Elementary Calculus Manual

Alfredo Sánchez Alberca (asalber@ceu.es)

Sep 2016

Department of Applied Math and Statistics
CEU San Pablo

License terms

Contents

Analytic geometry

Some phenomena of Nature can be described by a number and a unit of measurement.

Definition (Scalar)

A *scalar* is a number that expresses a magnitude without direction.

Examples The height or weight of a person, the temperature of a gas or the time it takes a vehicle to travel a distance.

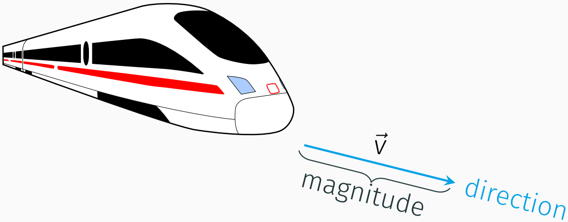
However, there are other phenomena that cannot be described adequately by a scalar. If, for instance, a sailor wants to head for seaport and only knows the intensity of wind, he will not know what direction to take. The description of wind requires two elements: intensity and direction.

Definition (Vector)

A *vector* is a number that expresses a magnitude and has associated an orientation and a sense.

Examples The velocity of a vehicle or the force applied to an object.

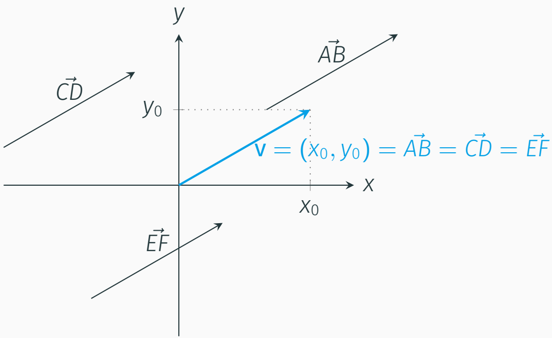
Geometrically, a vector is represented by an directed line segment, that is, an arrow.



Vector representation

An oriented segment can be located in different places in a Cartesian space. However, regardless of where it is located, if the length and the direction of the segment does not change, the segment represents always the same vector.

This allows to represent all vectors with the same origin, the origin of the Cartesian coordinate system. Thus, a vector can be represented by the Cartesian *coordinates* of its final end in any Euclidean space.

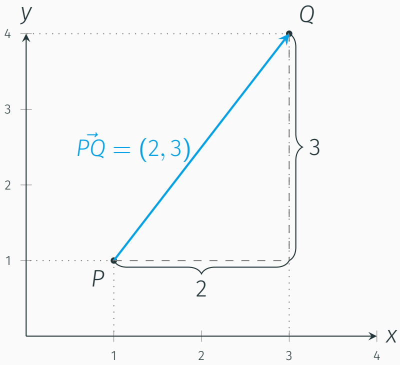


Vector from two points

Given two points P and Q of a Cartesian space, the vector that starts at P and ends at Q has coordinates $\vec{PQ} = Q - P$.

Example Given the points $P = (1, 1)$ and $Q = (3, 4)$ in the real plane $\mathbb{R}^2$, the coordinates of the vector that start at P and ends at Q are

$$\vec{PQ} = Q - P = (3, 4) - (1, 1) = (3 - 2, 4 - 1) = (2, 3).$$



Definition (Module of a vector)

Given a vector $\mathbf{v} = (v_1, \cdots, v_n)$ in $\mathbb{R}^n$, the *module* of $\mathbf{v}$ is

$$|\mathbf{v}| = \sqrt{v_1^2 + \cdots + v_n^2}.$$

The module of a vector coincides with the length of the segment that represents the vector.

