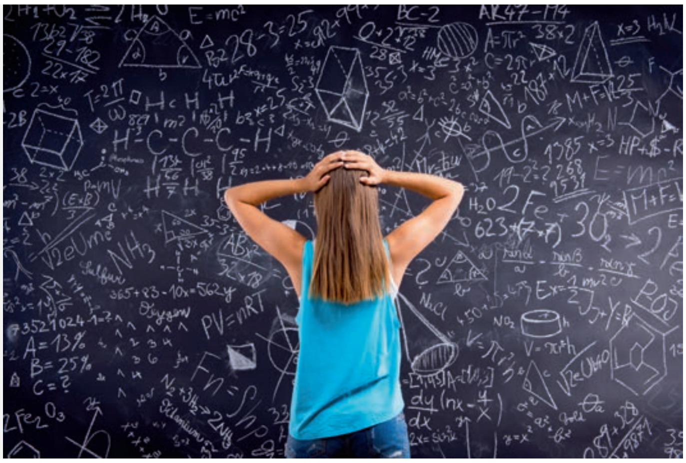


Unit 1

Calculus

STUDY GOALS

On completion of this unit, you will have learned …

… how to differentiate and integrate functions of a single variable.

… how to perform partial differentiation and multiple integrals for functions with multiple variables.

… how to approximate a function in a Taylor series.

… the basic concepts of calculus of variations.

1. Calculus

Introduction

Functions express relationships between variables. For example, in standard notation, the function $y = f(x)$ formalizes how a value y, called the dependent variable, varies with respect to another value x, called the independent variable. The letter f is the name of the function.

The rate at which the dependent variable changes with respect to the independent variable is of particular interest, both mathematically and for applications. One example of such a rate of change is the change in distance with respect to time — also known as velocity. The method for finding this rate when given a function is called differentiation.

Differentiation is a powerful tool that is used to understand the relationship between variables. In the case of multiple independent variables, for example, $z = f(x, y)$, partial differentiation allows us to explore the rate of change of the function with respect to each independent variable.

The operation of differentiation can often be "undone" via an operation called integration. Integration can be seen as the inverse of differentiation and therefore, it is often called the anti-derivative. Intuitively, this means that if we have a formula that expresses the speed of a particle with respect to time, we can often construct a formula for the displacement of the particle (i.e., the distance it has traveled).

In calculus of variations, the concept of differentiation is extended to functionals, which are maps from a set of functions to the set of real numbers. In this sense, functionals are functions of functions. The calculus of variations looks for which input function maximizes or minimizes the dependent (output) variable.

Good textbooks that further cover this subject area are e. g. (Deisenroth, Faisal & Ong, 2020, Chap. 5), (Strang, 2017, Chap. 2-5, 7, 8, 13, 14) and (Loomis & Sternberg, 2014, Chap. 3, 8).

1.1 Differentiation and Integration

Derivatives of Functions of a Single Variable

We are often interested in how a **function** changes with respect to its argument. For example, we could imagine a travelling car with position s at a given time t. We know from our everyday experience that at each time, t, a car has a velocity, $v(t)$, which measures how fast the car is travelling at time t. Over a given time interval, Δt, the average velocity describes the rate at which the car travels the distance for that interval of time

$$v(t) = \frac{\Delta s}{\Delta t},$$

(1.1)

Function
A function is a relation between two sets that associates every element of one set to exactly one element of the other set.

where Δs is the change in position of the car or the distance covered by the car. In equation 1.1, the time t is the argument of the function v, which establishes a relationship between the distance covered by the car and the time it takes to cover that distance.

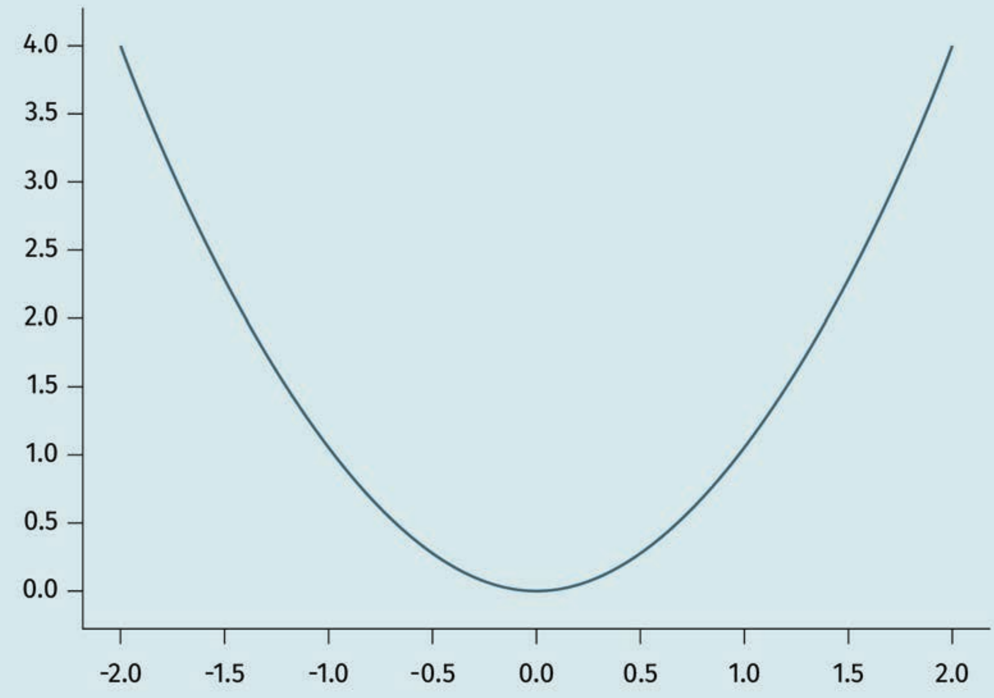
More generally, we often wish to find the rate of change of a general function $f(x)$, where f depends on some argument x. We will begin by considering functions that depend only on a single variable, such as $f(x) = x^2$, which is shown in the graphic that follows this explanation. Fix a given value of x and let us consider the value of the function, $f(x)$, as we change the input to a slightly different value, where we assume that the function is continuous and doesn't have any "kinks" or "jumps." For example, if we start at $x_0 = 1$ and move to $x_1 = 1.1$, the value of the function $f(x) = x^2$ will change from $f(x_0) = 1$ to $f(x_1) = 1.21$. Let us denote this change in x by Δx and write $x \to x + \Delta x$ to indicate that x changes from x to $x + \Delta x$. Then, the change in the value of the function f is $\Delta f = f(x + \Delta x)$ - $f(x)$. By making this increment Δx arbitrarily small, we can work out the rate of change of f at a single instant. That is the idea of a derivative — an instantaneous rate of change — and the limit operation allows us to formally capture this intuition. As we make the change in x smaller and smaller, written $\Delta x \to 0$, we can define the gradient or first derivative of the function f as

$$f'(x) \equiv \frac{df(x)}{dx} \equiv \lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x}.$$

(1.2)

The function is differentiable at x_a if, and only if, this limit exists at the point $x = x_a$. Note that if the limit does not exist at $x = x_a$, the function is not differentiable at this x_a. The definition 1.2 does not specify if we approach x from smaller values (so Δx is negative) or vice versa (in which case Δx is positive). This is because, in order for the limit to exist, the definition of the limit requires that the quotient $\frac{f(x+\Delta x) - f(x)}{\Delta x}$ approaches the same value, $f'(x)$, from both the left and right of x.

A Parabola: $f(x) = x^2$



Geometrically, the derivative $f'(x)$ can be interpreted as the slope of the line tangent to the function $f(x)$ at the point x.

Example

Find the first derivative of $f(x) = x^2$.

