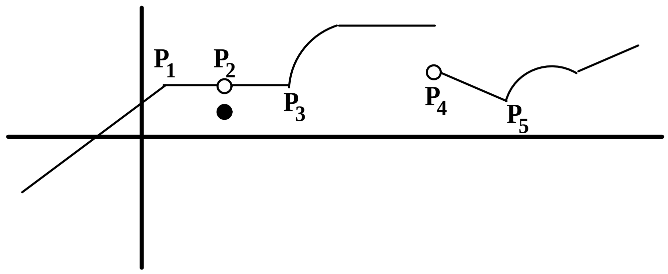
Continuity

We have used the term *continuity* in connection with many of the functions we have introduced. It is, roughly, synonymous with "unbroken"—by and large the elementary functions of calculus have graphs that can be drawn "without lifting the pencil from the paper". Some functions have breaks in their graphs, as does for example the function $f(x) = x/|x|$ that takes on the value 1 for $x > 0$ and the value -1 for $x < 0$. Drawing its graph requires "lifting the pencil" at the point $x = 0$ to jump over the break. Similarly the function whose graph is represented below is not continuous at the points P_2 and P_4. At P_2 the value of the function is "out of place", and at P_4 the graph takes a "jump". On the other hand at all other points the graph (and hence the function) is continuous. The points P_1, P_3 and P_5 are not breaks in the graph. The graph turns a sharp corner at these points, but the pencil need not be lifted.



Except in a few circumstances our intuition about continuity is quite sound. In a very precise sense we shall see that the Elementary Functions of calculus are, indeed, continuous at nearly every point and that the few places where they are not continuous are easily recognized. For example the rational function defined in Example 9 of Section 2.4 is continuous except at the two points $x = 2, 3$ where division by zero is excluded. And the function $\tan x = \sin x/\cos x$ is continuous everywhere except the points $x = \pi/2 + n \cdot \pi$ where the cosine function vanishes (cf. Example 2 in Section 1.4).

In preparation for the next definition, we will say that an *interior point* of a set of real numbers is a point that can be enclosed in an open interval that is contained in the set.

Definition 1: A function is *continuous* at an interior point c of its domain if $\lim_{x\to c} f(x) = f(c)$. If it is not continuous there, i.e. if either the limit does not exist or is not equal to $f(c)$ we will say that the function is *discontinuous* at c.

Note that we are requiring that

1. The function f is defined at the point $x = c$,
2. The point $x = c$ is an interior point of the domain of f,
3. $\lim_{x\to c} f(x)$ exists, call it L, and
4. $L = f(c)$.

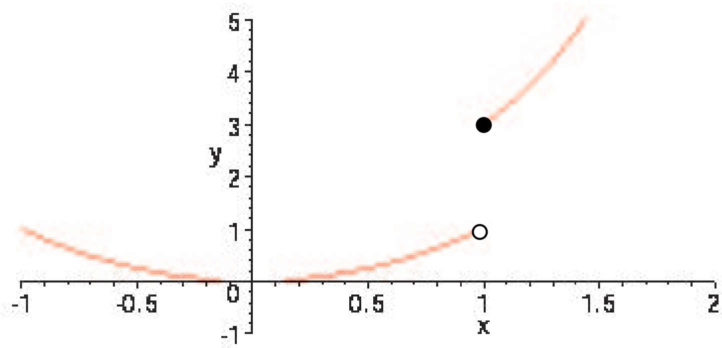
It is the notion of limit, once again, that enables us to capture precisely the idea of continuity. The existence of the limit of $f(x)$ as $x \to c$ says, in effect, that no matter how powerful a microscope we might use to examine the graph of $f(x)$ in the vicinity of $x = c$, we would never see a "break" in the graph. The right-hand and left-hand limits both exist at c and have the same value, and moreover that value coincides with the value of $f(c)$. The plot shown above illustrates a function that has two places where it is not continuous. At P_2 the limit exists (the right-hand and left-hand limits exist and are equal) but is not equal to the value of the function. At P_4 the limit does not exist (the left-hand limit does exist and is the value of the function, but the right-hand limit has a different value). In both cases we can characterize the graph as "broken" and the function as not continuous. We will see below that we can say further, that the discontinuity at P_2 is *removable* (in a sense it is less serious: we need only redefine the value of the function at that one point).

And we can say that the function is *left continuous* at P_4 because the left-hand limit exists and coincides with the value of the function.

