FOUNDATION OF SCIENTIFIC COMMUNICATION:
SLICE OF PIE

Different Derivatives for Physics

Author: *Remko Osseweijer*

Course: *Derivatives*
Date: *January 9, 2026*

1. Introduction

This short paper aims to provide an overview of the concept of differentiation of functions in one and more dimensions. This is by no means a sufficient introduction to the mathematical field of analysis, but it suffices for a rapid familiarisation with the topic at hand in the context of physical systems.

It begins to formally formulate the definition of a derivative and their rules, only to extend this concept to various generalisations. Sufficient physical examples are provided to enhance the discussion and engage the reader's mind. The paper ends by considering derivatives in various fields in physics, among which quantum mechanics and general relativity.

2. Normal derivative

2.1. Formal definition

Consider a continuous function $f : \mathbb{R} \to \mathbb{R}$ that takes a real number and outputs another real number, i.e. $f(x) = y$ for $x, y \in \mathbb{R}$. We recall (one of the many) definitions of continuity for the reader's sense of completeness.

Definition 2.1 (Continuity)**.** *Let $f : \mathbb{R} \to \mathbb{R}$ be a function. Then f is said to be* ***continuous at a point*** *$c \in \mathbb{R}$ if the following three conditions are satisfied.*

1. *The value $f(c) \in \mathbb{R}$ is well-defined;*
2. *The limit $\lim_{x\to c} f(x)$ exists, which is to say that $\lim_{x\to c^+} f(x) = \lim_{x\to c^-} f(x)$ holds;*
3. *The limit equates as $\lim_{x\to c} f(x) = f(c)$.*

We say that the function $f : \mathbb{R} \to \mathbb{R}$ is ***continuous*** *if f is continuous at all points $c \in \mathbb{R}$.*

Informally, this definition tells us that a function is continuous if you can draw its graph without lifting your pencil, ensuring that there are no gaps or sudden jumps. Admittedly, one would need to formally introduce the notion of a limit, but it is assumed that the intuition behind a limit is sufficient to forgo the formal definition. You may think of a limit $\lim_{x\to c} f(x)$ as infinitesimally approaching the value $f(c)$ (if defined) by sending x to c, but never quite reaching it. There is always a nonzero difference between $\lim_{x\to c} f(x)$ and $f(c)$, though this difference keeps growing smaller and smaller.

Having discussed what continuity is, let us consider such a continuous function $f : \mathbb{R} \to \mathbb{R}$, and let $x_1, x_2 \in \mathbb{R}$ be two real numbers. We can describe the slope $m \in \mathbb{R}$ between the two points x_1 and x_2 as the ratio of how much quicker the function grows compared to the original separation, namely

$$m = \frac{f(x_2) - f(x_1)}{x_2 - x_1} = \frac{y_2 - y_1}{x_2 - x_1} = \frac{\Delta y}{\Delta x},$$

where we define $f(x_1) = y_1$ and $f(x_2) = y_2$, and $\Delta y = y_2 - y_1$ and $\Delta x = x_2 - x_1$, to recast the slope into a familiar form. This slope tells us how the function increases or decreases between two points; it being quantified by a straight line with slope m.

But what if we are interested in the instantaneous change at some point, say at $x \in \mathbb{R}$? How does $f(x)$ change when we stay infinitesimally close to x? Evidently, we run head-first into a problem, namely that the denominator of the definition of the slope approaches zero! This is catastrophic, as the slope will approach infinity in value. Our only hope is that $f(x_2) - f(x_1)$ counteracts the divergence caused by $x_2 - x_1$, yielding a finite value for the slope m. This is precisely what motivates derivatives. We provide a formal definition.