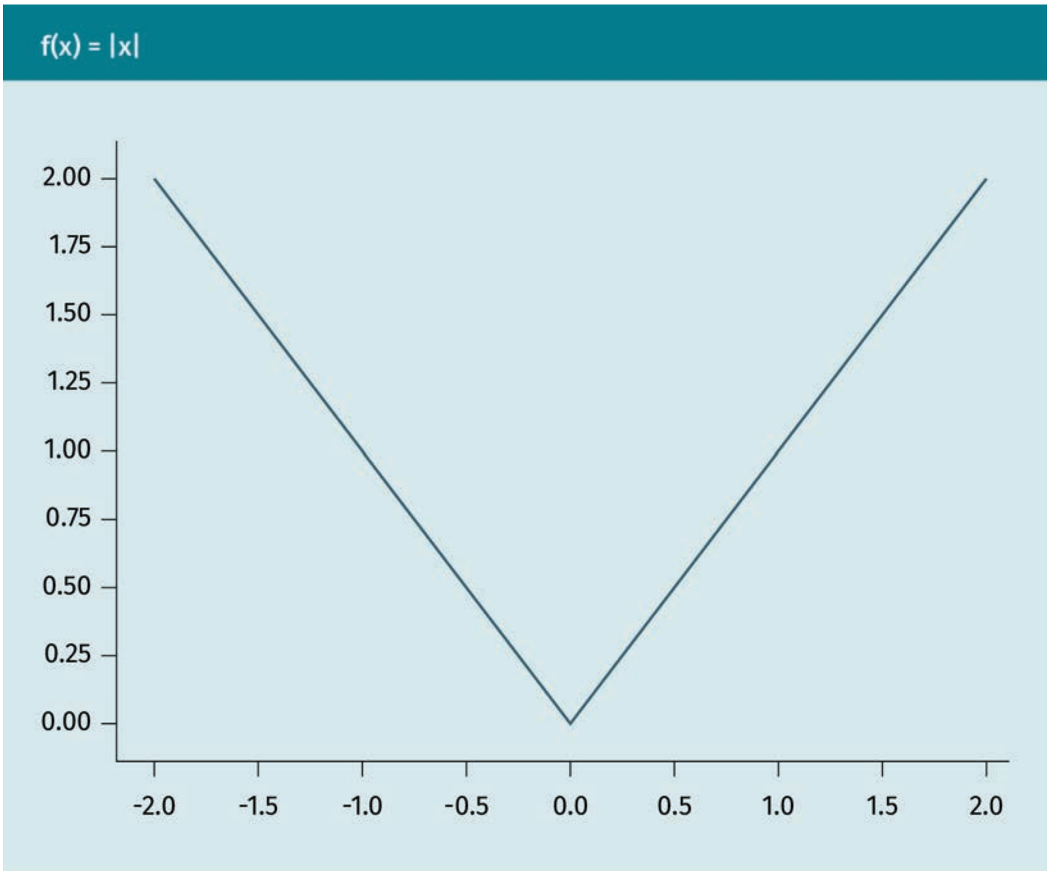
Using definition 1.2,

$$\begin{aligned}
f'(x) &= \lim_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x} \\
&= \lim_{\Delta x \to 0} \frac{(x+\Delta x)^2 - x^2}{\Delta x} \\
&= \lim_{\Delta x \to 0} \frac{x^2 + 2x\Delta x + (\Delta x)^2 - x^2}{\Delta x} \\
&= \lim_{\Delta x \to 0} \frac{2x\Delta x + (\Delta x)^2}{\Delta x} \\
&= \lim_{\Delta x \to 0} \frac{\Delta x(2x+\Delta x)}{\Delta x} \\
&= \lim_{\Delta x \to 0} (2x+\Delta x) \\
&= 2x
\end{aligned}$$

Here, we have observed that Δx becomes infinitessimally small as it approaches zero, but is always non-zero and can therefore be cancelled from the numerator and denominator.

Be aware that to be differentiable at x_a, a function must be continuous at x_a (or else the limit will not exist at that point), but merely being continuous everywhere does not necessarily mean that a function is differentiable everywhere, as shown in the following figure. As x approaches 0 from the left, the derivative (which is the slope, in this case) is –1, but if we approach $x = 0$ from the right, the limit of the quotient $\frac{f(x+\Delta x) - f(x)}{\Delta x}$ is +1. The right and left hand limits do not agree, so the limit, and therefore the derivative of $f(x) = |x|$, is not defined at $x = 0$.

Using definition 1.2 in combination with the laws of limits, one can find derivatives of many fundamental functions. For reference, here are the derivatives of some important functions where $n > 0$ is a natural number and a is a real-valued constant.



$$\frac{d}{dx}x^n = nx^{n-1} \qquad \frac{d}{dx}e^{ax} = ae^{ax}$$

$$\frac{d}{dx}\ln(ax) = \frac{d}{dx}(\ln(a) + \ln(x)) = \frac{1}{x} \qquad \frac{d}{dx}\sin(ax) = a\cos(ax)$$

$$\frac{d}{dx}\cos(ax) = -a\sin(ax) \qquad \frac{d}{dx}\tan(ax) = \frac{a}{\cos^2(ax)}$$

Higher order derivatives

Derivatives are themselves functions, therefore we can consider their rates of change. We call derivatives of derivatives of a function f higher order derivatives of f, and they are obtained using the definition of the derivative in the same way. For the second derivative, we use definition 1.2 but replace the function $f(x)$ with the first derivative $f'(x)$ as follows:

$$f''(x) \equiv \frac{df'(x)}{dx} \equiv \lim_{\Delta x \to 0} \frac{f'(x + \Delta x) - f'(x)}{\Delta x}, \tag{1.3}$$

where, again, f'' is defined if, and only if, the limit exists. More generally, we can define the n^{th} derivative of $f(x)$ to be

$$f^{(n)}(x) \equiv \frac{df^{(n-1)}}{dx} \equiv \lim_{\Delta x \to 0} \frac{f^{(n-1)}(x + \Delta x) - f^{(n-1)}(x)}{\Delta x}, \tag{1.4}$$

whenever the limit exists.

Stationary points
Looking again at the first graphic depicting a parabola, we notice that the point $(0, 0)$ is special; the value of the function on either side of point $x = 0$ is greater than at $x = 0$. In other words, at $x = 0$, f achieves a local minimum. Graphically, we observe that the line tangent to the graph of f at this point is horizontal — its slope is equal to zero. To reiterate: the slope of the line tangent to f at $x = 0$, which is $f'(0)$, has a derivative at that point with a value of zero.

Points where the derivative is equal to zero, such as the point described above, are called stationary points. After examining a number of examples, we see that f' is often, but not always, equal to zero at a local minimum. The other possibility, illustrated by the previous graphic, is that the slope of the tangent line, the derivative, is undefined at the local extrema. For $f(x) = |x|$, $(0, 0)$ is a critical point, defined to be a place where the derivative is zero or does not exist, but it is not a stationary point.

Note that there are three different stationary points. They are as follows:

- the function f has a maximum at a stationary point at $x = a$ if $f'(a) = 0$ and $f''(a) < 0$,
- the function f has a minimum at a stationary point at $x = a$ if $f'(a) = 0$ and $f''(a) > 0$, and
- a stationary point at $x = a$ is called a saddle point if $f'(a) = 0$ and f'' changes sign at this point.

Note that the maximum and minimum found this way may not be the global maximum or minimum of the function, but rather a local extremum at the stationary point.

Rules of Differentiation

Differentiation of functions with a constant
Some functions are composed of a constant and a variable part, e.g. $f(x) = a \cdot g(x)$ where a is an arbitrary constant and $g(x)$ is some function that depends on x. The derivative is given by

$$\frac{d}{dx}f(x) = f'(x) = a\frac{d}{dx}g(x) = ag'(x)\,.$$

Differentiation of products
Previously in this section, differentiation rules for some functions with simple structures were discussed. However, in many cases, we are interested in the rates of change of functions that are more complicated.

As a first example, we will investigate how to differentiate functions that can be written as products of two other functions, namely functions of the form $f(x) = u(x) \cdot v(x)$. The idea is that if we know how to differentiate u and v, and how to use that information together with the product structure to find a derivative of f, we can avoid applying the definition of the derivative. We could, from this perspective, reexamine $f(x) = x^2$, noting that we could write it as $f(x) = x \cdot x$. Slightly more complicated examples are $g(x) = x^2 \cdot \sin(x)$, which we could decompose as $g(x) = u(x)$ where $u(x) = x^2$ and $v(x) = \sin(x)$. Such a decomposition is not unique; we could consider any functions u and v whose product is $x^2 \cdot \sin(x)$. However, the idea behind decomposing the original function $f(x)$ into two functions, u and v, is to choose u and v that are easier to differentiate than f. Then, if we have a general method to calculate the derivative of a product, we can apply that method to f in order to make taking the derivative easier than if we were to calculate it using definition and equation 1.2. This general method, called the product rule, is obtained from the definition (See equation 1.2) as follows. First, let's simplify the difference $f(x{+}\Delta x) - f(x)$. This results in

$$\begin{aligned} f(x+\Delta x) - f(x) &= u(x+\Delta x) \cdot v(x+\Delta x) - u(x) \cdot v(x) \\ &= u(x+\Delta x)[v(x+\Delta x) - v(x)] + v(x)[u(x+\Delta x) - u(x)] \,. \end{aligned}$$

Note that we added and subtracted $v(x)u(x{+}\Delta x)$ in order to be able to factor. Substituting the result of our simplification into the definition of the derivative, we obtain

$$\begin{aligned} \frac{df}{dx} &= \lim_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x} \\ &= \lim_{\Delta x \to 0} \left\{ u(x+\Delta x)\left[\frac{v(x+\Delta x) - v(x)}{\Delta x}\right] + v(x)\left[\frac{u(x+\Delta x) - u(x)}{\Delta x}\right]\right\} . \end{aligned}$$

As Δx approaches zero, $u(x{+}\Delta x)$ approaches $u(x)$ and the terms in the square brackets become the derivatives of the functions u and v respectively. Hence, the formula for the derivative of a product of functions, called the product rule, is given by

$$f' \equiv \frac{df}{dx} \equiv \frac{d}{dx}\Big[u(x)v(x)\Big] = u(x)\frac{dv(x)}{dx} + v(x)\frac{du(x)}{dx} = uv' + vu' \,. \tag{1.5}$$

Using this rule repeatedly, the derivative of products of three or more differentiable functions can be obtained as follows:

$$\begin{aligned} &\text{Given } f(x) = u(x)v(x)w(x), \\ &f'(x) = \frac{df}{dx} = u\frac{d}{dx}(vw) + vw\frac{d}{dx}u \\ &\qquad\quad = uv\frac{dw}{dx} + uw\frac{dv}{dx} + vw\frac{du}{dx} \,. \end{aligned}$$

Example

Find the derivative of $f(x) = x^2 \sin(x)$.

Using definition 1.5 with $u(x) = x^2$ and $v(x) = \sin(x)$, we get

$$\begin{aligned}\frac{d}{dx}x^2\sin(x) &= x^2\frac{d}{dx}\sin(x) + \sin(x)\frac{d}{dx}x^2 \\ &= x^2\cos(x) + 2x\sin(x)\,.\end{aligned}$$

The chain rule

Many functions can be written as compositions of functions, namely as functions whose inputs are functions themselves. For example, $f(x) = (x-1)^2$ can be written as $f(x) = u^2(x)$ where $u(x) = x - 1$. Write this as $f(u(x))$.