Examples Let $\mathbf{u} = (3, 4)$ be a vector in $\mathbb{R}^2$, then its module is

$$|\mathbf{u}| = \sqrt{3^2 + 4^2} = \sqrt{25} = 5$$

Let $\mathbf{v} = (4, 7, 4)$ be a vector in $\mathbb{R}^3$, then its module is

$$|\mathbf{v}| = \sqrt{4^2 + 7^2 + 4^2} = \sqrt{81} = 9$$

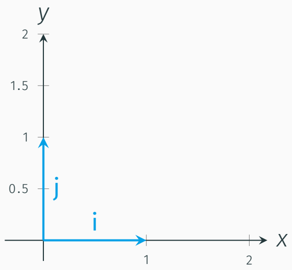
Definition (Unit vector)

A vector $\mathbf{v}$ in $\mathbb{R}^n$ is a *unit vector* if its module is one, that is, $|\mathbf{v}| = 1$.

The unit vectors with the direction of the coordinate axes are of special importance and they form the *standard basis*.

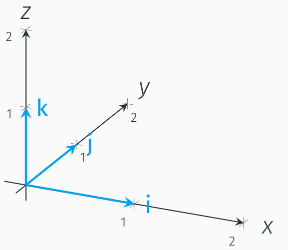
In $\mathbb{R}^2$ the standard basis is formed by two vectors

$$\mathbf{i} = (1, 0) \text{ and } \mathbf{j} = (0, 1)$$



In $\mathbb{R}^3$ the standard basis is formed by three vectors

$$\mathbf{i} = (1, 0, 0), \mathbf{j} = (0, 1, 0) \text{ and } \mathbf{k} = (0, 0, 1)$$



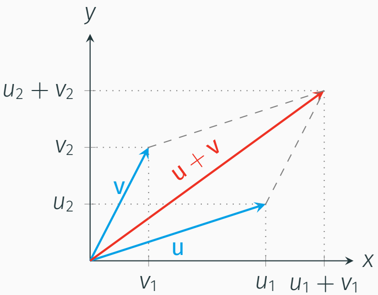
Definition (Sum of two vectors)

Given two vectors $\mathbf{u} = (u_1, \cdots, u_n)$ y $\mathbf{v} = (v_1, \cdots, v_n)$ de $\mathbb{R}^n$, the *sum* of $\mathbf{u}$ and $\mathbf{v}$ is

$$\mathbf{u} + \mathbf{v} = (u_1 + v_1, \dots, u_n + v_n).$$

Example Let $\mathbf{u} = (3, 1)$ and $\mathbf{v} = (2, 3)$ two vectors in $\mathbb{R}^2$, then the sum of them is

$$\mathbf{u} + \mathbf{v} = (3 + 2, 1 + 3) = (5, 4).$$



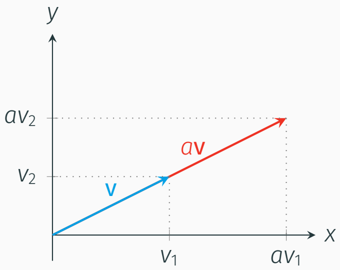
Definition (Product of a vector by a scalar)

Given a vector $\mathbf{v} = (v_1, \cdots, v_n)$ in $\mathbb{R}^n$, and a scalar $a \in \mathbb{R}$, the *product* of $\mathbf{v}$ by a is

$$a\mathbf{v} = (av_1, \ldots, av_n).$$

Example Let $\mathbf{v} = (2, 1)$ a vector in $\mathbb{R}^2$ and $a = 2$ a scalar, then the product of a by $\mathbf{v}$ is

$$a\mathbf{v} = 2(2, 1) = (4, 2).$$

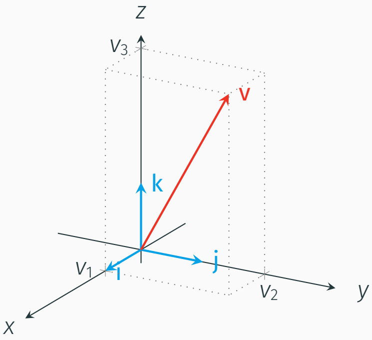


Expressing a vector as a linear combination of the standard basis

The sum of vectors and the product of vector by a scalar allow us to express any vector as a linear combination of the standard basis.

