

2. Derivatives

The **derivative** of a function f is another function, f' , defined as

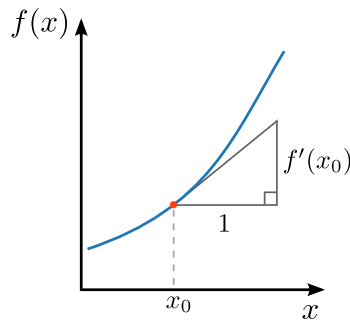
$$f'(x) \equiv \frac{df}{dx} \equiv \lim_{\delta x \rightarrow 0} \frac{f(x + \delta x) - f(x)}{\delta x}. \quad (2.1)$$

This kind of expression is called a **limit expression** because it involves a limit (in this case, the limit $\delta x \rightarrow 0$).

If the derivative exists (i.e., the above limit expression is well-defined) over some domain of x , we say that f is **differentiable** in that domain. It can be shown that a differentiable function is automatically continuous.

If f is differentiable, we can define its second-order derivative f'' as the derivative of f' . Third-order and higher-order derivatives are defined similarly.

Graphically, the derivative represents the slope of the graph of $f(x)$, as shown in the figure below. The second derivative represents the curvature; in this figure, the curve is upward-curving so the second derivative is positive.



2.1 Properties of derivatives

2.1.1 Rules for limit expressions

Let us briefly review the mathematical rules governing limit expressions.

First, the limit of a linear superposition is equal to the linear superposition of limits. Given two constants a_1 and a_2 and two functions f_1 and f_2 ,

$$\lim_{x \rightarrow c} [a_1 f_1(x) + a_2 f_2(x)] = a_1 \lim_{x \rightarrow c} f_1(x) + a_2 \lim_{x \rightarrow c} f_2(x). \quad (2.2)$$

Second, limits obey a product rule and a quotient rule:

$$\lim_{x \rightarrow c} [f_1(x) f_2(x)] = \left[\lim_{x \rightarrow c} f_1(x) \right] \left[\lim_{x \rightarrow c} f_2(x) \right] \quad (2.3)$$

$$\lim_{x \rightarrow c} \left[\frac{f_1(x)}{f_2(x)} \right] = \frac{\lim_{x \rightarrow c} f_1(x)}{\lim_{x \rightarrow c} f_2(x)}. \quad (2.4)$$

As a special exception, the product rule and quotient rule are inapplicable if they result in $0 \times \infty$, ∞/∞ , or $0/0$, which are undefined. As an example of why such combinations are problematic, consider this:

$$\lim_{x \rightarrow 0} x = \lim_{x \rightarrow 0} \left[x^2 \frac{1}{x} \right] \stackrel{?}{=} \lim_{x \rightarrow 0} [x^2] \lim_{x \rightarrow 0} \left[\frac{1}{x} \right] = 0 \times \infty \quad (??)$$

In fact, the true value of the limit expression is 0; the second step was incorrect, since it involved an attempt to apply the product rule where it is inapplicable.

2.1.2 Composition rules for derivatives

Using the rules for limit expressions, we can derive the following elementary composition rules for derivatives:

$$\frac{d}{dx} [\alpha f(x) + \beta g(x)] = \alpha f'(x) + \beta g'(x) \quad (\text{linearity}) \quad (2.5)$$

$$\frac{d}{dx} [f(x)g(x)] = f(x)g'(x) + f'(x)g(x) \quad (\text{product rule}) \quad (2.6)$$

$$\frac{d}{dx} [f(g(x))] = f'(g(x))g'(x) \quad (\text{chain rule}) \quad (2.7)$$

These can all be proven by direct substitution into the definition of the derivative, and taking appropriate orders of limits. With the aid of these rules, we can prove various standard results, such as the “power rule” for derivatives:

$$\frac{d}{dx} [x^n] = nx^{n-1}, \quad n \in \mathbb{N}. \quad (2.8)$$

Moreover, the linearity of the derivative operation implies that derivatives “commute” with sums, i.e. you can move them to the left or right of summation signs. This is a very useful feature. For example, we can use it to prove that the exponential is its own derivative, as follows:

$$\frac{d}{dx} [\exp(x)] = \frac{d}{dx} \sum_{n=0}^{\infty} \frac{x^n}{n!} \quad (2.9)$$