The essential idea of the chain rule is that we differentiate the outer function f with respect to the inner function u to get $f'(u)$, leaving the inner function alone. Then differentiate the inner function u with respect to x to get $u'(x)$ and multiply the two together:

$$\frac{df}{dx} = \frac{df}{du} \cdot \frac{du}{dx}\,. \tag{1.6}$$

This is known as the chain rule because we “chain” the derivatives together. The concept can be easily extended to functions of functions of functions, and so on. We only need to repeatedly apply the chain rule until we reach the independent variable.

Example

Find the derivative of $f(x) = (x-1)^2$.

We can write this as $f(x) = u^2\,(x)$, where $u(x) = x{-}1$. Using theorem 1.6 we obtain

$$\begin{aligned}\frac{df}{dx} &= \frac{df}{du} \cdot \frac{du}{dx} \\ &= 2u\frac{du}{dx} \\ &= 2u \cdot 1 \\ &= 2(x-1)\,.\end{aligned}$$

The chain rule can also be used to calculate the derivative of functions of the form $f(x) = 1/v(x)$. Rather than writing this as a quotient, we can express it as a composition of functions, $f(x) = v^{-1}(x)$ (noting that this is the –1 power, not the inverse), and then apply the chain rule

$$\begin{aligned}\frac{df}{dx} &= \frac{df}{dv}\frac{dv}{dx} \\ &= -v^{-2}\frac{dv}{dx} \\ &= -\frac{1}{v^2(x)}\frac{dv}{dx}\end{aligned}$$

where we have used the elementary derivative $\frac{d}{dx}x^n = nx^{n-1}$.

Differentiation of quotients
In some cases, the function we want to take the derivative of can be written in the form of a quotient of two functions, such as $f(x) = \frac{u(x)}{v(x)}$. One way to create a rule in order to calculate derivatives for such functions is to combine the product rule in equation 1.5 with the chain rule, and write the product as $f(x) = u(x)[1/v(x)]$. Applying the product rule, we get

$$\begin{aligned}\frac{df}{dx} &= \frac{d}{dx}\left(\frac{u(x)}{v(x)}\right) \\ &= u\frac{d}{dx}\left(\frac{1}{v(x)}\right) + \frac{1}{v(x)}\frac{d}{dx}u(x)\,.\end{aligned}$$

Using the chain rule to evaluate $\frac{d}{dx}\left(\frac{1}{v(x)}\right)$ as above, we obtain

$$\frac{df}{dx} = u\left(-\frac{dv(x)/dx}{v(x)^2}\right) + \frac{du(x)/dx}{v(x)}\,.$$

The “prime” notation for the derivative yields an expression that is easier to read

$$f' = \left(\frac{u}{v}\right)' = \frac{vu' - uv'}{v^2}, \tag{1.7}$$

where $u = u(x)$ and $v = v(x)$.

Integrals of Functions of a Single Variable

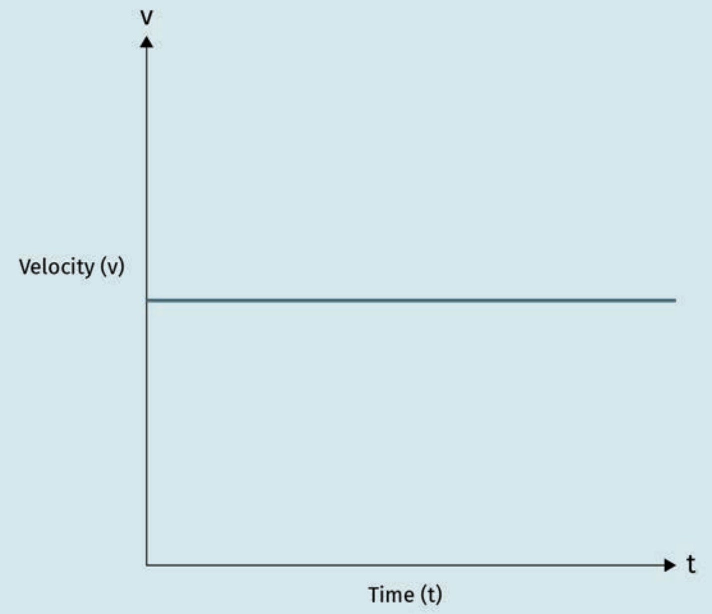
Integrals as area under the curve
In the first part of this unit, we focused on rates of change of functions of a single variable, and developed the first derivative, or gradient, as a tool to mathematically investigate rates of change. Returning to the example of the car traveling on a road, we found that we could express the average velocity as the distance traveled by the car in a given amount of time,

$$v = \frac{\Delta s}{\Delta t},$$

where Δs is the change in position (the distance) over time interval Δt. We considered progressively shorter time intervals in order to investigate the instantaneous rate of change

$$v(t) = \frac{ds}{dt}.$$

Constant Velocity Against Time



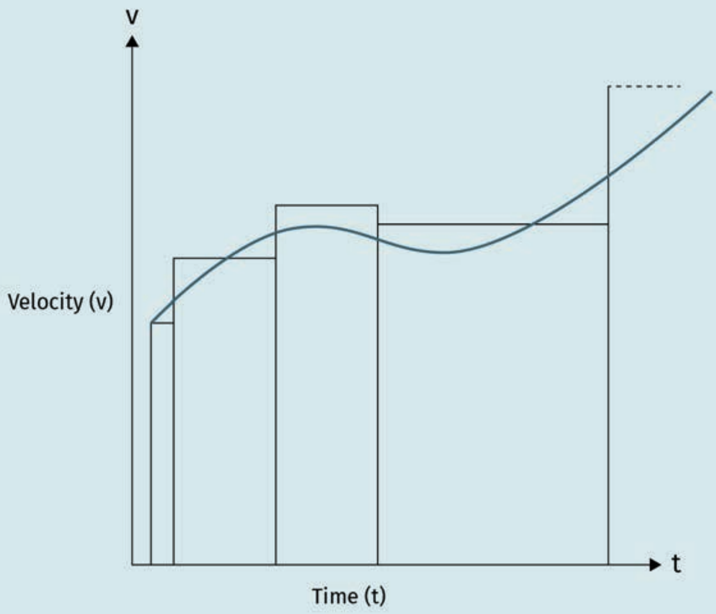
The first derivative of the position with respect to time is the instantaneous velocity.

If we know a function for the velocity $v(t)$, it is natural to wonder if it is possible to calculate the distance Δs the car has traveled over a given time interval, Δt.

If the velocity is constant, this is intuitively clear. We have $\Delta s = v\Delta t$, as illustrated in the figure above. Note that the area of the rectangle $v \cdot \Delta t$ corresponds to change in the position Δs. In most cases, the velocity of a car will not be constant, and therefore we would like a more general way to express the distance as a (continuous) function of velocity and time.

Informally, we can do this by looking at very small sub-intervals rather than considering the whole interval Δt at once. We will then assume that the velocity is constant over these small rectangles in order to get an approximation of the distance the car travels over each small interval. Each time, we will use our constant approximation of the velocity multiplied by the length of time to get the area of the small rectangle, as shown in the following figure. We know that we have a small error in each interval, but the smaller the intervals, the smaller the error becomes. To find the total distance the car has traveled, we combine the total areas of all the rectangles that represent the contributions from each small time interval.

Variable Velocity Against Time



More formally, consider an arbitrary function $f(x)$ of a single variable x that is defined over the interval $a \leq x \leq b$. Following the approach above, we divide the interval $[a, b]$ into many sub-intervals by introducing intermediate points ξ_i so that $a = \xi_1 < \xi_2 < \ldots < \xi_n = b$. The lengths of the intervals $(\xi_i - \xi_{i-1})$ are the lengths of the rectangles on the x-axis and the $f(\xi_i)$ are the heights of the rectangles. The sum S,

$$S = \sum_{i=1}^{n} f(x_i)(\xi_i - \xi_{i-1}), \tag{1.8}$$

is the area of all of the rectangles. The area under some curves over certain intervals is not finite; consider $f(x) = \frac{1}{x}$ over the interval [0, 1]. Therefore, as we take more and more intervals, i.e. as we consider the limit of S as n approaches ∞, the sum S may or may not converge to a finite limit. If this limit exists, the limit of the sum is the definite integral I of the function $f(x)$ in the interval $[a, b]$,

$$I = \int_a^b f(x) dx\,.$$

(1.9)

If the limit does not exist, the integral is undefined.