In $\mathbb{R}^3$, for instance, a vector with coordinates $\mathbf{v} = (v_1, v_2, v_3)$ can be expressed as the linear combination

$$\mathbf{v} = (v_1, v_2, v_3) = v_1\mathbf{i} + v_2\mathbf{j} + v_3\mathbf{k}.$$



Definition (Dot product of two vectors)

Given the vectors $\mathbf{u} = (u_1, \cdots, u_n)$ and $\mathbf{v} = (v_1, \cdots, v_n)$ in $\mathbb{R}^n$, the *dot product* of $\mathbf{u}$ and $\mathbf{v}$ is

$$\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + \cdots + u_n v_n.$$

Example Let $\mathbf{u} = (3, 1)$ and $\mathbf{v} = (2, 3)$ two vectors in $\mathbb{R}^2$, then the dot product of them is

$$\mathbf{u} \cdot \mathbf{v} = 3 \cdot 2 + 1 \cdot 3 = 9.$$

It holds that

$$\mathbf{u} \cdot \mathbf{v} = |\mathbf{u}||\mathbf{v}| \cos \alpha$$

where α is the angle between the vectors.

Definition (Parallel vectors)

Two vectors $\mathbf{u}$ and $\mathbf{v}$ are *parallel* if there is a scalar $a \in \mathbb{R}$ such that

$$\mathbf{u} = a\mathbf{v}.$$

Example The vectors $\mathbf{u} = (-4, 2)$ and $\mathbf{v} = (2, -1)$ in $\mathbb{R}^2$ are parallel, as there is a scalar -2 such that

$$\mathbf{u} = (-4, 2) = -2(2, -1) = -2\mathbf{v}.$$

Definition (Orthogonal and orthonormal vectors)

Two vectors $\mathbf{u}$ and $\mathbf{v}$ are *orthogonal* if their dot product is zero,

$$\mathbf{u} \cdot \mathbf{v} = 0.$$

If in addition both vectors are unit vectors, $|\mathbf{u}| = |\mathbf{v}| = 1$, then the vectors are *orthonormal*.

Orthogonal vectors are perpendicular, that is the angle between them is right.

Example The vectors $\mathbf{u} = (2, 1)$ and $\mathbf{v} = (-2, 4)$ in $\mathbb{R}^2$ are orthogonal, as

$$\mathbf{uv} = 2 \cdot -2 + 1 \cdot 4 = 0,$$

but they are not orthonormal since $|\mathbf{u}| = \sqrt{2^2 + 1^2} \neq 1$ and $|\mathbf{v}| = \sqrt{-2^2 + 4^2} \neq 1$.

The vectors $\mathbf{i} = (1, 0)$ and $\mathbf{j} = (0, 1)$ in $\mathbb{R}^2$ are orthonormal, as

$$\mathbf{ij} = 1 \cdot 0 + 0 \cdot 1 = 0, \quad |\mathbf{i}| = \sqrt{1^2 + 0^2} = 1, \quad |\mathbf{j}| = \sqrt{0^2 + 1^2} = 1.$$

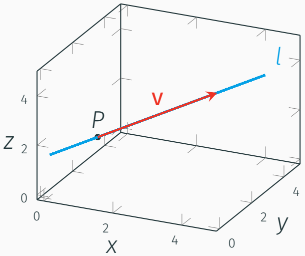
Definition (Vectorial equation of a straight line)

Given a point $P = (p_1, \ldots, p_n)$ and a vector $\mathbf{v} = (v_1, \ldots, v_n)$ of $\mathbb{R}^n$, the *vectorial equation of the line l* that passes through the point P with the direction of $\mathbf{v}$ is

$$l : X = P + t\mathbf{v} = (p_1, \ldots, p_n) + t(v_1, \ldots, v_n) = (p_1 + tv_1, \ldots, p_n + tv_n), \quad t \in \mathbb{R}.$$

Example Let l the line of $\mathbb{R}^3$ that goes through $P = (1, 1, 2)$ with the direction of $\mathbf{v} = (3, 1, 2)$, then the vectorial equation of l is

$$l : X = P + t\mathbf{v} = (1, 1, 2) + t(3, 1, 2) =$$
$$= (1 + 3t, 1 + t, 2 + 2t) \quad t \in \mathbb{R}.$$