$$= \sum_{n=0}^{\infty} \frac{d}{dx} \frac{x^n}{n!} \quad (2.10)$$

$$= \sum_{n=1}^{\infty} \frac{x^{n-1}}{(n-1)!} \quad (2.11)$$

$$= \exp(x). \quad (2.12)$$

Derivatives also commute with limits. For example, we can use this to prove that the exponential is its own derivative, using the alternative definition of the exponential function discussed in Exercise 1 of Chapter 1:

$$\frac{d}{dx} [\exp(x)] = \frac{d}{dx} \lim_{n \rightarrow \infty} \left(1 + \frac{x}{n}\right)^n \quad (2.13)$$

$$= \lim_{n \rightarrow \infty} \frac{d}{dx} \left(1 + \frac{x}{n}\right)^n \quad (2.14)$$

$$= \lim_{n \rightarrow \infty} \left(1 + \frac{x}{n}\right)^{n-1} \quad (2.15)$$

$$= \exp(x). \quad (2.16)$$

2.2 Taylor series

A function is **infinitely differentiable** at a point x_0 if all orders of derivatives (i.e., the first derivative, the second derivative, etc.) are well-defined at x_0 . If a function is infinitely differentiable at x_0 , then near that point it can be expanded in a **Taylor series**:

$$f(x) \leftrightarrow \sum_{n=0}^{\infty} \frac{(x-x_0)^n}{n!} \left[\frac{d^n f}{dx^n} \right] (x_0) \quad (2.17)$$

$$= f(x_0) + (x-x_0)f'(x_0) + \frac{1}{2}(x-x_0)^2 f''(x_0) + \dots \quad (2.18)$$

Here, the “zeroth derivative” refers to the function itself. The Taylor series can be derived by assuming that $f(x)$ can be written as a general polynomial involving terms of the form $(x - x_0)^n$, and then using the definition of the derivative to find the series coefficients.

Here are some common Taylor series:

$$\frac{1}{1-x} = 1 + x + x^2 + x^3 + \dots \quad \text{for } |x| < 1 \quad (2.19)$$

$$\ln(1-x) = -x - \frac{x^2}{2} - \frac{x^3}{3} - \frac{x^4}{4} - \dots \quad \text{for } |x| < 1 \quad (2.20)$$

$$\sin(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \dots \quad (2.21)$$

$$\cos(x) = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \dots \quad (2.22)$$

$$\exp(x) = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots \quad (2.23)$$

$$\sinh(x) = x + \frac{x^3}{3!} + \frac{x^5}{5!} + \frac{x^7}{7!} + \dots \quad (2.24)$$

$$\cosh(x) = 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \frac{x^6}{6!} + \dots \quad (2.25)$$

Note that it is possible for a function to have a divergent Taylor series, or a Taylor series that converges to a different value than the function itself. The conditions under which such breakdowns occur is a complicated topic that we will not delve into.

For certain functions, however, the Taylor series are exact. The Taylor series for the exponential, Eq. (2.23), is exactly the same as the definition of the exponential itself. Likewise, for Eqs. (2.21)–(2.25), the Taylor series can be shown to converge to the value of the function for all $x \in \mathbb{R}$, which means that the series form is *exactly equivalent* to the function itself. (Thanks to this, it is common for math textbooks to start out by *defining* the trigonometric functions in series form, and then deriving their geometric meanings, rather than vice versa.)

2.3 Ordinary differential equations

Differential equations are equations that contain derivatives. For example, the equation

$$\frac{df}{dx} = f(x) \quad (2.26)$$

involves both f and its first derivative. It is called an **ordinary differential equation**, meaning that it contains a derivative with respect to a single variable x , rather than multiple variables.

Solving a differential equation means finding a function that satisfies the equation. There is no single method for solving differential equations. Sometimes, we can guess the solution; for example, by trying different elementary functions, we can discover that the above differential equation is solved by

$$f(x) = A \exp(x). \quad (2.27)$$

Certain classes of differential equation can be solved using techniques like Fourier transforms, Green’s functions, etc., some of which will be taught in this course. On the other hand, many differential equations have no known exact analytic solution.