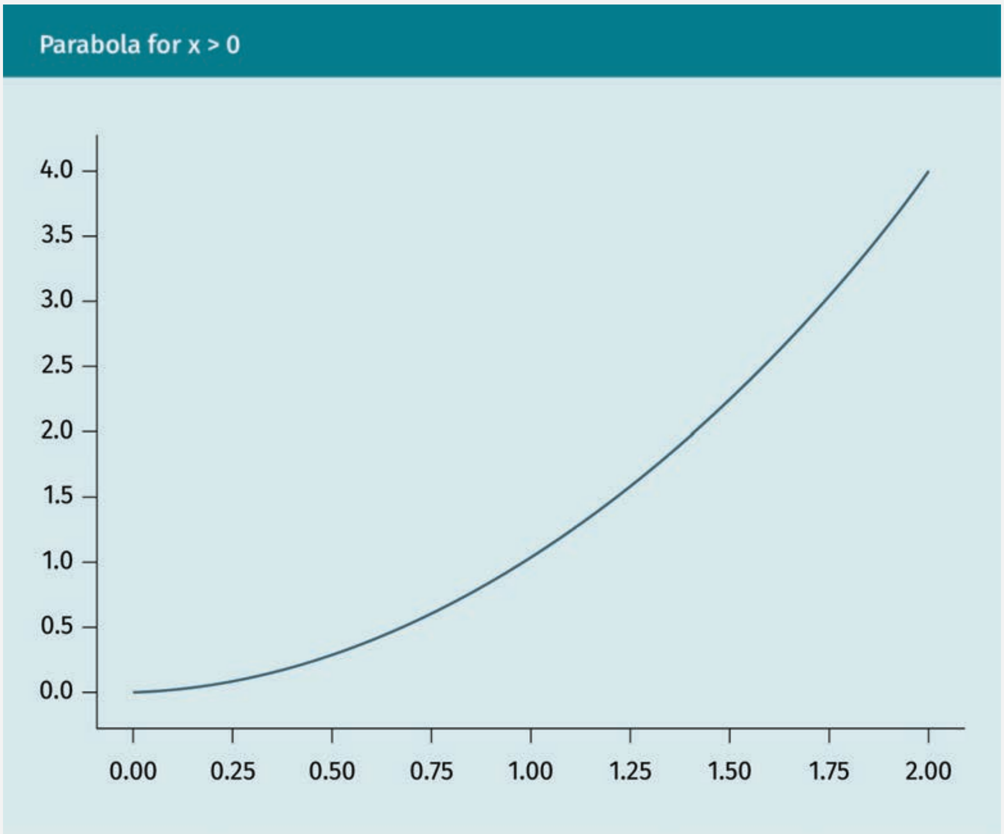
For closed, finite intervals, the question of whether this limit exists — whether the function f is integrable over the given interval — hinges on whether the function f is continuous over that interval.

For continuous functions over a finite interval $[a, b]$ this limit, the integral, always exists.

Example

Evaluate the integral $I = \int_0^b x^2 dx$.

The function $f(x) = x^2$, called the integrand, is shown below.



The first step toward computing this integral, or determining whether this limit exists, is to divide the interval $[0, b]$ into n rectangles of uniform width w. Next, we evaluate the function $f(x) = x^2$ at the left hand endpoint of each sub-interval to determine the height of each rectangle. We could also have taken the value at the right hand endpoint or any in the middle — the limit does not depend on this choice. The area of the i^{th} rectangle is then $w \cdot (iw)^2 = i^2 w^3$. The total area of our approximation, A, is then given by

$$A = \sum_{i=1}^{n} i^2 w^3 .$$

The term w^3 is a constant with respect to the index of summation, i, so we can factor it out of the sum operator as follows:

$$A = w^3 \sum_{i=1}^{n} i^2 .$$

Recall that the sum $\sum_{i=1}^{n} i^2$ has the closed form

$$\sum_{i=1}^{n} i^2 = \frac{1}{6} n(n+1)(2n+1),$$

and hence the area of our approximation is

$$A = w^3 \frac{1}{6} n(n+1)(2n+1) .$$

When constructing the rectangles, we divided the interval $[0, b]$ into intervals of the same length, namely $w = \frac{b}{n}$. Therefore, we can substitute into our expression for A and reduce to get

$$\begin{aligned} A &= \left(\frac{b}{n}\right)^3 \frac{1}{6} n(n+1)(2n+1) \\ &= \frac{b^3}{6} \frac{n(n+1)(2n+1)}{n^3} \\ &= \frac{b^3}{6} \frac{(n+1)(2n+1)}{n^2} \\ &= \frac{b^3}{6} \frac{\left(2n^2+3n+1\right)}{n^2} \\ &= \frac{b^3}{6} \left(2 + \frac{3}{n} + \frac{1}{n^2}\right) . \end{aligned}$$

As we increase the number of intervals without bound, or as $n \to \infty$, the value of the sum in the above expression approaches 2 and thus $\frac{b^3}{3}$ is the value of the finite integral $I = \int_0^b x^2 dx = \frac{1}{3} b^3$.

Using the properties of limits and finite sums, as above, one can see that the following properties hold:

$$\int_a^b 0dx = 0 \tag{1.10}$$

$$\int_a^a f(x)dx = 0 \tag{1.11}$$

$$\int_a^b [f(x) + g(x)]dx = \int_a^b f(x)dx + \int_a^b g(x)dx \tag{1.12}$$

$$\int_a^c f(x)dx = \int_a^b f(x)dx + \int_b^c f(x)dx, \text{ for all } b \in [a, c] \tag{1.13}$$

If we set $c = a$ in the last expression, we can derive the identity

$$\int_a^b f(x)dx = -\int_b^a f(x)dx\,.$$

Integrals as the inverse of differentiation

So far we have introduced integrals over finite intervals $[a, b]$, where the bounds a and b are fixed. We can formally define the function $F(x)$ to be

$$F(x) = \int_a^x f(u)du\,. \tag{1.14}$$

To see how integration is related to differentiation, we evaluate the function F at position $x+\Delta x$ and apply equation 1.13 to get

$$\begin{aligned} F(x + \Delta x) &= \int_a^{x+\Delta x} f(u)du \\ &= \int_a^x f(u)du + \int_x^{x+\Delta x} f(u)du \\ &= F(x) + \int_x^{x+\Delta x} f(u)du\,. \end{aligned}$$

If we divide both sides by Δx and bring $F(x)$ to the left side, the equation reads

$$\frac{F(x+\Delta x)-F(x)}{\Delta x}=\frac{1}{\Delta x}\int_{x}^{x+\Delta x} f(u)du\,.$$

Considering the limit as Δx approaches zero of both sides, this becomes

$$\frac{dF(x)}{dx}=f(x), \tag{1.15}$$

or, written with the definition of $F(x)$ substituted in,

$$\frac{d}{dx}\left[\int_{a}^{x} f(u)du\right]=f(x)\,. \tag{1.16}$$

This says that the derivative of the integral gives back the original integrand. This very important result is called the Fundamental Theorem of Calculus, and it has a second part, which relates the definite integral to the antiderivative. Let's explore it now.

The above discussion did not depend on any attribute of the arbitrary constant a. Hence, the inverse of differentiation is not unique. However, any two inverse functions $F_1(x)$ and $F_2(x)$ differ at most by a constant, so it is written as

$$\int f(x)dx=F(x)+c \tag{1.17}$$

for the family of functions with derivative $f(x)$. Recall that $\frac{d}{dx}c=0$. This is the indefinite integral of $f(x)$, and c is called the constant of integration.

The antiderivative $F(x)$ can also be used to evaluate definite integrals. Let x_0 be an arbitrary fixed point x_0 in (a, b) and consider equation 1.13 to obtain

$$\int_{a}^{b} f(x)dx=\int_{a}^{x_0} f(x)dx+\int_{x_0}^{b} f(x)dx \tag{1.18}$$

$$=\int_{x_0}^{b} f(x)dx+\int_{a}^{x_0} f(x)dx$$

(1.19)

$$\int_a^b f(x)dx = \int_{x_0}^b f(x)dx - \int_{x_0}^a f(x)dx$$

(1.20)

$$= F(b) - F(a)\,.$$

(1.21)

Integrals with infinite bounds of integration

The somewhat intuitive definition of the integral as the area under a curve or as an inverse function does not allow for bounds of integration that are infinite. However, we can extend the definition to include these cases with the observation that

$$\int_a^\infty f(x)dx = \lim_{b \to \infty} \int_a^b f(x)dx = \lim_{b \to \infty} F(b) - F(a),$$

(1.22)

where the limit as b approaches ∞ is evaluated after the integral is calculated.

Evaluation of integrals

Unfortunately, unlike differentiation, many integrals cannot be evaluated easily and there are few simple rules which can be used. Some examples of indefinite integrals are given below. Note that u is typically a function $u(x)$, and $du = u'(x)dx$.

$$\int u^n du = \frac{u^{n+1}}{n+1} + c \qquad (n \neq -1)$$
$$\int \frac{du}{u} = \ln|u| + c$$
$$\int a^u du = \frac{a^u}{\ln a} + c$$
$$\int e^u du = e^u + c$$
$$\int \cos u du = \sin u + c$$
$$\int \sin u du = -\cos u + c$$
$$\int \cosh u du = \sinh u + c$$
$$\int \sinh u du = \cosh u + c$$
$$\int \frac{du}{\cos^2 u} = \tan u + c$$
$$\int \frac{du}{\sin^2 u} = -\cot u \quad + c$$
$$\int \frac{du}{u^2 + a^2} = \frac{1}{a} \arctan \frac{u}{a} + c$$
$$\int \frac{du}{\sqrt{a^2 - u^2}} = \arcsin \frac{u}{a} + c$$

Formulae for a large number of integrals can be found in tables of integrals. In order to evaluate unknown integrals, we generally try to transform integrals into forms that are easier to evaluate. For reference, here are a few "techniques of integration" that might help.

- Logarithmic integration: Integrals for which the integrand can be written as the quotient of the derivative of a function, and that same function can be evaluated as

$$\int \frac{f'(x)}{f(x)} dx = \ln|f(x)| + c$$

- Decomposition: When the integrand is a linear combination of integrable functions, we can split the integral of the sum into a sum of simpler integrals:

$$\int \sum_{i=1}^{n} a_i f_i(x) dx = \sum_{i=1}^{n} a_i \int f_i(x) dx$$

- Substitution: If the integrand can be parameterized in terms of a different variable or function $x = u(t)$, we can often utilize the substitution:

$$\int_a^b f(x)dx = \int_{u(a)}^{u(b)} f(u(t))\frac{du(t)}{dt}dt$$

The key to identifying integrals of this form is to find a suitable substitution function.

