Discountinuities for Functions of One and Two Variables

E. L. Lady

(September 5, 1998)

There are three ways that a function can be discontinuous at a point. **(1)** The function can be undefined at the given point, even though it does have a limit there. **(2)** The limit of the function at the given point may not exist. (Note: This includes the case where the limit is $\pm\infty$, since these are not real numbers.) **(3)** The function may be defined and have a limit at the point, but the limit and the value of the function do not agree.

In practice, though, when one talks about discontinuities, one is almost always thinking about case **(2)**. Cases **(1)** and **(3)** simply mean that one has not assigned the "correct" value to the function at the given point. These are called "removeable discontinuities," since one can fix the discontinuity simply by assigning a more appropriate value to the function at the given point.

But case **(2)** means that there's a problem at the given point which cannot be fixed (without drastically changing the function).

Functions of One Variable. One way case **(2)** can occur for a function of one variable is a "jump discontinuity," where the limit from the left and from the right at a given point exist, but do not exist. It is easy to contrive functions with jump discontinuities, but this sort of discontinuity almost never occurs in functions defined by the usual type of formulas found in calculus books. Two examples of functions with jump discontinuities at $x = 0$ are

$$f(x) = \frac{|x|}{x}$$
$$g(x) = \tan^{-1}\left(\frac{1}{x}\right)$$

Here $f(x)$ has a limit of $+1$ as x approaches 0 from the right and -1 as x approaches 0 from the left. And $g(x)$ has a limit of $\pi/2$ as x approaches 0 from the right and a limit of $-\pi/2$ as x approaches 0 from the left.

Usually, though, when one talks about a discontinuity for a function $f(x)$ of one variable, one usually thinks of a place where the function "blows up." In fact, the functions one encounters in calculus courses and most commonly in elementary applications are constructed by combining polynomials, trig functions, logarithms, and the exponential function. For instance,

$$f(x) = \frac{e^{x^3}\sin(4x^2+5)}{x^2-3x+2}.$$

would be a fairly complicated example of such a function. (This function blows up at $x = 1$ and at $x = 2$.) Such functions are "analytic." This means (roughly speaking) that differentiability is so strongly built into the structure of the function that the function is differentiable even if one allows complex numbers as values of x.

Functions with formulas of this kind normally have discontinuities only at points where a denominator becomes 0, or where one is attempting to take the logarithm of 0 or the tangent of $\pi/2$ or doing something else which is clearly bad. At points like this, a function will usually blow up.

In Complex Analysis, a singularity where the function blows up is called a **pole** of the function. To say that a function $f(x)$ has a pole at a point $x = a$ is to say that for values of x very close to a, $|f(x)|$ will be extremely large. However there is another, somewhat less common, type of singularity for a function of a complex variable. This is called an **essential singularity**. As x gets very close to an essential singularity of a function $f(x)$, the values of $f(x)$ will not keep getting larger and larger. Instead, the values will fluctuate through a range of different possibilities, without stabilizing. (Sometimes one cannot see this without looking at complex values for the variable x — something one doesn't learn to do in a basic calculus course.) The most straightforward example of an essential singularity is the behavior of $f(x) = \sin x$ as x "approaches infinity." In fact, as x gets larger and larger, $\sin x$ certainly does not "blow up," since $|\sin x| \le 1$ for all real values of x, even when x is extremely large. Instead, $\sin x$ keeps oscillating over the complete range of values between -1 and $+1$.

Now the concept of a "discontinuity at infinity" does not even arise in a basic calculus course. However it is not too difficult to use the sine function to construct examples of essential singularities at finite points. For instance, one can define

$$f(x) = \sin\left(\frac{1}{x}\right)$$

or

$$g(x) = \sin\left(\frac{x+5}{x-2}\right).$$

The function f has an essential singularity at $x = 0$ and g has an essential singularity at $x = 2$.

To understand this, consider the behavior of $f(x) = \sin(1/x)$, for example, as x gets close and closer to 0. (At $x = 0$ itself, it is, of course, undefined.) The function f does not blow up at 0. In fact, $f(x)$ does not get extremely large since $|\sin(1/x)|$ can never be larger than 1. But what happens is that as x gets closer and closer to 0, the values for $\sin(1/x)$ fluctuate more and more unstably.