Example—The following ordinary differential equation describes a damped harmonic oscillator:

$$\frac{d^2x}{dt^2} + 2\gamma\frac{dx}{dt} + \omega_0^2x(t) = 0. \quad (2.28)$$

In this case, $x(t)$ is the function, and t is the input variable. This is unlike our previous notation where x was the input variable, so don't get confused! This equation is obtained by applying Newton's second law to an object moving in one dimension, subject to both a damping force and a restoring force, with $x(t)$ denoting the position as a function of time t . We will study Eq. (2.28) in detail later in the course.

2.3.1 Specific solutions and general solutions

When confronted with an ordinary differential equation, the first thing you should check is the highest derivative appearing in the equation. This is called the **order** of the differential equation. If the equation has order N , its **general solution** contains N **free parameters** that can be assigned any value. Therefore, if you happen to guess a solution, but that solution does not contain N free parameters, then you know the solution isn't the most general one.

For example, the ordinary differential equation

$$\frac{df}{dx} = f(x) \quad (2.29)$$

has order one. We have previously guessed the solution $f(x) = A \exp(x)$, which has one free parameter, A . So we know our work is done: there is no solution more general than the one we found.

A **specific solution** to a differential equation is a solution containing no free parameters. One way to get a specific solution is to start from a general solution, and assign a value to every free parameter. In physics problems, the assigned values are commonly determined by **boundary conditions**. For an ordinary differential equation of order N , we need independent boundary N conditions to define a specific solution.

2.4 Partial derivatives

So far, we have focused on functions which take a single input. Functions can also take multiple inputs; for instance, we can have a function f that takes two input numbers, x and y , and outputs a number $f(x, y)$. We can define a **partial derivative** as a derivative of a function with respect to one of the inputs, keeping all other inputs fixed.

For example, the function

$$f(x, y) = \sin(2x - 3y^2) \quad (2.30)$$

has the following partial derivatives:

$$\frac{\partial f}{\partial x} = 2 \cos(2x - 3y^2), \quad \frac{\partial f}{\partial y} = -6y \cos(2x - 3y^2). \quad (2.31)$$

2.4.1 Change of variables

We saw in Section 2.1.2 that single-variable functions obey a derivative composition rule,

$$\frac{d}{dx} f(g(x)) = g'(x) f'(g(x)). \quad (2.32)$$

This composition rule can be generalized to partial derivatives, which is often required when performing a **change of coordinates**. Suppose we have a function $f(x, y)$, and we want to re-express (x, y) in a different coordinate system (u, v) . Each coordinate in the old system may depend on *both* coordinates in the new system:

$$x = x(u, v), \quad y = y(u, v). \quad (2.33)$$

Expressed in the new coordinates, the function is

$$F(u, v) \equiv f(x(u, v), y(u, v)). \quad (2.34)$$

It can be shown that the transformed function's partial derivatives obey the composition rule

$$\frac{\partial F}{\partial u} = \frac{\partial f}{\partial x} \frac{\partial x}{\partial u} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial u} \quad (2.35)$$

$$\frac{\partial F}{\partial v} = \frac{\partial f}{\partial x} \frac{\partial x}{\partial v} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial v}. \quad (2.36)$$

On the right-hand side of these equations, the partial derivatives are meant to be expressed in terms of the new coordinates (u, v) . For example,

$$\frac{\partial f}{\partial x} \equiv \left. \frac{\partial f}{\partial x} \right|_{x=x(u,v), y=y(u,v)}$$

The generalization of this rule to more than two inputs is straightforward. For a function $f(x_1, \dots, x_N)$, a change of coordinates $x_i = x_i(u_1, \dots, u_N)$ involves the composition

$$F(u_1, \dots, u_N) = f(x_1(u_1, \dots, u_N), \dots), \quad (2.37)$$

$$\frac{\partial F}{\partial u_i} = \sum_{j=1}^N \frac{\partial f}{\partial x_j} \frac{\partial x_j}{\partial u_i}. \quad (2.38)$$

Example—In two dimensions, Cartesian and polar coordinates are related by

$$x = r \cos \theta, \quad y = r \sin \theta. \quad (2.39)$$

Given a function $f(x, y)$, we can re-write it in polar coordinates as $F(r, \theta)$. The partial derivatives in polar coordinates are then given by

$$\frac{\partial F}{\partial r} = \frac{\partial f}{\partial x} \frac{\partial x}{\partial r} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial r} = \frac{\partial f}{\partial x} \cos \theta + \frac{\partial f}{\partial y} \sin \theta. \quad (2.40)$$

$$\frac{\partial F}{\partial \theta} = \frac{\partial f}{\partial x} \frac{\partial x}{\partial \theta} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial \theta} = -\frac{\partial f}{\partial x} r \sin \theta + \frac{\partial f}{\partial y} r \cos \theta. \quad (2.41)$$

2.4.2 Partial differential equations

A **partial differential equation** is a differential equation involving multiple partial derivatives (as opposed to an ordinary differential equation, which involves derivatives with respect to a single variable).