- Integration by parts: Recall the product rule:

$$\frac{d}{dx}\left[u \cdot v\right] = uv' + u'v$$

Integration by parts enables us to split the integral into parts which are easier to solve. Rearranging equation 1.5 (the product rule)

$$\frac{d}{dx}(uv) = u\frac{dv}{dx} + v\frac{du}{dx} + c$$

to

$$\frac{dv}{dx} = \frac{d}{dx}(uv) - v\frac{du}{dx}.$$

Integrating both sides we obtain

$$\int uv'dx = uv - \int vu'dx.$$

The “art” of solving the integral is to choose the functions u and v so that the remaining integral becomes easier to solve.

Example

Evaluate the integral $\int_a^b x\cos x dx$.

Noting that the integrand is a product of x and $\cos x$, we solve this using integration by parts and choose $u = x$ and $v' = \cos x$, and thus get the result that $v = \sin x$ and $du = dx$. Substituting, we get

$$\begin{aligned}\int_a^b x\cos x dx &= \int_a^b x(\sin x)' dx \\ &= x\ \sin x\Big|_a^b - \int_a^b \sin x dx \\ &= (x\ \sin x + \cos x)|_a^b \\ &= b\ \sin b + \cos b - (a\ \sin a + \cos a)\,.\end{aligned}$$

Example

Evaluate the integral $\int \frac{1}{x^2 + x} dx\,.$

First, we note that the denominator x^2+x can be factored as $x(x+1)$. Using partial fraction decomposition, we get

$$\begin{aligned}\int \frac{1}{x^2 + x} dx &= \int \frac{1}{x(x + 1)} dx \\ &= \int \left(\frac{1}{x} - \frac{1}{x + 1}\right) dx \\ &= \ln x - \ln(x + 1) + c \\ &= \ln\left(\frac{x}{x + 1}\right) + c\end{aligned}$$

where we have split the difference inside the integral into a sum of integrals and used the fact that $\ln\left(\frac{a}{b}\right) = \ln a - \ln b$. However, in general, we need to be careful to consider that the argument of the logarithm is not defined for negative numbers.

Taylor approximation

Taylor's Theorem
The theorem is named after Brook Taylor who expressed this relationship in 1712.

A very useful application of derivatives and integrals is **Taylor's theorem**. Taylor's theorem provides an approximation to a function in the vicinity of a given point x_0 as a sum. Taylor's theorem requires that the function $f(x)$ is continuous and that all of the derivatives $f'(x)$,$f''(x)$, ..., up to order $f^{(n)}(x)$ exist in order to generate an n^{th} degree polynomial approximation to $f(x)$ near x_0. Using equation 1.21 we can express $f(x)$ as

$$\int_a^{a + \epsilon} f'(x) dx = f(a + \epsilon) - f(a) \tag{1.23}$$

where x and $x - \epsilon$ are in the vicinity of a. This can be written as

$$f(a + \epsilon) = f(a) + \int_a^{a + \epsilon} f'(x) dx\,.$$

(1.24)

Assuming that ϵ is very small, we can assume f'(x) is approximately equal to f'(a), and hence

$$f(a+\epsilon) \approx f(a) + \epsilon f'(a)$$

(1.25)

holds. We can express this in terms of x and a, assuming that we stay close to the point a, to get the approximation

$$f(x) \approx f(a) + (x-a)f'(a)\,.$$

(1.26)

The approximation given by equation 1.26 is called the linear approximation to $f(x)$ near $x = a$. It is the tangent line approximation to the function f. By using more information about f, namely by constructing a function that also agrees with f on higher order derivatives at $x = a$, we can obtain an even better approximation. That is the general idea of the Taylor approximation of degree n. Because f is n-differentiable, we can apply the approximation to each of the derivatives of f to obtain

$$\begin{aligned} f'(x) &\approx f'(a) + (x-a)f''(a), \\ f''(x) &\approx f''(a) + (x-a)f'''(a), \end{aligned}$$

and similarly,

$$f^{(n-1)}(x) \approx f^{(n-1)}(a) + (x-a)f^{(n)}(a)\,.$$

We can now substitute the estimate of f'(x) into equation 1.24 and obtain

$$\begin{aligned} f(a+\epsilon) &\approx f(a) + \int_a^{a+\epsilon}[f'(a) + (x-a)f''(a)]dx \\ &= f(a) + \epsilon f'(a) + \frac{\epsilon^2}{2}f''(a)\,. \end{aligned}$$

This process can be repeated iteratively as long as the higher order derivatives exist, which yields the n^{th}-degree Taylor polynomial approximation. Expressing again in terms of x and a, we can write:

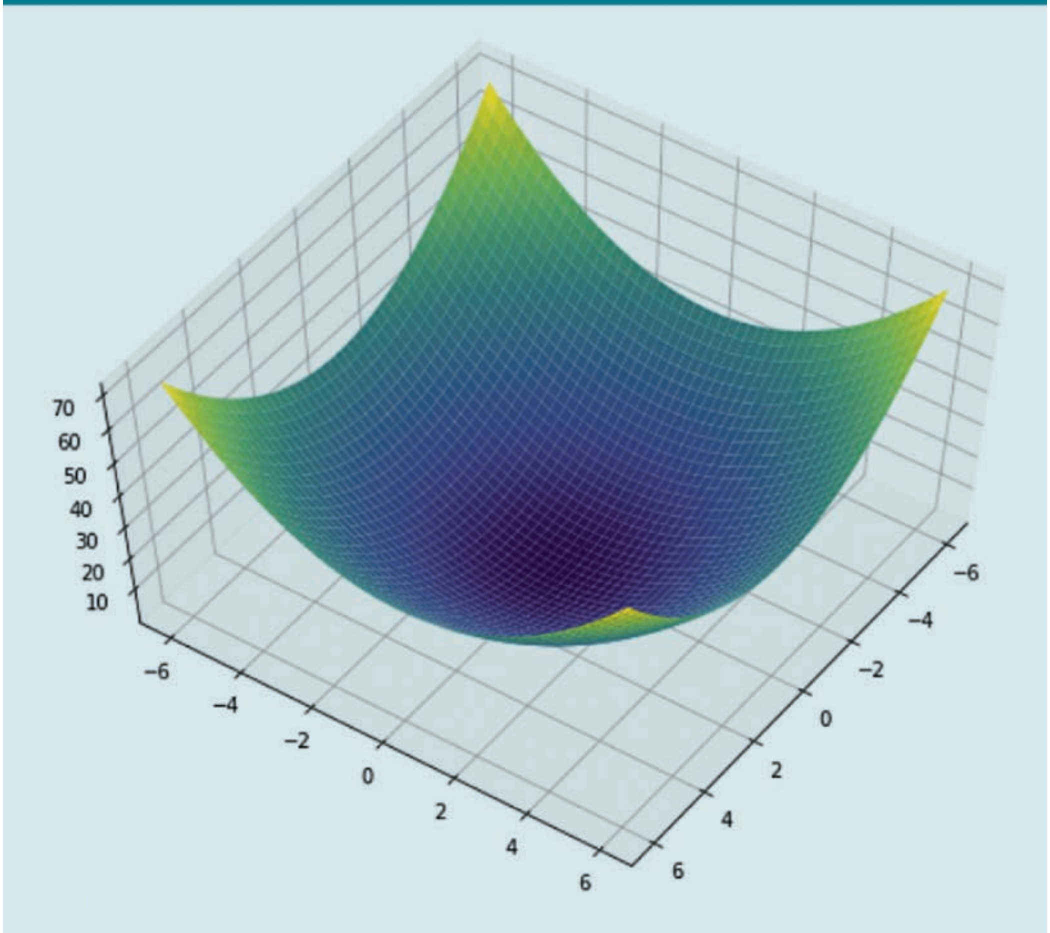
$$f(x) \approx f(a) + (x-a)f'(a) + \frac{(x-a)^2}{2!}f''(a) + \cdots + \frac{(x-a)^n}{n!}f^{(n)}(a)\,.$$

(1.27)

1.2 Partial Differentiation

In the previous section, we considered derivatives of functions of a single variable, i.e. $f^{(n)}(x) = \frac{d^n}{dx}f(x)$. More generally, we can consider rates of change of functions that depend on more than one variable. We can write $f(x_1, x_2, . . . , x_n)$ for a function that depends on n variables $x_1, x_2, . . . , x_n$. An example of a function that depends on two variables x and y, $f(x, y) = x^2 + y^2$, is shown in the following figure. The function is well-defined for each pair (x, y). For example, $f(1, 1)$ = 2.