For instance, if $x = .001$ then $\sin(1/x) \approx .8269$, but for $x = .00101$, $\sin(1/x) \approx -.477$. (Note that

$$1/.001 = 1,000 \approx 318\pi + .974,$$

so $\sin(1.001) \approx \sin .974 = .8269$. And

$$1/.00101 \approx 990.1 \approx 315\pi + .498,$$

so $\sin 1/.00101 \approx -\sin .498 = -.477$.) For practical purposes, when x is extremely close to 0 one might as well think of $\sin(1/x)$ as a random number between -1 and 1, since it is almost impossible to compute it accurately to even one decimal place because a very slight error in then rounding of x will cause such a large fluctuation. If one now imagines the graph of $\sin(1/x)$ near $x = 0$ (i. e. near the y-axis), one realizes that at a certain point close to the y-axis it will simply start looking like a solid black blur filling the vertical space between $y = -1$ and $y = 1$, and this image can never be resolved, no matter how much one magnifies the graph. Thus $\sin(1/x)$ has no limit at $x = 0$ and therefore is discontinuous there, although it is continuous at all other values of x.

A less conspicuous example of an essential singularity is the function $g(x) = e^{1/x}$. One quickly sees that if one considers only positive values of x, then

$$\lim_{x \to +0} e^{1/x} = e^{\infty} = \infty.$$

However, technically $g(x)$ does not have a pole at 0, since by looking at negative x one gets

$$\lim_{x \to -0} e^{1/x} = e^{-\infty} = 1/e^{\infty} = 0.$$

In fact, if one considers complex numbers as values for x, then it can be seen that the behavior of $e^{1/x}$ at 0 is very similar to the behavior of $\sin 1/x$.

The very weirdness of $f(x) = \sin(1/x)$ suggests that it is not the kind of function which one would expect to arise in applications. In practice, for functions of one variable one will not often be mistaken if one interprets the term “discontinuity” as a point where a function blows up.

Functions of Two Variables. A function $f(x,y)$ of two or more variables (if given by a reasonable formula) will also usually also be continuous except at points where a denominator becomes zero or something else happens which is clearly illegal. And just as in the case of one variable, a function clearly blows up at points where the denominator is zero but the numerator is not. The only problematic limits are the ones where a numerator and denominator both approach zero, or one gets some other indeterminate form.

What makes the theory of functions of several variables different is that when a discontinuity does occur, it is quite common for the function not to blow up but instead to behave rather weirdly.

The classic example of this is the behavior of the function

$$f(x,y) = \frac{2xy}{x^2+y^2}$$

at the point $(x,y) = (0,0)$. Since the denominator is 0, there is the possibility of a discontinuity, but since the numerator is also zero, it is not obvious what will happen.

Now the interesting (and misleading) thing here is that if one holds y fixed, then the function

$$g(x) = \frac{2xy}{x^2+y^2} \qquad (y \text{ fixed})$$

is a continuous function of x. This is clear if y is not zero, since in that case the denominator never becomes 0. But it is also true when $y=0$, because in that case, $g(x)=0$ for $x \neq 0$, so that $x=0$ is a "removable discontinuity" for g which can be eliminated by giving $g(0)$ the value 0 (which amounts to giving $f(0,0)$ the value 0). Likewise, if one holds x fixed, the function

$$h(y) = \frac{2xy}{x^2+y^2} \qquad (x \text{ fixed})$$

is a continuous function of y with $h(0)=0$. So it would seem pretty clear that

$$\lim_{(x,y)\to(0,0)} \frac{2xy}{x^2+y^2} = 0 \qquad (?)$$

This is, however, incorrect. There are, in fact, points (x,y) arbitrarily close to $(0,0)$ where $f(x,y)$ is not at all close to 0. Consider, for instance, a point (x,x) with x non-zero but extremely small. One sees that

$$f(x,x) = \frac{2x^2}{x^2+x^2} = 1,$$

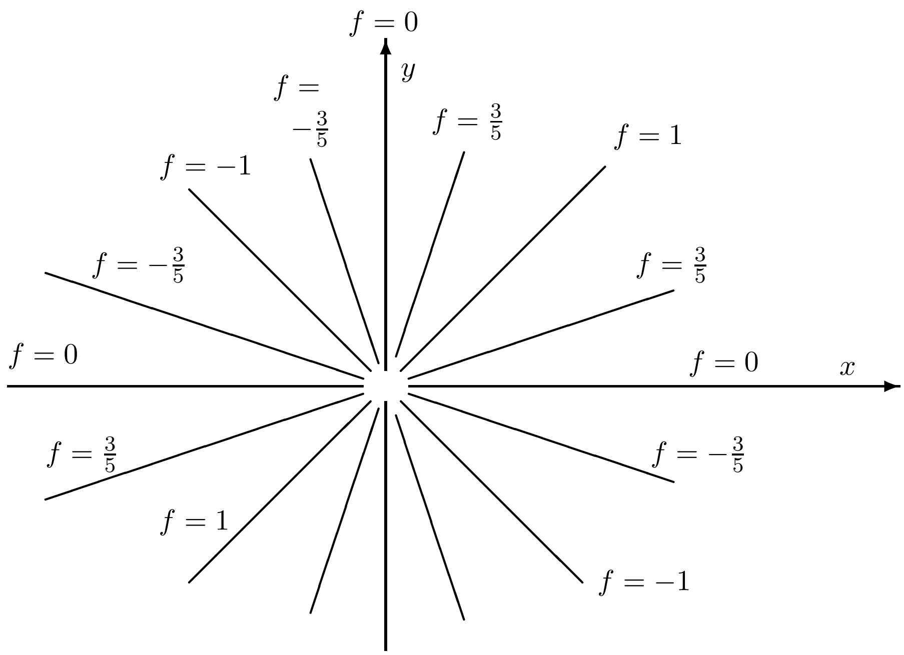
no matter how small x is.