Definition 2.2 (Differentiability)**.** *Let $f : \mathbb{R} \to \mathbb{R}$ be a continuous function. We say that f is **differentiable at** $a \in \mathbb{R}$ if the limit*

$$\lim_{h \to 0} \frac{f(a+h) - f(a)}{h}$$

*exists. The value of this limit is called the **derivative of f at** a, and we denote it as $f'(a)$. We say that f is **differentiable (everywhere)** when f is differentiable for every $x \in \mathbb{R}$ and $f'(x)$ is called the **derivative of** f.*

The expression for the derivative follows immediately from considering the slope of the function between the points a and $a + h$ where we then take the limit of sending h to zero. Incidentally, derivatives describe a notion of *change* of the steepness of the original function.

If the derivative $f' : \mathbb{R} \to \mathbb{R}$ of a continuous function $f : \mathbb{R} \to \mathbb{R}$ is itself continuous, then we could reapply the definition of the derivative, which, if well-defined, will give the derivative f'' of the derivative f', also called the **second (order) derivative** f'' of the function f. We then say that f is **twice differentiable**. We could repeat this procedure to obtain higher order derivatives. If we can proceed indefinitely, we say that f is **infinitely differentiable** or **smooth**, as it will be continuous and differentiable no matter how many times we take the derivative.

Different notations for the derivative of a function $f : \mathbb{R} \to \mathbb{R}$ at a point $a \in \mathbb{R}$ aside from the standard $f'(a)$ include $\left.\frac{\mathrm{d}f(x)}{\mathrm{d}x}\right|_{x=a}$ or $\frac{\mathrm{d}f(x)}{\mathrm{d}x}(a)$ (seen a lot in Taylor expansions), $f^{(1)}(a)$ (generalises easily to higher order derivatives), $\dot{f}(a)$ (usually when considering time), $f_x(a)$ (useful when considering multiple variables), and various types using $Df(a)$ or $D_x f(a)$. One could exhaust a lifetime studying the different types of notation used for derivatives.

Example 2.3. In physics, we describe the position of a box on a one-dimensional line using the coordinate $x \in \mathbb{R}$. The velocity of the box is defined as the instantaneous change of position with respect to time, thus defined as the derivative $\frac{\mathrm{d}x}{\mathrm{d}t}$ or $\dot{x}$. The acceleration is defined as the change of velocity with respect to time, $\frac{\mathrm{d}v}{\mathrm{d}t}$, which can be expressed using the definition of velocity as $a = \frac{\mathrm{d}v}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}\frac{\mathrm{d}x}{\mathrm{d}t} = \frac{\mathrm{d}^2x}{\mathrm{d}t^2}$, or simply $a = \ddot{x}$.

2.2. Properties

In a physics major, you see the definition of the derivative once and only once. Fortunately for us, there are differentiation rules that significantly ease the process of taking derivatives. We will briefly cover these rules here, giving the reader the opportunity to prove them all using the limit definition of the derivative.

[properties]

3. Higher-Dimensional derivatives

In the previous section, we strictly considered a continuous function $f : \mathbb{R} \to \mathbb{R}$. We would like to extend our formalism to deal with real vector-valued functions $\boldsymbol{f} : \mathbb{R} \to \mathbb{R}^n$. The formal definition of the derivative of a vector-valued function $\boldsymbol{f}$ is eerily similar, namely,

$$\boldsymbol{f}'(x) = \lim_{h \to 0} \frac{\boldsymbol{f}(x-h) - \boldsymbol{f}(x)}{h}.$$

The fun part is writing it out. A vector-valued function $\boldsymbol{f} : \mathbb{R}^n \to \mathbb{R}$ can be explicitly written as a vector, unsurprisingly, with n real-valued functions $f_i : \mathbb{R} \to \mathbb{R}$ as components,

$$\boldsymbol{f}(x) = \begin{pmatrix} f_1(x) \\ f_2(x) \\ \vdots \\ f_n(x) \end{pmatrix}.$$

The derivative follows then component-wise,

$$\boldsymbol{f}'(x) = \begin{pmatrix} \frac{\partial f_1(x)}{\partial x} \\ \vdots \\ \frac{\partial f_n(x)}{\partial x} \end{pmatrix}.$$

Thus, taking the derivative of a vector-valued function just requires us to apply our knowledge of derivatives of real-valued functions n-fold. There is not really much news here to uncover.