A Function Depending on Two Variables f(x, y) = x² + y²



Previously, we discussed that the derivative of a function of a single variable is related to the change or gradient of that function. As we consider more variables, we want to know how the function changes as each of the variables change individually, imagining, for example, how x changes as y is held constant. Considering again the function $f(x, y) = x^2 + y^2$, it has a specific gradient in all directions of the xy plane. As a special case, we consider it when we move in either the x or the y direction, for example, along the x or y axis. We move along one direction, for example the x axis, and keep the value of the other variable(s), in this case y, constant as we observe the change of the function. These derivatives are called partial derivatives indicating that we only observe the “partial” change of the function along one of the variables. Similar to the definition of the derivative with respect to a single variable (equation 1.2), we define the partial derivative to be:

$$\frac{\partial f}{\partial x} = \lim_{\Delta x \to 0} \frac{f(x + \Delta x, y) - f(x, y)}{\Delta x},$$

and

(1.28)

$$\frac{\partial f}{\partial y} = \lim_{\Delta y \to 0} \frac{f(x, y + \Delta y) - f(x, y)}{\Delta y},$$

(1.29)

where the symbol ∂ indicates that this differentiation is performed partially with respect to a single variable while the other variables are kept constant. To make this explicit, it is often written as

$$\left(\frac{\partial f}{\partial x}\right)_y \quad \text{and} \quad \left(\frac{\partial f}{\partial y}\right)_x$$

(1.30)

in order to indicate which variable is considered in the derivative (the one in the partial derivative expression) and which is kept constant (the one outside the parentheses). Just as there are many notations for the derivative in the single variable case, there are also many ways to indicate partial derivatives. The following are some common short-hand notations for the partial derivative of f with respect to x:

$$\frac{\partial f}{\partial x} = f_x = \partial_x f\,.$$

(1.31)

One can calculate higher order derivatives, provided that the relevant limits exist, and are calculated in the same way. Some possibilities in the case of two-variables are

$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) = \left(\frac{\partial^2 f}{\partial x^2}\right) = f_{xx},$$
$$\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right) = \left(\frac{\partial^2 f}{\partial y^2}\right) = f_{yy},$$
$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = \left(\frac{\partial^2 f}{\partial x \partial y}\right) = f_{xy}, \text{ and}$$
$$\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = \left(\frac{\partial^2 f}{\partial y \partial x}\right) = f_{yx}.$$

Note that under sufficient continuity conditions, the relation

$$\left(\frac{\partial^2 f}{\partial x \partial y}\right) = \left(\frac{\partial^2 f}{\partial y \partial x}\right)$$

holds.

Example

Find f_x and f_y, the first partial derivatives of $f(x, y) = 3x^2y^2 + y$.

First, we calculate the partial derivative with respect to x, treating y as a constant to obtain

$$\frac{\partial f}{\partial x} = 6xy^2.$$

For the partial derivative with respect to y, we now treat x as a constant and find

$$\frac{\partial f}{\partial y} = 6x^2y + 1.$$

Total Differential

The definition of the partial derivatives allows us to examine the rate of change of a function along, for example, the x or y axes. We now want to investigate the rate of change if we move in any direction in the domain.

In a case where we have functions of two variables x and y, we move Δx in the x direction and Δy in the y direction. Following the approach we have taken previously, we can evaluate

$$
\begin{aligned}
\Delta f &= f(x+\Delta x, y+\Delta y) - f(x,y) \\
&= f(x+\Delta x, y+\Delta y) - f(x, y+\Delta y) + f(x, y+\Delta y) - f(x,y) \\
&= \left[\frac{f(x+\Delta x, y+\Delta y) - f(x, y+\Delta y)}{\Delta x}\right]\Delta x + \\
&\quad \left[\frac{f(x, y+\Delta y) - f(x,y)}{\Delta y}\right]\Delta y
\end{aligned}
$$

where we have performed the algebraic trick of adding and subtracting the same term, namely $-f(x,\ y+\Delta y) + f(x,\ y+\Delta y) = 0$ in the middle step in order to factor into the desired quotients, and we have also multiplied by $\frac{\Delta x}{\Delta x} = 1$ and $\frac{\Delta y}{\Delta y} = 1$. The term in the first square brackets describes the change of the function $f(x,\ y)$ if we move a step Δx in the x direction, the term in the second square bracket corresponds in the y direction. If we let $\Delta x \to 0$ and $\Delta y \to 0$ on both sides, the terms in the square brackets are the partial derivatives defined in equation 1.28, and we obtain the total differential of a function $f(x,\ y)$, which is then given by

$$
df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy\,. \tag{1.32}
$$

For functions of n variables, the formula above is extended accordingly to

$$
df = \frac{\partial f}{\partial x_1}dx_1 + \frac{\partial f}{\partial x_2}dx_2 + \cdots + \frac{\partial f}{\partial x_n}dx_n\,. \tag{1.33}
$$

Chain Rule

When a function of a single variable could be expressed as a composition of functions, we used the chain rule (recall equation 1.6) to differentiate it. The same approach can be applied to functions with several variables.

For example, in the case of a function $f(x,\ y)$, the variables x and y are now functions of another variable u and we wish to find the derivative with respect to u, i.e. $\frac{df}{du}$. Starting from the total derivative in equation 1.32 we obtain

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

The same approach can be taken if the functions are nested in more than one level, i.e. instead of $f(u(x))$ one might have $f(u(v(x)))$, and the chain rule can be used to calculated the derivative, e.g.

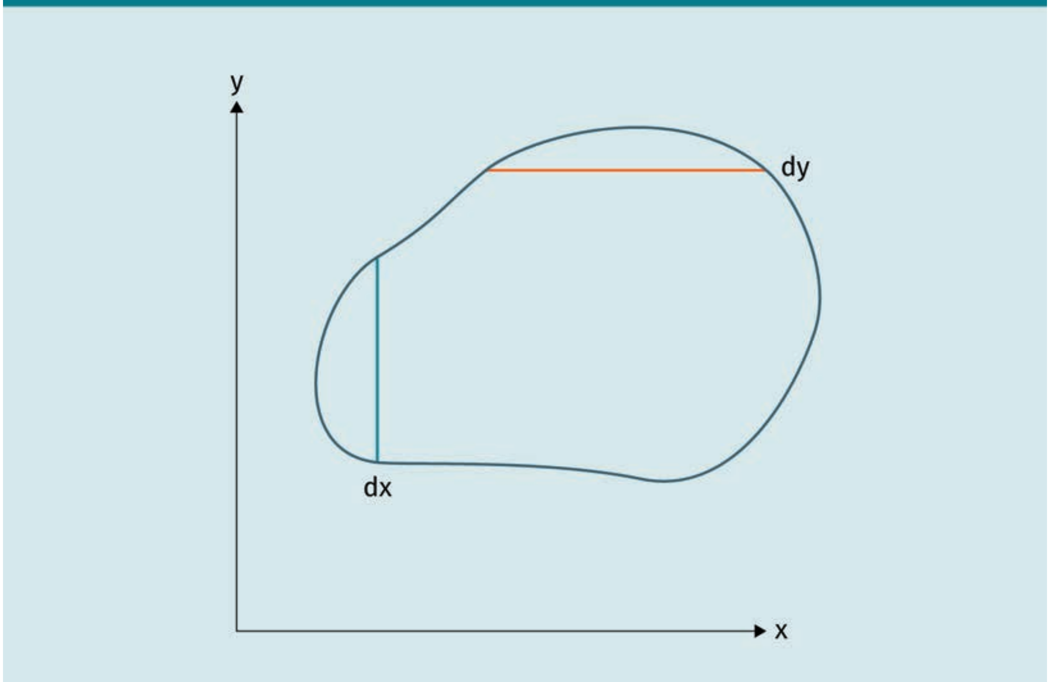
$$
\frac{df(u(v(x)))}{dx} = \frac{\partial f}{\partial u}\frac{\partial u}{\partial v}\frac{dv}{dx}\,.
$$

(1.34)

1.3 Multiple Integrals

Previously in this unit, we introduced integrals over functions of a single variable. Recall that the definite integral measures the area under the curve given by 28 function $f(x)$ over the interval $[a, b]$. We later extended this explanation to encompass indefinite integration. Furthermore, we interpreted the integration as the inverse of differentiation.

Schematic View of a Function f(x,y) which is to be Integrated



Again, in the multivariate case, we will approach integration as a limit of approximations, focusing on the case of two variables x and y first. We wish to find the volume enclosed by the x, y-plane and the function $f(x, y)$ with specific bounds in the x- and y- directions, represented by a region R enclosed by a contour C. Following the approach described previously in this unit, we divide the area R inside the curve into N areas of ΔA_p with $p = 1, 2, \ldots, N$ and define the sum

$$S = \sum_{p=1}^{N} f(x_p, y_p) \Delta A_p$$

to express the approximate volume, where ΔA_p is the area of the base and $f(x_p,\, y_p)$ is the height of cell p. Again, we consider many areas like this, i.e. we let $N \to \infty$, implying $\Delta A_p \to 0$. Similar to the case of a single variable, if the above sum has a finite limit or value, we say that this limit is the value of the double integral over $f(x,\, y)$ over some region R:

$$I = \int_R f(x, y)dA,$$

(1.35)

where dA is an infinitesimally small area in the $x,\, y$ plane where the function $f(x,\, y)$ is evaluated. So far, we have not made any assumption about the small area ΔA considered in the above sum. If we choose small rectangles in x and y direction, we can write $\Delta A = \Delta x \Delta y$, and when $\Delta x \to 0$ and $\Delta y \to 0$, we can write

$$I = \int\int_R f(x, y)dxdy$$

(1.36)

as the double integral. For such integrals, it sometimes matters whether we integrate with respect to x or to y first. It is frequently helpful to draw a picture to see which variable could be taken more easily to depend on the other. If x can be easily expressed as a function of y, we might choose to take small areas in the direction of width dy first. That gives us

$$I = \int_{y\,=\,c}^{y\,=\,d}\left[\int_{x\,=\,x_1(y)}^{x\,=\,x_2(y)} f(x, y)dx\right]dy\,.$$

(1.37)

In this case, the bounds of the inner integral are the parametrization of the boundary curve C, expressed as $x = x_1(y)$ and $x = x_2(y)$. In the first step, y is treated as a constant as the inner integral over x is evaluated. The next step of the computation, the outer integral, is evaluated between the bounds $y = c$ and $y = d$ just as in the single variable case as there are no x’s left in the expression.