Parametric and Cartesian equations of a line

From the vectorial equation of a line $l : X = P + t\mathbf{v} = (p_1 + tv_1, \dots, p_n + tv_n)$ is easy to obtain the coordinates of the the points of the line with n *parametric equations*

$$x_1(t) = p_1 + tv_1, \dots, x_n(t) = p_n + tv_n$$

from where, if $\mathbf{v}$ is a vector with non-null coordinates ($v_i \neq 0\ \forall i$), we can solve for t and equal the equations getting the *Cartesian equations*

$$\frac{x_1 - p_1}{v_1} = \cdots = \frac{x_n - p_n}{v_n}$$

Example Given a line with vectorial equation $l : X = (1, 1, 2) + t(3, 1, 2) = (1 + 3t, 1 + t, 2 + 2t)$ in $\mathbb{R}^{\nvDash}$, its parametric equations are

$$x(t) = 1 + 3t, \quad y(t) = 1 + t, \quad z(t) = 2 + 2t,$$

and the Cartesian equations are

$$\frac{x - 1}{3} = \frac{y - 1}{1} = \frac{z - 2}{2}$$

Point-slope equation of a line in the plane

In the particular case of the real plane $\mathbb{R}^2$, if we have a line with vectorial equation $l : X = P + t\mathbf{v} = (x_0, y_0) + t(a, b) = (x_0 + ta, y_0 + tb)$, its parametric equations are

$$x(t) = x_0 + ta, \quad y(t) = y_0 + tb$$

and its Cartesian equation is

$$\frac{x - x_0}{a} = \frac{y - y_0}{b}.$$

From this, moving b to the other side of the equation, we get

$$y - y_0 = \frac{b}{a}(x - x_0),$$

or renaming $m = b/a$,

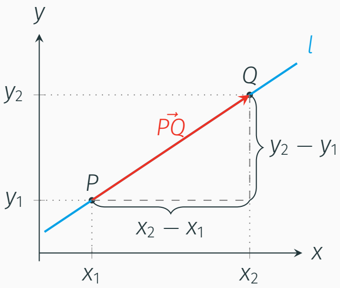
$$y - y_0 = m(x - x_0).$$

This equation is known as the *point-slope equation* of the line.

Definition (Slope of a line in the plane)

Given a line $l : X = P + t\mathbf{v}$ in the real plane $\mathbb{R}^2$, with direction vector $\mathbf{v} = (a, b)$, the *slope* of l is b/a.

Recall that given two points $P = (x_1, y_1)$ and $Q = (x_2, y_2)$ on the line l, we can take as a direction vector the vector from P to Q, with coordinates $\vec{PQ} = Q - P = (x_2 - x_1, y_2 - y_1)$. Thus, the slope of l is $\dfrac{y_2 - y_1}{x_2 - x_1}$, that is, the ratio between the changes in the vertical and horizontal axes.

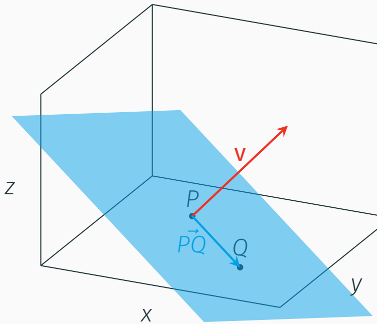


Vector equation of a plane in space

To get the equation of a plane in the real space $\mathbb{R}^3$ we can take a point of the plane $P = (x_0, y_0, z_0)$ and an orthogonal vector to the plane $\mathbf{v} = (a, b, c)$. Then, any point $Q = (x, y, z)$ of the plane satisfies that the vector $\vec{PQ} = (x - x_0, y - y_0, z - z_0)$ is orthogonal to $\mathbf{v}$, and therefore their dot product is zero.

$$\vec{PQ} \cdot \mathbf{v} = (x - x_0, y - y_0, z - z_0)(a, b, c) = a(x - x_0) + b(y - y_0) + c(z - z_0) = 0.$$

This equation is known as the *vector equation of the plane*.



From the vector equation of the plane we can get

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0 \Leftrightarrow ax + by + cz = ax_0 + by_0 + cz_0,$$

that, renaming $d = ax_0 + by_0 + cz_0$, can be written as

$$ax + by + cz = d,$$

and is known as the *scalar equation of the plane*.