An example of a partial differential equation encountered in physics is Laplace's equation,

$$\frac{\partial^2 \Phi}{\partial x^2} + \frac{\partial^2 \Phi}{\partial y^2} + \frac{\partial^2 \Phi}{\partial z^2} = 0, \quad (2.42)$$

which describes the electrostatic potential $\Phi(x, y, z)$ at position (x, y, z) , in the absence of any electric charges.

Partial differential equations are considerably harder to solve than ordinary differential equations. In particular, their boundary conditions are more complicated to specify: whereas each boundary condition for an ordinary differential equation consists of a single number (e.g., the value of $f(x)$ at some point $x = x_0$), each boundary condition for a partial differential equation consists of a *function* (e.g., the values of $\Phi(x, y, z)$ along some curve $g(x, y, z) = 0$).

2.5 Exercises

1. Show that if a function is differentiable, then it is also continuous.

2. Prove that the derivative of $\ln(x)$ is $1/x$.

[solution available]

3. Using the definition of non-natural powers, prove that

$$\frac{d}{dx}[x^y] = yx^{y-1}, \quad \text{for } x \in \mathbb{R}^+, y \notin \mathbb{N}. \quad (2.43)$$

4. Consider $f(x) = \tanh(\alpha x)$.

(a) Sketch $f(x)$ versus x , for two cases: (i) $\alpha = 1$ and (ii) $\alpha \gg 1$.

(b) Sketch the derivative function $f'(x)$ for the two cases, based on your sketches in part (A) (i.e., without evaluating the derivative directly).

(c) Evaluate the derivative function, and verify that the result matches your sketches in part (B).

5. Prove geometrically that the derivatives of the sine and cosine functions are:

$$\frac{d}{dx} \sin(x) = \cos(x), \quad \frac{d}{dx} \cos(x) = -\sin(x). \quad (2.44)$$

Hence, derive their Taylor series.

6. For each of the following functions, derive the Taylor series around $x = 0$:

(a) $f(x) = \ln[\alpha \cos(x)]$, to the first 3 non-vanishing terms.

(b) $f(x) = \cos[\pi \exp(x)]$, to the first 4 non-vanishing terms.

(c) $f(x) = \frac{1}{\sqrt{1 \pm x}}$, to the first 4 non-vanishing terms. Keep track of the signs (i.e., \pm versus \mp).

7. For each of the following functions, sketch the graph and state the domains over which the function is differentiable:

(a) $f(x) = |\sin(x)|$

(b) $f(x) = [\tan(x)]^2$

(c) $f(x) = \frac{1}{1 - x^2}$

8. Let $\vec{v}(x)$ be a *vectorial* function which takes an input x (a number), and gives an output value \vec{v} that is a 2-component vector. The derivative of this vectorial function is defined in terms of the derivatives of each vector component:

$$\vec{v}(x) = \begin{bmatrix} v_1(x) \\ v_2(x) \end{bmatrix} \Rightarrow \frac{d\vec{v}}{dx} = \begin{bmatrix} dv_1/dx \\ dv_2/dx \end{bmatrix}. \quad (2.45)$$

Now suppose $\vec{v}(x)$ obeys the vectorial differential equation

$$\frac{d\vec{v}}{dx} = \mathbf{A}\vec{v}, \quad (2.46)$$

where

$$\mathbf{A} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \quad (2.47)$$

is a matrix that has two distinct real eigenvectors with real eigenvalues.

- (a) How many independent numbers do we need to specify for the general solution?
- (b) Let \vec{u} be one of the eigenvectors of \mathbf{A} , with eigenvalue λ :

$$\mathbf{A}\vec{u} = \lambda\vec{u}. \quad (2.48)$$

Show that $\vec{v}(x) = \vec{u}e^{\lambda x}$ is a specific solution to the vectorial differential equation. Hence, find the general solution.

[solution available]