Alternatively, we can first evaluate the integral over y and then over x, as

$$I = \int_{x\,=\,a}^{x\,=\,b}\left[\int_{y\,=\,y_1(x)}^{y\,=\,y_2(x)} f(x, y)dy\right]dx\,.$$

(1.38)

Example

Evaluate the integral $\mathrm{I} = \int\int_{\mathrm{R}} \mathrm{x}^2\mathrm{y}\,\mathrm{dxdy}$ where R is given by a triangular area bounded by $\mathrm{x} = 0$, $\mathrm{y} = 0$, $\mathrm{x} + \mathrm{y} = 1$.

First, we carry out the integration over y, which means that we keep x fixed. In this case, the limits on y are $\mathrm{y} = 0$ and $\mathrm{y} = 1$ - x. Given the constraint $\mathrm{x}+\mathrm{y} = 1$, the maximum value of x is $\mathrm{x} = 1$ for $\mathrm{y} = 0$. The integral is then written as

$$\mathrm{I} = \int_{\mathrm{x}\,=\,0}^{\mathrm{x}\,=\,1}\left[\int_{\mathrm{y}\,=\,0}^{\mathrm{y}\,=\,1\,-\,\mathrm{x}} \mathrm{x}^2\mathrm{ydy}\right]\mathrm{dx}\,.$$

We evaluate the inner integral first, treating x as a constant, to get

$$\int_{\mathrm{y}\,=\,0}^{\mathrm{y}\,=\,1\,-\,\mathrm{x}} \mathrm{x}^2\mathrm{ydy} = \frac{1}{2}\mathrm{x}^2\mathrm{y}^2\Big|_{\mathrm{y}\,=\,0}^{\mathrm{y}\,=\,1\,-\,\mathrm{x}} = \frac{1}{2}\mathrm{x}^2(1-\mathrm{x})^2\,.$$

This result is now inserted into the outer integral as follows:

$$\begin{aligned}
&\int_{\mathrm{x}\,=\,0}^{\mathrm{x}\,=\,1}\frac{1}{2}\mathrm{x}^2(1-\mathrm{x})^2\mathrm{dx}\\
&= \frac{1}{2}\int_{\mathrm{x}\,=\,0}^{\mathrm{x}\,=\,1}\mathrm{x}^2\mathrm{dx} - \frac{1}{2}\int_{\mathrm{x}\,=\,0}^{\mathrm{x}\,=\,1}2\mathrm{x}^3\mathrm{dx} + \frac{1}{2}\int_{\mathrm{x}\,=\,0}^{\mathrm{x}\,=\,1}\mathrm{x}^4\mathrm{dx}\\
&= \frac{1}{2}\frac{1}{3}\mathrm{x}^3\Big|_0^1 - \frac{1}{4}\mathrm{x}^4\Big|_0^1 + \frac{1}{2}\frac{1}{5}\mathrm{x}^5\Big|_0^1\\
&= \frac{1}{6} - \frac{1}{4} + \frac{1}{10}\\
&= \frac{1}{60}\,.
\end{aligned}$$

In case of more than two variables, the same notation can be extended accordingly, such as

$$\int\int\int_{\mathrm{V}}\mathrm{f}(\mathrm{x},\mathrm{y},\mathrm{z})\mathrm{dxdydz} \tag{1.39}$$

where, in the case of three variables, we integrate over a specific volume rather than an area.

1.4 Calculus of Variations

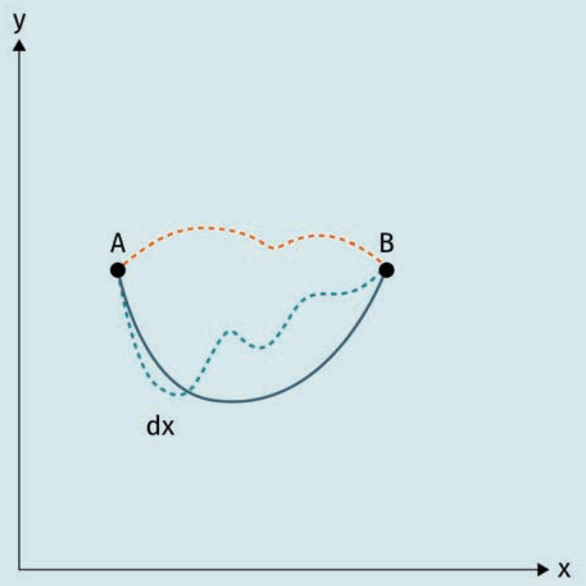
Previously, we introduced the idea of local extrema and how to use stationary points to find them. We can apply the same ideas to more than one variable. In fact, we can even extend this idea to look for input functions that give extrema (maxima and minima), rather than input values that give extrema.

This is the idea behind calculus of variations. In most cases, we want to minimize or maximize a given quantity that depends on a family of input functions; the calculus of variations provides a method for finding a function $f(x)$ which yields the extreme value.

As a concrete example, we could imagine a rope that is attached to two points, A and B, as shown in the following figure, but otherwise hangs freely under the influence of **gravity**. We expect that the rope will hang down in a shape such as the one indicated by the solid line, and not take any other shape (e.g., those suggested by the two dotted lines) — that is, as long as there is no external force other than gravity, and all initial motion has come to a rest. In this example, the rope is fixed at the points A and B, so we have two constraints (not including the length of the rope) which we take as constant. As the gravitational force acts on each part of the rope, the rope will take the shape where the total potential energy, expressed by the integral over all small segments of the rope, is minimal. We wish to find the function $y(x)$ be the function that describes the shape of the hanging rope with the minimal potential energy.

Gravity
Gravity is one of the natural forces caused by the mass of objects, resulting in them being pulled towards each other.

Illustration of the Concept of Calculus of Variations with a Rope Hanging From Points A and B



To introduce the calculus of variations, we start with the integral

$$I = \int_a^b F(y, y', x) dx,$$

(1.40)

where a, b, and F are given by the nature of the problem we wish to consider. This integral depends on the function $y(x)$. In the example of the rope, the limits a and b of the integral are fixed: they correspond to the endpoints of the rope at which the rope is attached, for example, to two poles.

We call such functions (ones that take other functions as their input and result in a scalar as their output) functionals. Here, I is a functional of $y(x)$, which we denote by

$$I = I[y(x)] \, .$$

(1.41)

We use square brackets to indicate that I is a functional rather than a function of $\mathbb{R}^n$. We then look for the curves $y(x)$, which are the stationary value(s) of the integral I, and determine whether such curves are extrema of the integral. The integral may have one or more stationary points.

A stationary point $y(x)$ of the functional $I[y(x)]$ is a point where the functional I does not change if the $y(x)$ is perturbed by a small amount. In the case of the rope, this would be the function that describes the physical shape the rope takes if we fix it at two points and let it hang under the influence of gravity. Because $y(x)$ is a stationary point of the integral $I[y(x)]$, if we change

$$y(x) \to y(x) + \epsilon\eta(x)$$

(1.42)

by a small amount ε using any (sufficiently well-behaved) function $\eta(x)$, we require that the value of I does not change, i.e.,

$$\left.\frac{dI}{d\epsilon}\right|_{\epsilon\,=\,0} = 0 \forall \eta(x) \, .$$

(1.43)

We now insert the above equation 1.42 into the integral in equation 1.40:

$$I[y(x), \epsilon] = \int_a^b F\Big(y + \epsilon\eta, y' + \epsilon\eta', x\Big) dx \, .$$

We generally assume that all functions are well behaved, especially when considering situations related to physical examples.

Taylor Series with Multiple Variables

We have already encountered Taylor series for the case of the single variable in Eqn.(1.27) and used this to expand a function into a series around some point. This approach can be generalized to several variables. For example, for a function that depends on two variables x and y, we can write the corresponding second degree Taylor polynomial as:

$$\begin{aligned} f(x,y) = f(x_0,y_0) &+ \frac{\partial f}{\partial x}\Delta x + \frac{\partial f}{\partial y}\Delta y \\ &+ \frac{1}{2!}\left[\frac{\partial^2 f}{\partial x^2}(\Delta x)^2 + 2\frac{\partial^2 f}{\partial x \partial y}\Delta x \Delta y + \frac{\partial^2 f}{\partial y^2}(\Delta y)^2\right] \end{aligned}$$

where we evaluate the derivatives around some point (x_0,y_0)

and $\Delta x = x - x_0$ and $\Delta y = y - y_0$.