Example Given the point $P = (2, 1, 1)$ and the vector $\mathbf{v} = (2, 1, 2)$, the vector equation of the plane that passes through P and is orthogonal to $\mathbf{v}$ is

$$(x - 2, y - 1, z - 1)(2, 1, 2) = 2(x - 2) + (y - 1) + 2(z - 1) = 0,$$

and its scalar equation is

$$2x + y + 2z = 7.$$

Differential calculus with one real variable

Definition (Increment of a variable)

An *increment* of a variable x is a change in the value of the variable; is it is denoted Δx. The increment of a variable x along an interval $[a, b]$ is given by

$$\Delta x = b - a.$$

Definition (Increment of a function)

The *increment* of a function $y = f(x)$ along an interval $[a, b] \subseteq Dom(f)$ is given by

$$\Delta y = f(b) - f(a).$$

Example The increment of x along the interval $[2, 5]$ is $\Delta x = 5 - 2 = 3$ and the increment of the function $y = x^2$ along the same interval is $\Delta y = 5^2 - 2^2 = 21$.

The study of a function $y = f(x)$ requires to understand how the function changes, that is, how the dependent variable y changes when we change the independent variable x.

Definition (Average rate of change)

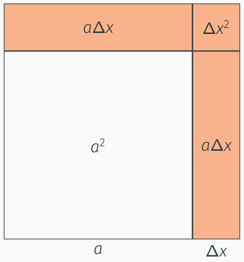
The *average rate of change* of a function $y = f(x)$ in an interval $[a, a+\Delta x] \subseteq Dom(f)$, is the quotient between the increment of y and the increment of x in that interval, and is denoted by

$$\text{ARC } f[a, a+\Delta x] = \frac{\Delta y}{\Delta x} = \frac{f(a+\Delta x) - f(a)}{\Delta x}.$$

Let $y = x^2$ be the function that measures the area of a metallic square of side length x.

If at any given time the side of the square is a, and we heat the square uniformly increasing the side by dilatation a quantity Δx, how much will the area of the square increase?

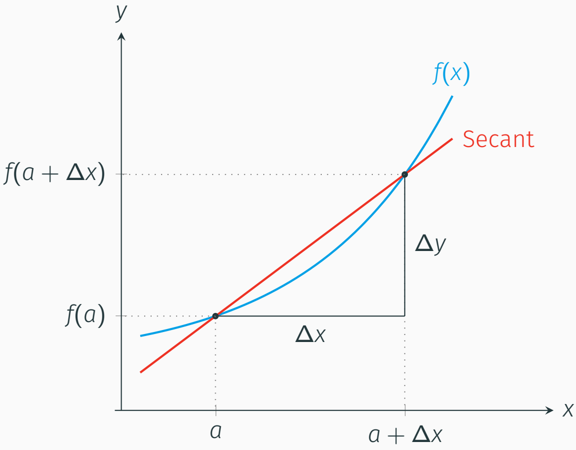
$$\begin{aligned}\Delta y = f(a+\Delta x) - f(a) = (a+\Delta x)^2 - a^2 = \\ = a^2 + 2a\Delta x + \Delta x^2 - a^2 = 2a\Delta x + \Delta x^2.\end{aligned}$$



What is the average rate of change in the interval $[a, a+\Delta x]$?

$$\text{ARC } f[a, a+\Delta x] = \frac{\Delta y}{\Delta x} = \frac{2a\Delta x + \Delta x^2}{\Delta x} = 2a + \Delta x.$$

The average rate of change of a function $y = f(x)$ in an interval $[a, a + \Delta x]$ is the slope of the *secant* line to the graph of f through the points $(a, f(a))$ and $(a + \Delta x, f(a + \Delta x))$.



Often it is interesting to study the rate of change of a function, not in an interval, but in a point.

Knowing the tendency of change of a function in an instant can be used to predict the value of the function in nearby instants.

