice, but they tend to short-change the math.

2. Manheim, Jerome H., *The Genesis of Point Set Topology*, MacMillan Co., New York, 1964

   This was the first book I read about the significance of Fourier Series and its impact on the development of mathematics, in particular point set topology (and Cantor's theory of transfinite

[2] JOS: The original term in *Wikipedia* is "open set", which turns out to be equivalent, but I thought less intuitive at this point. The use of neighborhoods here relates more directly to the ideas we have been discussing.

numbers). But it also probably only makes sense to someone who already has a background in calculus. Moreover, it is long out-of-print and there is no online digital version.

3. Bressoud, David M., *A Radical Approach to Real Analysis*, Second Edition, Mathematical Association of America, Washington D.C., 2007

   This is an excellent book. But it is a real math book, that is, it is developing, and expecting the student to learn, real advanced mathematics, but from a historical, rather than logical, perspective. It begins with the crisis caused by Fourier Series in the beginning of the 19th century and then shows how mathematicians developed a deeper understanding of the mathematics of infinite series and differentiation to solve the problems.

4. Bressoud, David M., *A Radical Approach to Lebesgue's Theory of Integration*, Cambridge University Press, 2008.

   This is another excellent book. But again it is a real math book. It extends the development of mathematics in response to the crisis caused by Fourier Series to the origin of point set topology for the real line and to a deeper understanding of integration, first with Riemann integration and then with Lebesgue integration.