Example 3.1. .

4. Partial derivatives

Having considered real-valued functions $f : \mathbb{R} \to \mathbb{R}$ and real vector-valued functions $\boldsymbol{f} : \mathbb{R} \to \mathbb{R}^n$, it is but natural to consider real multivariable functions $f : \mathbb{R}^n \to \mathbb{R}$, or $f(x, y, z, \dots)$ where we now explicitly denote the input variables. In this case, it is quite straightforward to define a derivative for a particular variable. We define the **partial**

derivatives,

$$\begin{aligned}\frac{\partial f}{\partial x} &= \lim_{h\to 0}\frac{f(x+h,y,z,\dots)-f(x,y,z,\dots)}{h},\\ \frac{\partial f}{\partial y} &= \lim_{h\to 0}\frac{f(x,y+h,z,\dots)-f(x,y,z,\dots)}{h},\\ \frac{\partial f}{\partial z} &= \lim_{h\to 0}\frac{f(x,y,z+h,\dots)-f(x,y,z,\dots)}{h},\dots.\end{aligned}$$

The partial derivatives measure the change in a direction while holding all the other variables constant. In some sense, you can simply ignore all the other variables while taking the partial derivative, and you put them back later.

The fun part comes when considering more than one partial derivatives, more particularly when we try mixing them, for example,

$$\frac{\partial^2 f(x,y,z,\dots)}{\partial x^2},\frac{\partial^2 f(x,y,z,\dots)}{\partial x\partial y},\frac{\partial^2 f(x,y,z,\dots)}{\partial y\partial x},\dots.$$

Example 4.1. .

4.1. Gradient

Let us consider a multivariable function $f:\mathbb{R}^n\to\mathbb{R}$ and explicitly write out the variables as $f(x_1,\dots,x_n)$. The partial derivatives are then given by $\frac{\partial f}{\partial x_1},\frac{\partial f}{\partial x_2},\dots,\frac{\partial f}{\partial x_n}$. We could suggestively group these in vector notation as,

$$\begin{pmatrix}\frac{\partial f}{\partial x_1}\\ \frac{\partial f}{\partial x_2}\\ \vdots\\ \frac{\partial f}{\partial x_n}\end{pmatrix}.$$

The previous observation actually motivates us to define the following operator,

$$\nabla=\begin{pmatrix}\frac{\partial}{\partial x_1}\\ \frac{\partial}{\partial x_2}\\ \vdots\\ \frac{\partial}{\partial x_n}\end{pmatrix}.$$

This object is often called the **del** or the **nabla**; we will stick with the latter. It looks like a vector, but formally it needs a function to operate on. Letting ∇ act on our function $f:\mathbb{R}^n\to\mathbb{R}$ gives the required,

$$\nabla f=\begin{pmatrix}\frac{\partial}{\partial x_1}\\ \frac{\partial}{\partial x_2}\\ \vdots\\ \frac{\partial}{\partial x_n}\end{pmatrix}f=\begin{pmatrix}\frac{\partial f}{\partial x_1}\\ \frac{\partial f}{\partial x_2}\\ \vdots\\ \frac{\partial f}{\partial x_n}\end{pmatrix}.$$

The combination ∇f is called the **gradient** of f and it gives us a short-hand way of grouping partial derivatives together. It is a vector-valued function representing the derivatives in all the directions of the function f.

If we pursue our self-fed delusion of nabla being a vector, we can play with operations on vectors. For this, we need a vector-valued function with multiple variables.