Definition (Instantaneous rate of change and derivative)

The *instantaneous rate of change* of a function $f(x)$ at a point $x = a$, is the limit of the average rate of change of f in the interval $[a, a + \Delta x]$, when Δx tends to 0, and is denoted by

$$\mathrm{IRC}\, f(a) = \lim_{\Delta x \to 0} \mathrm{ARC}\, f[a, a + \Delta x] = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \to 0} \frac{f(a + \Delta x) - f(a)}{\Delta x}$$

When this limit exists, the function f is said to be *differentiable* at the point a, and its value is called the *derivative* of f at a, and it is denoted $f'(a)$ (Lagrange's notation) or $\frac{df}{dx}(a)$ (Leibniz's notation).

Let's take again the function $y = x^2$ that measures the area of a metallic square of side length x.

If at any given time the side of the square is a, and we heat the square uniformly increasing the side, what is the tendency of change of the area in that moment?

$$\begin{aligned}\mathrm{IRC}\, f(a) &= \lim_{\Delta x\to 0}\frac{\Delta y}{\Delta x} = \lim_{\Delta x\to 0}\frac{f(a+\Delta x)-f(a)}{\Delta x} = \\ &= \lim_{\Delta x\to 0}\frac{2a\Delta x+\Delta x^2}{\Delta x} = \lim_{\Delta x\to 0} 2a+\Delta x = 2a.\end{aligned}$$

Thus,

$$f'(a) = \frac{df}{dx}(a) = 2a,$$

indicating that the area of the square tends to increase the double of the side.

The derivative of a function $f'(a)$ shows the growth rate of f at point a:

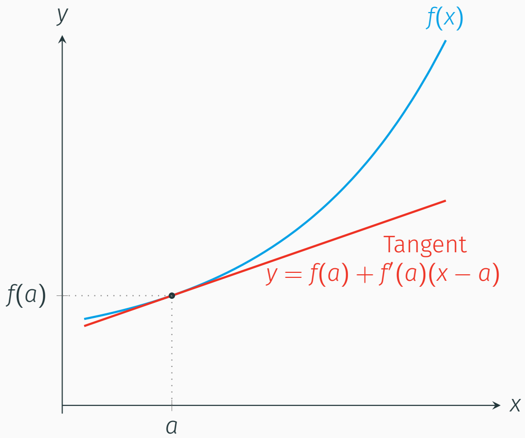
- $f'(a) > 0$ indicates an increasing tendency (y increases as x increases).
- $f'(a) < 0$ indicates a decreasing tendency (y decreases as x increases).

Example A derivative $f'(a) = 3$ indicates that y tends to increase triple of x at point a. A derivative $f'(a) = -0.5$ indicates that y tends to decrease half of x at point a.

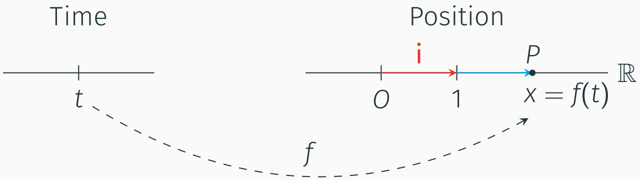
Geometric interpretation of the derivative

The instantaneous rate of change or derivative of a function $y = f(x)$ at $x = a$ is the slope of the *tangent line* to the graph of f at point $(a, f(a))$. Thus, the equation of the tangent line to the graph of f at the point $(a, f(a))$ is

$$y - f(a) = f'(a)(x - a) \Leftrightarrow y = f(a) + f'(a)(x - a)$$



Assume that the function $y = f(t)$ describes the position of an object moving in the real line at time t. Taking as reference the coordinates origin O and the unitary vector $\mathbf{i} = (1)$, we can represent the position of the moving object P at every moment t with a vector $\vec{OP} = x\mathbf{i}$ where $x = f(t)$.



Remark It also makes sense when f measures other magnitudes as the temperature of a body, the concentration of a gas, or the quantity of substance in a chemical reaction at every moment t.

Kinematic interpretation of the average rate of change

In this context, if we take the instants $t = a$ and $t = a + \Delta t$, both in $\text{Dom}(f)$, the vector

$$\mathbf{v}_m = \frac{f(a + \Delta t) - f(a)}{\Delta t}$$

is known as the *average velocity* of the trajectory f in the interval $[a, a + \Delta t]$.