Example 1: Is the function

$$f(x) = \begin{cases} x^2 & x < 1 \\ x^3 + 2 & 1 \leq x \end{cases}$$

continuous at $x = 1$? We certainly notice that the graph of f is broken at $x = 1$ (note the graph below). Note that $f(1) = 3$. $\lim_{x\to 1^+} f(x) = \lim_{x\to 1^+}(x^3 + 2) = 3$. And $\lim_{x\to 1^-} f(x) = \lim_{x\to 1^-}(x^2) = 1$. The limit as $x \to 1$ does not exist since the right-hand and left-hand limits differ, thus the function is not continuous at $x = 1$. We could, however, say that the function is *right continuous* at $x = 1$ since the right-hand limit is equal to the function value (cf Definition 2, below).



Definition 2: A function f is *right continuous* at a point c if it is defined on an interval $[c, d]$ lying to the right of c and if $\lim_{x\to c^+} f(x) = f(c)$. Similarly it is *left continuous* at c if it is defined on an interval $[d, c]$ lying to the left of c and if $\lim_{x\to c^-} f(x) = f(c)$.

In Definition 1, the notion of continuity was defined only at *interior* points of the domain. Definition 2 enables us to speak as well of the continuity of a function at endpoints of its domain. For example we often define functions with restricted domains as with $f(x) = x^2$, $1 \leq x \leq 2$. And we certainly wish to assert that this function is continuous on the closed interval $[1, 2]$. The following definition accomplishes this.

Definition 3: A function f is continuous at a point $x = c$ if c is in the domain of f and:

1. If $x = c$ is an interior point of the domain of f, then $\lim_{x\to c} f(x) = f(c)$.
2. If $x = c$ is not an interior point of the domain but is an endpoint of the domain, then f must be right or left continuous at $x = c$, as appropriate. (Note: Such a point c that is not an interior point but is an endpoint of the domain will correspond to some interval of the form $[c, b)$ or $(b, c]$ that is in the domain.)

Thus, from Definition 3 we can conclude that a function f is continuous on a closed interval $[a, b]$ if it is continuous at every interior point of the interval, is right continuous at a, and is left continuous at b.

Here are several definitions that govern how we speak about the notion of continuity.

Definition 4: A function f is said to be a *continuous function* if it is continuous at every point of its domain.

Definition 5: A *point of discontinuity* of a function f is a point in the domain of f at which the function is not continuous.

Caution: Note that, according to these definitions, a function such as $f(x) = 1/x$ is a continuous function, in spite of the fact that its graph consists of two pieces. At every point where the function *is* defined it is continuous. At the single point $x = 0$ where it is *not* defined continuity is not an issue. In particular we cannot call $x = 0$ a discontinuity of f because it is not in the domain of f. (To say that *f is not continuous at $x = 0$* is not the same as saying *$x = 0$ is a discontinuity of f*. The latter requires that 0 be in the domain of f whereas the former says nothing about whether 0 is in the domain of f.) This is a potential source of confusion, but we will be very consistent in this use of language.

In beginning this section we pointed out that most of the functions we meet in calculus and in applications are continuous wherever they are defined. In particular all polynomials, rational functions, trigonometric functions, the absolute value function, and the exponential and logarithm functions are continuous. Moreover all functions built from these using the operations of addition, subtraction, multiplication, division, composition, and taking inverses are also continuous.

Example 2: All of the following functions are continuous everywhere on their domains:

$$5x^3 - 2x + 1 \qquad \frac{3x+1}{x^2-5} \qquad \sqrt{x+2} \qquad \sin 2\pi x$$

$$(x + \tan x)^{\frac{1}{3}} \qquad \frac{\cos x}{1+\tan x} \qquad |x^2 - 3|$$

Example 3: The rational function $f(x) = \frac{x^2-4}{x-2}$ is a continuous function. The only point not in its domain is $x = 2$. We notice, however, that $\lim_{x\to 2} = 4$ exists. It is as though the function "wants to be continuous" at this point with the value 4. If we oblige by extending the definition of the function we obtain its *continuous extension*

$$F(x) = \begin{cases} f(x) & \text{if } x \text{ is in the domain of } f \\ 4 & \text{if } x = 2 \end{cases}$$

We notice that $F(x)$ is none other than the function $x + 2$ obtained by factoring the numerator of $f(x)$ and simplifying.

Example 4: The function

$$f(x) = \begin{cases} \sin x & x \neq \pi/3 \\ 0 & x = \pi/3 \end{cases}$$

is discontinuous at $\pi/3$. This discontinuity is removable, however, by redefining the value of f at $\pi/3$. The "proper" value would be $\lim_{x\to\pi/3} f(x) = \lim_{x\to\pi/3} \sin x = \sin \pi/3 = \sqrt{3}/2$.