4.2. Divergence

Consider a vector-valued multivariable function $\boldsymbol{f} : \mathbb{R}^n \to \mathbb{R}^n$ that can be explicitly written as,

$$\boldsymbol{f}(x_1, \ldots, x_n) = \begin{pmatrix} f_1(x_1, \ldots, x_n) \\ f_2(x_1, \ldots, x_n) \\ \vdots \\ f_n(x_1, \ldots, x_n) \end{pmatrix}.$$

We can consider the inner product of nabla and this vector-valued function. Recall that the inner product of two vectors $\boldsymbol{A}, \boldsymbol{B}$ is given by,

$$\boldsymbol{A} \cdot \boldsymbol{B} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \cdot \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = a_1 b_1 + a_2 b_2 + \ldots a_n b_n = \sum_{i=1}^{n} a_i b_i.$$

The inner product of ∇ and $\boldsymbol{f}$ is thus given by

$$\begin{aligned} \nabla \cdot \boldsymbol{f} &= \begin{pmatrix} \frac{\partial}{\partial x_1} \\ \frac{\partial}{\partial x_2} \\ \vdots \\ \frac{\partial}{\partial x_n} \end{pmatrix} \cdot \begin{pmatrix} f_1(x_1, \ldots, x_n) \\ f_2(x_1, \ldots, x_n) \\ \vdots \\ f_n(x_1, \ldots, x_n) \end{pmatrix} = \frac{\partial f_1(x_1, \ldots, x_n)}{\partial x_1} + \cdots + \frac{\partial f_n(x_1, \ldots, x_n)}{\partial x_n} \\ &= \sum_{i=1}^{n} \frac{\partial f_i(x_1, \ldots, x_n)}{\partial x_i}. \end{aligned}$$

The inner product $\nabla \cdot \boldsymbol{f}$ is called the **divergence** of $\boldsymbol{f}$. It is the sum of the partial derivatives and is thus a real-valued function of the coordinates. What does it mean? Well, at any point in your space, it tells you whether more derivatives are toward the point or away from the point, which is quantised by the sign of $\nabla \cdot \boldsymbol{f}$. In other words, if the divergence is negative, the function behaves like a sink around the point, guiding the surrounding points towards the original point if the system were to be dynamic. If the divergence is instead positive, the function behaves as a source, pushing the surrounding points away from the original point. In essence, the divergence describes whether a point gives or takes from the vector field.

4.3. Curl

Another operation we can induce on vectors is the cross-product. For this, we can only consider three-dimensional (or seven-dimensional, but that is for another day) functions, $f : \mathbb{R}^3 \to \mathbb{R}^3$, which is explicitly written as,

$$\boldsymbol{f}(x,y,z) = \begin{pmatrix} f_x(x,y,z) \\ f_y(x,y,z) \\ f_z(x,y,z) \end{pmatrix}.$$

The cross-product between two vectors $\boldsymbol{A}$ and $\boldsymbol{B}$ is defined as

$$\boldsymbol{A} \times \boldsymbol{B} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \times \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} a_2 b_3 - a_3 b_2 \\ a_3 b_1 - a_1 b_3 \\ a_1 b_2 - a_2 b_1 \end{pmatrix}.$$

Using the cross-product, we can act with nabla on the three-dimensional vector-valued function $\boldsymbol{f}$ to obtain,

$$\nabla \times \boldsymbol{f} = \begin{pmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \\ \frac{\partial}{\partial z} \end{pmatrix} \times \begin{pmatrix} f_x(x,y,z) \\ f_y(x,y,z) \\ f_z(x,y,z) \end{pmatrix} = \begin{pmatrix} \frac{\partial f_y}{\partial z} - \frac{\partial f_z}{\partial y} \\ \frac{\partial f_z}{\partial x} - \frac{\partial f_x}{\partial z} \\ \frac{\partial f_x}{\partial y} - \frac{\partial f_y}{\partial x} \end{pmatrix}.$$

[]