Example A vehicle makes a trip from Madrid to Barcelona. Let $f(t)$ be the function that determine the position of the vehicle at every moment t. If the vehicle departs from Madrid (km 0) at 8:00 and arrives at Barcelona (km 600) at 14:00, then the average velocity of the vehicle in the path is

$$\mathbf{v}_m = \frac{f(14) - f(8)}{14 - 8} = \frac{600 - 0}{6} = 100 km/h.$$

In the same context of the linear motion, the derivative of the function $f(t)$ at the moment t_0 is the vector

$$\mathsf{v} = f'(a) = \lim_{\Delta t \to 0} \frac{f(a + \Delta t) - f(a)}{\Delta t},$$

that is known, as long as the limit exists, as the *instantaneous velocity* or simply *velocity* of the trajectory f at moment a.

That is, the derivative of the object position with respect to time is a vector field that is called *velocity along the trajectory f*.

Example Following with the previous example, what indicates the speedometer at any instant is the modulus of the instantaneous velocity vector at that moment.

If $y = c$, is a constant function, then $y' = 0$ at any point.

If $y = x$, is the identity function, then $y' = 1$ at any point.

If $u = f(x)$ and $v = g(x)$ are two differentiable functions, then

- $(u + v)' = u' + v'$
- $(u - v)' = u' - v'$
- $(u \cdot v)' = u' \cdot v + u \cdot v'$
- $\left(\frac{u}{v}\right)' = \frac{u' \cdot v - u \cdot v'}{v^2}$

Theorem (Chain rule)

If the function $y = f \circ g$ is the composition of two functions $y = f(z)$ and $z = g(x)$, then

$$(f \circ g)'(x) = f'(g(x))g'(x).$$

It is easy to prove this fact using the Leibniz notation

$$\frac{dy}{dx} = \frac{dy}{dz}\frac{dz}{dx} = f'(z)g'(x) = f'(g(x))g'(x).$$

Example If $f(z) = \sin z$ and $g(x) = x^2$, then $f \circ g(x) = \sin(x^2)$. Applying the chain rule the derivative of the composite function is

$$(f \circ g)'(x) = f'(g(x))g'(x) = \cos(g(x))2x = \cos(x^2)2x.$$

On the other hand, $g \circ f(z) = (\sin z)^2$, and applying the chain rule again, its derivative is

$$(g \circ f)'(z) = g'(f(z))f'(z) = 2f(z)\cos z = 2\sin z \cos z.$$

Theorem (Derivative of the inverse function)

Given a function $y = f(x)$ with inverse $x = f^{-1}(y)$, then

$$\left(f^{-1}\right)'(y) = \frac{1}{f'(x)} = \frac{1}{f'(f^{-1}(y))},$$

provided that f is differentiable at $f^{-1}(y)$ and $f'(f^{-1}(y)) \neq 0$.

Again, it is easy to prove this equality using the Leibniz notation

$$\frac{dx}{dy} = \frac{1}{dy/dx} = \frac{1}{f'(x)} = \frac{1}{f'(f^{-1}(y))}$$

The inverse of the exponential function $y = f(x) = e^x$ is the natural logarithm $x = f^{-1}(y) = \ln y$, so we can compute the derivative of the natural logarithm using the previous theorem and we get

$$\left(f^{-1}\right)'(y) = \frac{1}{f'(x)} = \frac{1}{e^x} = \frac{1}{e^{\ln y}} = \frac{1}{y}.$$

Example Sometimes it is easier to apply the chain rule to compute the derivative of the inverse of a function. In this example, as $\ln x$ is the inverse of e^x, we know that $e^{\ln x} = x$, so differentiating both sides and applying the chain rule to the left side we get

$$(e^{\ln x})' = x' \Leftrightarrow e^{\ln x}(\ln(x))' = 1 \Leftrightarrow (\ln(x))' = \frac{1}{e^{\ln x}} = \frac{1}{x}.$$

The main application of derivatives is to determine the variation (increase or decrease) of functions. For that we use the sign of the first derivative.

Theorem

Let $f(x)$ be a function with first derivative in an interval $I \subseteq \mathbb{R}$.