Definition 6: If c is a discontinuity of a function f, and if $\lim_{x\to c} f(x) = L$ exists, then c is called a *removable discontinuity.* The discontinuity is removed by defining $f(c) = L$.

Definition 7: If f is not defined at c but $\lim_{x\to c} f(x) = L$ exists, then f has a *continuous extension* to $x = c$ by defining $f(c) = L$.

The situations in Definitions 6 and 7 are very closely related. They differ only in that $f(c)$ is defined in the first case (but is the "wrong" value) and undefined in the second. In both cases, since the limit exists and has the value L, the lack of continuity at the point is cured by simply giving f the "correct" value L. This is not an idle observation. Our very useful technique for computing limits of the form "$\frac{0}{0}$" by performing some sort of algebraic simplification amounts to finding the continuous extension of a function.

Example 5: Find $\lim_{x\to -2}(x^3+8)/(x+2)$. Solution: rewrite the fraction as

$$\frac{x^3+8}{x+2} = \frac{(x+2)(x^2-2x+4)}{x+2} = x^2 - 2x + 4 \qquad \text{if } x \neq -2$$

The polynomial $x^2 - 2x + 4$ is thus seen to be the *continuous extension* of the given rational function to $x = -2$. The limit of the rational function as $x \to -2$ is the same as the limit of the polynomial as $x \to -2$. And the limit of the polynomial is the same as its *value* at $x = -2$ since it is continuous there. Thus the value of the limit is $(-2)^2 - 2(-2) + 4 = 12$. This is what lay behind our rather peculiar technique of first simplifying the rational expression and then "plugging in" the value $x = -2$ to get the limit.

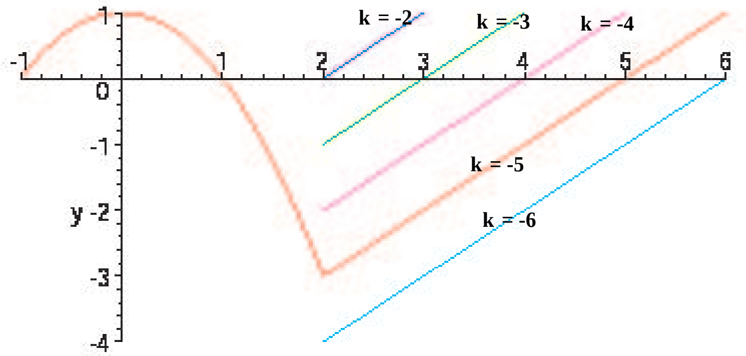
Example 6: Suppose that $f(x)$ is defined piecewise as

$$f(x) = \begin{cases} -x^2 + 1 & x < 2 \\ x + k & x > 2 \end{cases}$$

Let us find a value of the constant k such that f has a continuous extension to $x = 2$. We see that $\lim_{x\to 2^+} f(x) = \lim_{x\to 2^+}(x + k) = 2 + k$, whereas $\lim_{x\to 2^-} f(x) = \lim_{x\to 2^-}(-x^2 + 1) = -3$. Thus we must have $2 + k = -3$, or $k = -5$, for continuity. The desired continuous extension is then

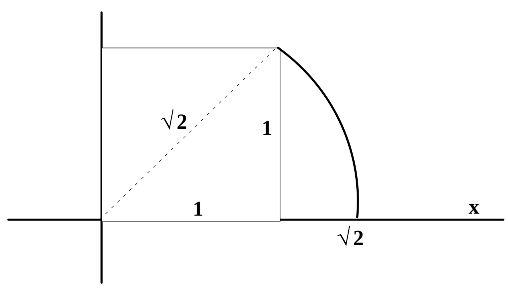
$$F(x) = \begin{cases} -x^2 + 1 & x < 2 \\ x - 5 & x \geq 2 \end{cases}$$

The function f to the left of $x = 2$ is defined to be the quadratic polynomial $y = -x^2 + 1$. To the right it is defined to be a straight line in the family of lines $y = x + k$, k a constant. Determining k so that the right and left hand limits are equal amounts to picking out the particular straight line that meets the parabola at the point $(2, -3)$, as shown in the graph below.

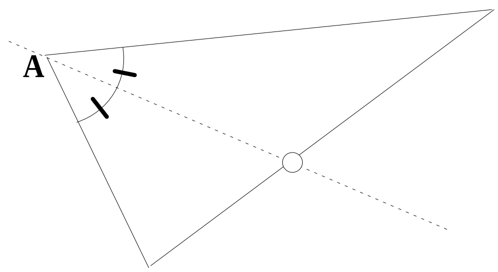


Continuity ... so what? Why such fuss over the notion of continuity? It seems like a perfectly natural property of any self-respecting function. It turns out that nearly all functions that anyone bothers with are continuous. And the only time discontinuity seems to raise its depraved head is when we take the trouble to define a function piecewise in some strange way.