4.4. Laplacian

4.5. d'Alembertian

4.6. Maxwell equations

5. Quantum derivatives

6. Covariant derivatives

In general relativity, it turns out that partial derivatives are not sufficient to describe changes while respecting the structure of spacetime. To see why, we briefly need to define a tensor as an object that transforms as a tensor. More specifically, a tensor is an object $T^{\mu_1 \dots \mu_k}{}_{\nu_1 \dots \nu_l}$ that transforms under a coordinate transformation $x^\mu \to x^{\mu'}$ as,

$$T^{\mu_1' \dots \mu_k'}{}_{\nu_1' \dots \nu_l'} = \frac{\partial x^{\mu_1'}}{\partial x^{\mu_1}} \cdots \frac{\partial x^{\mu_k'}}{\partial x^{\mu_k}} \frac{\partial x^{\nu_1}}{\partial x^{\nu_1'}} \cdots \frac{\partial x^{\nu_l}}{\partial x^{\nu_l'}} T^{\mu_1 \dots \mu_k}{}_{\nu_1 \dots \nu_l}.$$

Here, we use the Einstein convention, in which any repeated indices are summed over (in this case, the unprimed ones). In any equation, you can easily notice the mathematical validity of said equation by comparing free indices on either side (in this case, the primed ones).

Anything that transforms in this way is called a tensor. In particular, for a vector V^μ, we have the simpler transformation law, namely that it is an object that transforms as a vector,

$$V^{\mu'} = \frac{\partial x^{\mu'}}{\partial x^\mu} V^\mu.$$

For the partial derivative ∂_μ, we have the well-known chain rule under which it transforms,

$$\partial_{\mu'} = \frac{\partial}{\partial x^{\mu'}} = \frac{\partial x^\mu}{\partial x^{\mu'}} \frac{\partial}{\partial x^\mu} = \frac{\partial x^\mu}{\partial x^{\mu'}} \partial_\mu.$$

We can now consider the partial derivative acting on a vector $\partial_\mu V^\nu$. The transformation for this can be calculated readily,

$$\begin{aligned} \partial_{\mu'} V^{\nu'} = \frac{\partial}{\partial x^{\mu'}} V^{\nu'} &= \frac{\partial x^\mu}{\partial x^{\mu'}} \frac{\partial}{\partial x^\mu} \left(\frac{\partial x^{\nu'}}{\partial x^\nu} V^\nu \right) \\ &= \frac{\partial x^\mu}{\partial x^{\mu'}} \frac{\partial x^{\nu'}}{\partial x^\nu} \frac{\partial}{\partial x^\mu} V^\nu + \frac{\partial x^\mu}{\partial x^{\mu'}} \frac{\partial^2 x^{\nu'}}{\partial x^\nu \partial x^\mu} V^\nu \\ &= \frac{\partial x^\mu}{\partial x^{\mu'}} \frac{\partial x^{\nu'}}{\partial x^\nu} \partial_\mu V^\nu + \frac{\partial x^\mu}{\partial x^{\mu'}} \frac{\partial^2 x^{\nu'}}{\partial x^\nu \partial x^\mu} V^\nu. \end{aligned}$$

We used the product rule in the third equality. Notice that

$$\partial_{\mu'} V^{\nu'} \neq \frac{\partial x^\mu}{\partial x^{\mu'}} \frac{\partial x^{\nu'}}{\partial x^\nu} \partial_\mu V^\nu$$

because of the additional term, which means that $\partial_\mu V^\nu$ does not transform like a tensor and thus cannot be a tensor. The partial derivative fails its purpose and requires an adjustment. The resolution to this problem is found by defining the so-called **covariant derivative**, such that

$$\nabla_\mu V^\nu = \partial_\mu V^\nu + \Gamma^\nu_{\mu\lambda} V^\lambda,$$

where $\Gamma^\nu_{\mu\lambda}$ are called the **Christoffel symbols**. Notice that these symbols are not tensorial by construction, as they ought to cancel out the nontensorial part of the transformation. By explicitly transforming $\nabla_\mu V^\nu$, you can get an explicit expression for the Christoffel symbols $\Gamma^\nu_{\mu\lambda}$.