Derivatives and Differentials

Randy S

Abstract Many of the articles in this series assume familiarity with basic calculus. This article gives a brief review of some of the concepts, definitions, and theorems that are assumed to be familiar. This includes derivatives, partial derivatives, and differentials.

Contents

1 Derivative of a function: the concept

Consider a function f whose input is a real number x and whose output is a real number $f(x)$, such as the function depicted below. We can define a new function f' whose output $f'(x)$ is the *slope* of the graph of f at the horizontal position x. This new function f' is called the **derivative** of f.

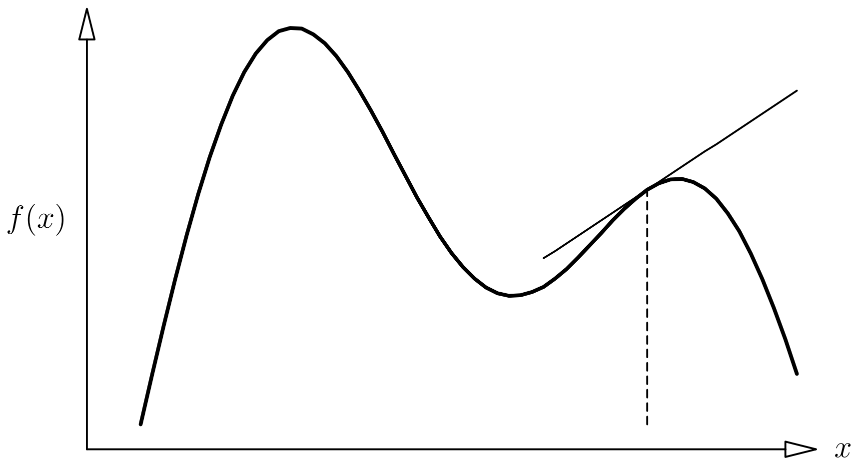

Figure 1 – The curved line is the graph of a function $f(x)$. The height of the dotted vertical line is the value of the function at a particular value of x. The slope of the tangent line is the derivative of the function at that value of x.

2 Derivative of a function: the definition

Figure 1 defines the derivative conceptually, but how can we express that definition mathematically? If the graph of f were a straight line, then we could choose any two points x and $x+\varepsilon$ and compute the slope using "rise over run":

$$\frac{f(x+\varepsilon)-f(x)}{(x+\varepsilon)-x}, \tag{1}$$

which may be simplified to

$$\frac{f(x+\varepsilon)-f(x)}{\varepsilon}. \tag{2}$$

If the graph of f is not a straight line, then the slope may be different at every point. In that case, equation (2) is only an approximation to the slope at x, but if the function is **smooth**, then the approximation can be made arbitrarily good by making ε arbitrarily small. The derivative is defined by taking the *limit* as $\varepsilon \to 0$. In symbols,[1]

$$\boxed{f'(x) \equiv \lim_{\varepsilon\to 0}\frac{f(x+\varepsilon)-f(x)}{\varepsilon}.} \tag{3}$$

If the ratio (2) does not approach a finite limit as ε approaches zero, then the derivative is undefined, and the function is not smooth at that point.

Equation (3) defines the **derivative** of f with respect to x. In this context, x is called the **independent variable**.

[1]The triple equals symbol "≡" means that the abbreviation on one side is defined by the expression on the other side.

3 Common examples

Here's a short list of the derivatives of some simple functions:[2]

- If $f(x) = x^n$ and $n > 0$, then $f'(x) = nx^{n-1}$.
- For $x \neq 0$: if $f(x) = x^n$ and $n < 0$, then $f'(x) = nx^{n-1}$.
- If $f(x) = \exp(x) \equiv e^x$, then $f'(x) = f(x)$.
- For $x > 0$: if $f(x) = \log(x)$, then $f'(x) = 1/x$.
- If $f(x) = \sin(x)$, then $f'(x) = \cos(x)$.
- If $f(x) = \cos(x)$, then $f'(x) = -\sin(x)$.