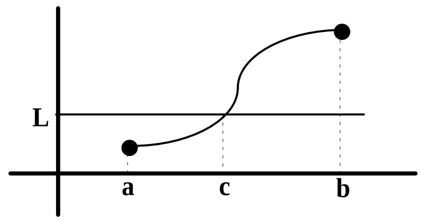
But it was not always so obvious. Pythagorus went bananas when he first noticed that the square root of 2 is not a rational number. That meant to him that a point on the x-axis could easily be constructed geometrically (see the figure below) that could not be "measured" by any number that is a quotient of integers. To the Pythagoreans, who did not believe in irrational numbers, this was a disturbing defect in nature. At issue was the *continuity* of the real line and the lack of richness in their number system for labeling every point on the line. It took humans more than 2000 years to arrive at our present-day notion of *real numbers*, filling in the gaps in the number system so that numbers correspond one-to-one to points on the geometrical line.



And Euclid, no slouch of a mathematician, stumbled over the same issue of continuity. In a number of proofs in his monumental axiomatic treatment of geometry he made the assumption that the bisector of any angle in a triangle must intersect the opposite side of the triangle. How could it possibly miss? (See the figure below.) But it was again thousands of years later when it was definitely shown that Euclid could not possibly have proved this apparently simple fact from his axioms. Again it was the notion of *continuity* that had raised its head. Euclid's axioms were not strong enough to establish the continuity of the number line. The following theorem, stated in modern language, is what enables us today to avoid Euclid's mistake.



The Intermediate Value Theorem: If a function f is continuous on a closed interval $[a,b]$, and if $f(a) < L < f(b)$ (or $f(a) > L > f(b)$), then there exists a point c in the interval $[a,b]$ such that $f(c) = L$.



Can we be forgiven for asking again "How can it possibly miss?" If the graph of f begins below the line $y = L$ and ends up above it, how can it possibly avoid passing through the line? At issue is continuity! The same issue over which Euclid stumbled. The proof of the theorem is subtle. It must show, in effect, that our system of real numbers today is finally rich enough to avoid "gaps" or "jumps" that would allow the graph to get from one side of the line $y = L$ to the other side of it without actually intersecting it. This is what continuity prevents. Within the notion of continuity lurks some two millennia of effort to understand geometry and our number system. This is what we are taking for granted when we say, from our vantage point in the 21st century, that "continuity is a simple and intuitive notion". It is ... but ....

Example 7: Show that the equation $x^5 - 3x + 1 = 0$ has a solution in the interval $[0,1]$. It is worth remarking that we have no algebraic method for solving this equation in exact form, similar to the quadratic formula for solving a second degree equation for example. But consider the function $f(x) = x^5 - 3x + 1$. It is a continuous function, hence continuous on the interval $[0,1]$. And clearly $f(0) = 1$ and $f(1) = -1$. The number $L = 0$ lies between 1 and -1, hence by the Intermediate Value Theorem the graph of f must intersect the line $L = 0$. I.e. there is a point c in the interval where $f(c) = 0$. So there is, indeed, a solution in the interval. Although we have no formula to express the solution, we know that one exists. Thus we can seek to find numerical approximations for the solution. This, indeed, is exactly what a computer program such as Maple does when it "solves" this equation.

Example 8: Does the equation $1/x = 0$ have a solution? Encouraged by Example 8 we may notice that the function $f(x) = 1/x$ has positive values when $x > 0$ and negative values when $x < 0$. In particular $f(-1) = -1$ and $f(1) = 1$. Can we conclude, then that there is a solution in the interval $[-1,1]$? In this case the answer is "no"! The function is not continuous on the interval $[-1,1]$, hence the Intermediate Value Theorem does not apply. And of course we should have known better. There is no real number whose reciprocal is zero!

Summary: Continuity is a simple and intuitive notion. It is precisely defined in terms of the notion of limit. We have seen that most functions of calculus are continuous, and that knowledge of this fact enables us to "turn the tables" and use continuity in the calculation of limits. Finally, the Intermediate Value Theorem (IVT) captures the essence of continuity—a continuous curve "has no gaps". It cannot go from one side of a straight line to the other without intersecting the line. Euclid would have danced with joy! And we can use the IVT as a tool in solving equations.

Applet: Continuity of Functions **Try it!**
Exercises: Problems **Check what you have learned!**
Videos: Tutorial Solutions **See problems worked out!**