In fact, $f(x,y)$ is constant on every straight line going through the origin. Consider a line $y = mx$. Substituting any point on this line into f, we get

$$f(x,mx) = \frac{2mx^2}{x^2 + mx^2} = \frac{2m}{1+m^2},$$

a value which is constant along the whole line. Indeed, it turns out that the lines through the origin are all level curves for $f(x,y)$. (Well, almost so, anyway. It can be seen that a level curve is actually the union of the two different lines $y = mx$ and $y = x/m$, since

$$f(x, \frac{x}{m}) = \frac{2x^2/m}{x^2 + \dfrac{x^2}{m^2}} = \frac{2/m}{1 + \dfrac{1}{m^2}} = \frac{2m}{m^2+1} = f(x,mx).)$$



It doesn't really seem right that the different level curves for a function should intersect. In fact, if one thinks a bit, one realizes that if two different level curves for a function intersect at a point, that would indicate that the function would have to have two different values at that particular point. Since a function can have only one value at a given point, what this means is that anywhere two different level curves of a function intersect has to be a discontinuity for that function.

Now the function $f(x,y)$ described above has a discontinuity at $(0,0)$ but does not blow up there. In fact, it is easy to see that $f(x,y)$ actually never gets larger than 1. To see this, consider an arbitrary point (x,y). This point certainly belongs to some straight line going through the origin (namely, the line connecting the origin to the point in question). And we have seen above that if m is the slope of that line, then $f(x,y) = f(x,mx) = 2m/(1+m^2)$. (If the line is vertical, then $x=0$ and $f(x,y) = f(0,y) = 0$.) And it's not too hard to show that the function $2m/(1+m^2)$ always takes values between -1 and $+1$.

This function $2xy/(x^2+y^2)$ is a lot stranger than it looks at first glance. To understand it better, and to better be able to see what the discontinuity at (0,0) looks like, it is useful to change to polar coordinates: $x = r\cos\theta$, $y = r\sin\theta$. One then gets

$$f(x,y) = \frac{2xy}{x^2+y^2} = \frac{2r^2\cos\theta\sin\theta}{r^2} = 2\sin\theta\cos\theta = \sin 2\theta.$$

This makes several things more clear. For one thing, it is now obvious that $|f(x,y)| \leq 1$ everywhere, as claimed above. Furthermore, since the value of f at a point (x,y) depends only on the argument θ, this makes it obvious why the function is constant on straight lines through the origin. Furthermore, it now doesn't take too much imagination to roughly visualize the graph $z = f(x,y)$. Namely, the graph looks somewhat like a circular staircase, except that instead of a staircase it's a ramp, and instead of steadily climbing up, its elevation moves up and down in something like a wave as it circles the z-axis. Furthermore, at the z-axis all the stairs collide (as it were), causing the discontinuity. (This is in fact the reason that circular staircases are always built with a hole in the middle.)

More generally, consider any function defined by a fraction

$$f(x,y) = \frac{p(x,y)}{q(x,y)},$$

where $p(x,y)$ and $q(x,y)$ are homogeneous polynomials of the same degree. In other words, each term in the numerator and the denominator has the same degree. An example would be

$$f(x,y) = \frac{x^3 - x^2y + 9xy^2}{5x^2y + 7y^3}.$$

Here, each term in the numerator and denominator has degree 3. Such a function can be algebraically rewritten so that it becomes a function of $\dfrac{y}{x}$:

$$f(x,y) = \frac{1 - \frac{y}{x} + 9(\frac{y}{x})^2}{5\frac{y}{x} + 7(\frac{y}{x})^3}.$$

This shows that the function takes a constant value along any line through the origin, since the value depends only on the slope $\dfrac{y}{x}$.

Since the origin belongs to all these lines, and since the function can have only one value at the origin, we see that there must be a discontinuity at the origin (except in the case where $f(x,y)$ is a constant function).