[2] I'm using $\log(x)$ to denote the natural log of x, so $\log(e^x) = x$. Sometimes the natural log is denoted "ln," but I prefer the notation "log" because it's easier to read. Sometimes engineers use "log" to denote the base-10 logarithm, but that's rarely used in theoretical physics, and when we do need to use it we can write it as $\log_{10}$ instead.

4 Useful properties

In this section, f' denotes the derivative of a function with respect to its input. The following properties are consequences of the definition (3). The first two properties say that the derivative operator is **linear**:

- If $h(x) = f(x) + g(x)$, then $h'(x) = f'(x) + g'(x)$.
- If $h(x) = c\,f(x)$ for some constant c, then $h'(x) = c\,f'(x)$.

The next property is the **product rule**, also called the **Leibniz rule**:

If $h = fg$, then $h' = f'g + g'f$.

The next property is the **chain rule**:

If $h(x) = f\big(g(x)\big)$, then $h'(x) = f'\big(g(x)\big)g'(x)$.

Section 6 shows another way of writing the chain rule that is easier to remember.

5 Various notations

The previous sections used the symbol f' to denote the derivative of f. The tick-mark is often called a "prime" (this has nothing to do with prime numbers), so f' is pronounced f-prime. This notation is *not* universal: sometimes the same notation f' is used for different things, and sometimes other notations are used for the derivative of f. Another common notation for the derivative of f is $\dot{f}$, especially when time is the independent variable.

Another way of writing the derivative of f is

$$\frac{df}{dx}(x) \qquad \text{or just} \qquad \frac{df}{dx}, \tag{4}$$

where x is the same symbol that is used for the input to f. This is called the **Leibniz** notation. This notation reflects the intuition that the derivative is (the limit of) a ratio, as in (1), with

$$\begin{aligned} df &= f(x+\varepsilon) - f(x) \\ dx &= (x+\varepsilon) - x. \end{aligned}$$

The prefix "d" stands for **differential** (section 9).

The derivative of f is also commonly written

$$\frac{d}{dx} f(x). \tag{5}$$

This notation is similar to the notation (4), but here the idea is that d/dx is an *operator* that we can apply to any function $f(x)$ to obtain its derivative. This kind of operator is called a **differential operator**. This notation is especially useful when we don't want to bother giving the function a name. For example, the derivative of the anonymous function x^3 is

$$\frac{d}{dx} x^3 = 3x^2.$$

6 The chain rule again

As before, let f' denote the derivative of a function with respect to its input. In particular, given a composite function $f(g(x))$, the notation $f'(g(x))$ means the derivative of f with respect to g, not with respect to x. This notation was used in section 4 to express the chain rule: if $h(x) = f\big(g(x)\big)$, then $h'(x) = f'\big(g(x)\big)g'(x)$.

The chain rule can be written in a nice way using the Leibniz notation: The derivative of the nested function $f\big(g(x)\big)$ with respect to the input x is

$$\frac{df}{dx} = \frac{df}{dg}\,\frac{dg}{dx}.$$

Rearranging this gives another useful form of the identity:

$$\frac{df}{dg} = \frac{df/dx}{dg/dx}.$$

The factor df/dg indicates the derivative of f with respect to its input, which is g. The nice thing about this notation is that it looks like a simple identity with fractions. The "factor" dg occurs in both the numerator and in the denominator, so it "cancels". Derivatives aren't really ordinary fractions (because the definition (3) involves taking a limit), but the chain rule works as though they were. This makes it easy to remember.

7 Analytic functions and Taylor series

So far, we've only considered the **first derivative** of a function. Since the result is another function, we can iterate this. The derivative of a derivative is called a **second derivative**, and so on. The second derivative of $f(x)$ is often denoted

$$\frac{d}{dx}\frac{d}{dx}f(x) \qquad \text{or} \qquad \left(\frac{d}{dx}\right)^2 f(x) \qquad \text{or} \qquad \frac{d^2 f}{dx^2}(x).$$

Example: for $n \geq 2$, the second derivative of x^n is

$$\frac{d^2}{dx^2}x^n = \frac{d}{dx}\left(\frac{d}{dx}x^n\right) = \frac{d}{dx}nx^{n-1} = (n-1)nx^{n-2}.$$

The notations f'' and $\ddot{f}$ are also common, with one prime per derivative or one dot per derivative.