We can write this as:

$$\begin{aligned} f(x,y) = f(x_0,y_0) &+ \left(\Delta x\frac{\partial}{\partial x} + \Delta y\frac{\partial}{\partial y}\right)f(x,y) \\ &+ \frac{1}{2!}\left(\Delta x\frac{\partial}{\partial x} + \Delta y\frac{\partial}{\partial y}\right)^2 f(x,y) \end{aligned}$$

Extending to higher derivatives, we can write the Taylor series for a function of two variables as:

$$f(x,y) = \sum_{n=0}^{\infty}\frac{1}{n!}\left[\left(\Delta x\frac{\partial}{\partial x} + \Delta y\frac{\partial}{\partial y}\right)^n f\left(x,y\right)\right]_{(x_0,\,y_0)}$$

We can further generalize this to any number of variables denoted by the vector $\vec{x}$:

$$f\left(\vec{x}\right) = f\left(\vec{x}_0\right) + \sum_i \frac{\partial f}{\partial x_i} + \frac{1}{2!}\sum_{i,j}\frac{\partial^2 f}{\partial x_i \partial x_j}\Delta x_i \Delta x_j + \cdots$$

Returning to the calculus of variations and the integral $I[y(x),\ \varepsilon]$, we can use the Taylor series with $\Delta y = \epsilon\eta$ and $\Delta y' = \epsilon\eta'$ and write the integral in the following way:

$$\begin{aligned} I[y(x),\epsilon] &= F(y+\epsilon\eta, y'+\epsilon\eta', x)dx \\ &= \int_a^b F(y,y',x)dx + \int_a^b\left(\frac{\partial F}{\partial y}\epsilon\eta + \frac{\partial F}{\partial y'}\epsilon\eta'\right)dx + \mathcal{O}(\epsilon^2)\,. \end{aligned}$$

(1.44)

In the following, we ignore all terms of order ε^2 and higher because ε is assumed to be a very small number. This means we consider the equation

$$I[y(x), \epsilon] = \int_a^b F(y, y', x)dx + \int_a^b \left(\frac{\partial F}{\partial y}\epsilon\eta + \frac{\partial F}{\partial y'}\epsilon\eta'\right)dx\,.$$

Now, recall that when we introduced the small perturbation in $y(x)$ in Equation 1.42, we said that this should not change the integral because we are at a stationary point. We expressed this more formally in Equation 1.43, where we demand that the integral I does not change if we change y a little bit by the term $\varepsilon\eta(x)$ for any choice of η.

This then implies that the second term must be equal to zero for any choice of $\eta(x)$, because ε is a small (but non-zero) number and we do not make any demands of the function $\eta(x)$ except that it be sufficiently well behaved, so we can take its derivative, integrate it, and so on. Then, because we demand that this holds for any small perturbation, the second part in the equation above must vanish, which we can write as

$$\delta I = \int_a^b \left(\frac{\partial F}{\partial y}\eta + \frac{\partial F}{\partial y'}\eta'\right)dx = 0,$$

(1.45)

where the notation δI is used to indicate the variation in the functional $I[y(x)]$ due to the change in $y(x) \rightarrow y(x) + \varepsilon\eta(x)$. Furthermore, ε is a small but non-zero number and can therefore be omitted from the above equation.

We now integrate the second part of the integral by parts, resulting in

$$\int_a^b \frac{\partial F}{\partial y'}\eta' dx = \frac{\partial F}{\partial y'}\eta\bigg|_a^b - \int_a^b \eta\frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right),$$

(1.46)

so the integral equation becomes

$$\frac{\partial F}{\partial y'}\eta\bigg|_a^b + \int_a^b \left[\frac{\partial F}{\partial y} - \frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right)\right]\eta(x)dx = 0\,.$$

(1.47)

We now impose the constraint that the endpoints a and b are fixed, as are $y(a)$ and $y(b)$ – recalling our initial example of the freely hanging rope under the influence of gravity, where the rope is fixed at its two attached points.

Since $y(a)$ and $y(b)$ are fixed, we also require that, at these points, $\eta(a) = 0$ and $\eta(b) = 0$: if we "wiggle" the rope a bit, i.e., change $y(x)$, the endpoints remain unchanged. It follows that the first term in the above equation vanishes. Since equation 1.47 must be equal to zero for any choice of $\eta(x)$, this implies that the function in the integral must be zero, namely that

$$\frac{\partial F}{\partial y} = \frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right). \tag{1.48}$$

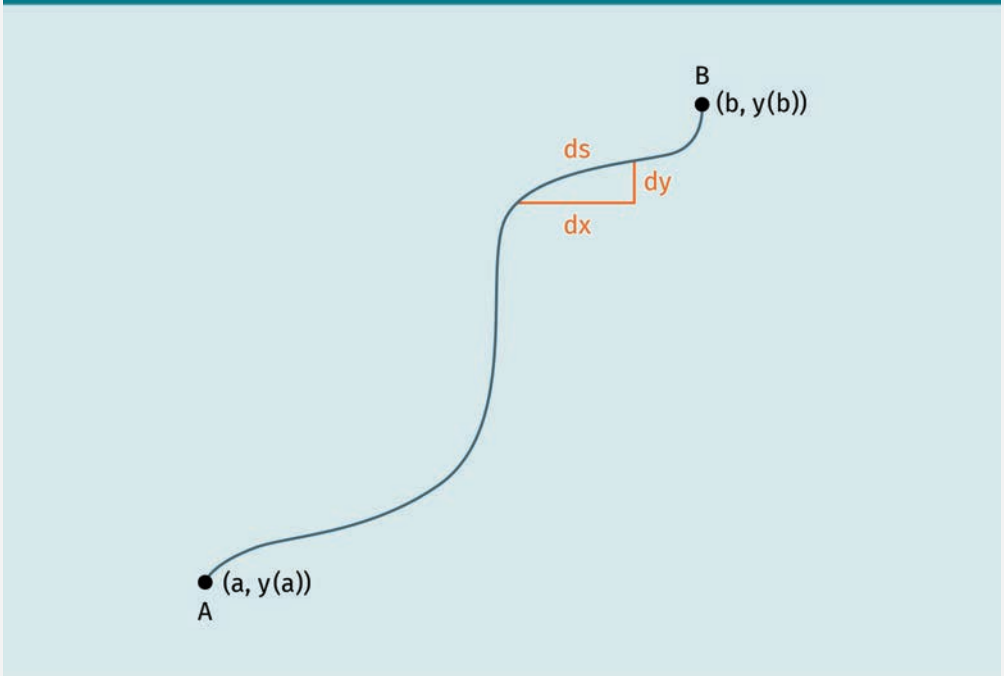
Equation 1.48 is known as the Euler-Lagrange equation.

Example

Show that the shortest path between two points is a straight line.

We start by specifying the initial and final points that will be connected with an arbitrary path; initial point A is given by the coordinates $(a, y(a))$ and the final point B, is given by the coordinates $(b, y(b))$, as shown below:

Two Points A, B Connected By a Path



For any small segment of the path, the length can be approximated by a straight line using the distance formula

$$ds = \sqrt{(dx)^2 + (dy)^2},$$

where we assume that dx and dy are small enough to justify the approximation of the small triangle for ds. Factoring out dx, the equation above can be written as

$$ds = \sqrt{1 + (y')^2}dx\,. \tag{1.49}$$

The total length of the line is given by the integral

$$L = \int_a^b ds = \int_a^b \sqrt{1 + y'^2}dx, \tag{1.50}$$

where the integration takes place along the path between the two points. We now calculate the path which leads to a stationary point for L, in this case a minimum which gives the shortest connection between the points A and B. We start from the Euler-Lagrange equation 1.48 and note that the function in the integral L does not depend on y explicitly. This implies that

$$\frac{\partial F}{\partial y} = 0,$$

so the Euler-Langrange equation can be written as

$$\frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right) = 0, \tag{1.51}$$

which in turn implies that

$$\frac{\partial F}{\partial y'} = c, \tag{1.52}$$

for some constant c. We now take the derivative of the function $\sqrt{1 + (y')^2}$ with respect to y' and obtain

$$c = \frac{\partial F}{\partial y'} = \frac{y'}{\sqrt{1 + (y')^2}}, \tag{1.53}$$

recalling that $\sqrt{w}$ can be written as $w^{1/2}$.

We now solve the equation

$$c = \frac{y'}{\sqrt{1 + (y')^2}}$$

for dy so that we can integrate both sides and obtain an explicit formula for y as follows:

$$c = \frac{y'}{\sqrt{1 + (y')^2}} \tag{1.54}$$

$$c^2 = \frac{(y')^2}{1 + (y')^2} \tag{1.55}$$

$$c^2\left(1 + y'^2\right) = (y')^2 \tag{1.56}$$

$$c^2 = (y')^2 - c^2(y')^2 \tag{1.57}$$

$$= (1 - c^2)(y')^2 \tag{1.58}$$

$$c = \sqrt{1 - c^2}\frac{dy}{dx} \tag{1.59}$$

$$\frac{c}{\sqrt{1 - c^2}}dx = dy\,. \tag{1.60}$$

Integrating both sides yields

$$y = \frac{c}{\sqrt{1 - c^2}}x + k \tag{1.61}$$

for some constant k by noting that the term $\frac{c}{\sqrt{1 - c^2}}$ is constant and $\int dx = x$.

As expected, the above equation is indeed a straight line of the form $y = mx + b$ with $m = \frac{c}{\sqrt{1-c^2}}$ and constant $k = b$.

Summary

In this unit, we have seen functions of a single variable $f(x)$, as well as multivariate functions, such as $f(x, y)$. Differentiation is a tool for studying the rate of change of a function with respect to a given variable. In the case of multivariate functions, the partial derivatives indicate how much the function changes along the x- or y-axis, for example, while the total differential extends this idea to the rate of change of a function in any arbitrary direction. Integration of the functions of one variable was introduced as the area enclosed by the function and can be interpreted as the inverse of the derivative. The integral is therefore often called the antiderivative. The Taylor expansion can be used to approximate a given function at a specific point. Finally, the calculus of variation extends the concepts of differentiation and integration to functions whose inputs are themselves functions.

Knowledge Check

Did you understand this unit?

You can check your understanding by completing the questions for this unit on the learning platform.

Good luck!