A function $f(x)$ is called **analytic** at x_0 if the series

$$f(x_0) + f'(x_0)(x-x_0) + f''(x_0)\frac{(x-x_0)^2}{2!} + f'''(x_0)\frac{(x-x_0)^3}{3!} + \cdots \tag{6}$$

converges to $f(x)$ for all x in some nonempty neighborhood of x_0. This is called a **Taylor series**. The simplest example of an analytic function is a polynomial, for which the Taylor series terminates after a finite number of terms because all of the higher derivatives are zero. Even if the function isn't a polynomial, we can artificially truncate its Taylor series to obtain a polynomial that gives a good approximation to the original function f when x is sufficiently close to x_0.

8 Partial derivatives

If $f(x_1, x_2, ..., x_N)$ is a function of two or more independent variables, then the **partial derivative** of f with respect to x_n, denoted

$$\frac{\partial f}{\partial x_n}$$

is simply the derivative of f with respect to x_n with all of the other indepenent variables treated as constants. As an example, consider the function

$$f(x,y) = x^2 + y^3 x. \tag{7}$$

The partial derivatives of this function are

$$\frac{\partial f}{\partial x}(x,y) = 2x + y^3 \qquad\qquad \frac{\partial f}{\partial y}(x,y) = 3y^2 x.$$

The definition of the partial derivative with respect to one variable depends on which set of variables we're treating as independent. To illustrate this, consider the function (7) again. Instead of using x and y as the independent variables, we could use x and $u \equiv y - x$ as the independent variables. The important message is that the definition of $\partial/\partial x$ depends on which set we use: in one case we treat y as a constant, and in the other case we treat u as a constant. Explicitly:

$$\begin{aligned}\left[\frac{\partial}{\partial x}(x^2 + y^3 x)\right]_{y=u+x} &= 2x + (u+x)^3 \\ \frac{\partial}{\partial x}\big(x^2 + (u+x)^3 x\big) &= 2x + (u+x)^3 + 3(u+x)^2 x.\end{aligned}$$

On the first line, the independent variables are x and y when the derivative $\partial/\partial x$ is applied. On the second line, the independent variables are x and u instead. The answers are different, even though both answers are eventually written in terms of x and u. Both answers are *correct*, but they are answering different questions.

9 Differentials

Let $f(x,y)$ be a smooth function of x and y, and let $x(s), y(s)$ be smooth functions of s. Then we can define a function

$$f\big(x(s),\, y(s)\big).$$

According to the chain rule (section 4), the derivative of this function with respect to s is

$$\frac{d}{ds}f\big(x(s),\, y(s)\big) = \frac{\partial f}{\partial x}\frac{dx}{ds} + \frac{\partial f}{\partial y}\frac{dy}{ds}. \tag{8}$$

Intuitively, the functions $x(s), y(s)$ describe a **curve** in the x-y plane **parameterized by** s, and we can think of (8) as the derivative of f along that parameterized curve. If we pretend for a moment that a derivative were simply a ratio, then we could cancel the factors of ds to get

$$df = \frac{\partial f}{\partial x}\, dx + \frac{\partial f}{\partial y}\, dy. \tag{9}$$

We can think of (9) as a way of writing (8) without committing to a particular parameterized curve. The formal quantities df, dx, and dy are called the **differentials** of f, x, and y, respectively.

We can also recover this intuitive picture from the more sophisticated concept of a **differential form**, which uses the same notation (9). Differential forms are introduced in article 09894.

10 References in this series

Article **09894** (https://cphysics.org/article/09894):
"Tensor Fields on Smooth Manifolds" (version 2023-05-08)