- *If $\forall x \in I\ f'(x) > 0$ then f is increasing on I.*
- *If $\forall x \in I\ f'(x) < 0$ then f is decreasing on I.*

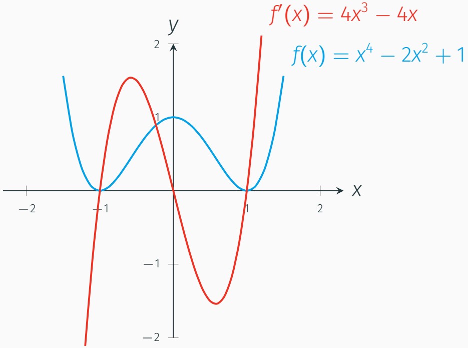
If $f'(a) = 0$ then a is known as a *critical point* or *stationary point*. At this point the function can be increasing, decreasing or neither increasing nor decreasing.

Example The function $f(x) = x^2$ has derivative $f'(x) = 2x$; it is decreasing on $\mathbb{R}^-$ as $f'(x) < 0\ \forall x \in \mathbb{R}^-$ and increasing on $\mathbb{R}^+$ as $f'(x) > 0\ \forall x \in \mathbb{R}^+$.

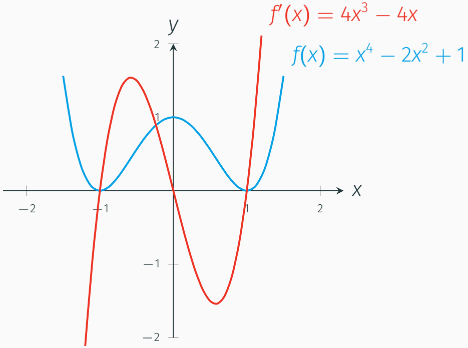
It has a critical point at $x = 0$, as $f'(0) = 0$; at this point the function is neither increasing nor decreasing.

Remark A function can be increasing or decreasing on an interval and not

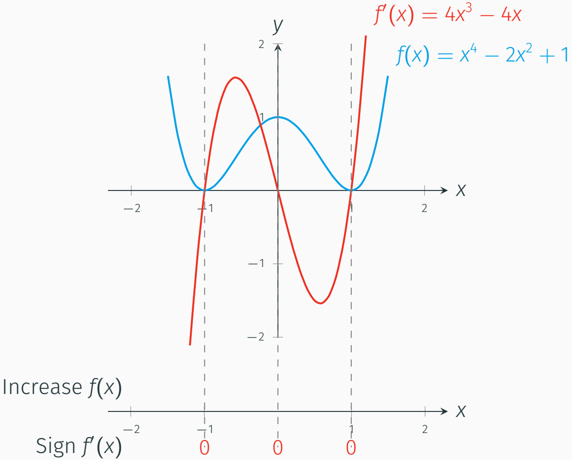
Let us analyze the increase and decrease of the function $f(x) = x^4 - 2x^2 + 1$. Its first derivative is $f'(x) = 4x^3 - 4x$.



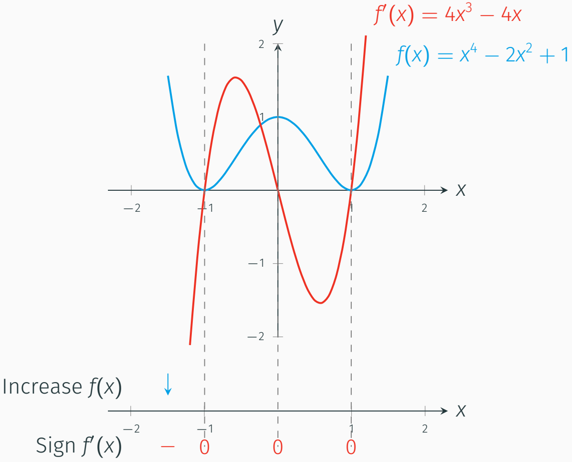
Let us analyze the increase and decrease of the function $f(x) = x^4 - 2x^2 + 1$. Its first derivative is $f'(x) = 4x^3 - 4x$.



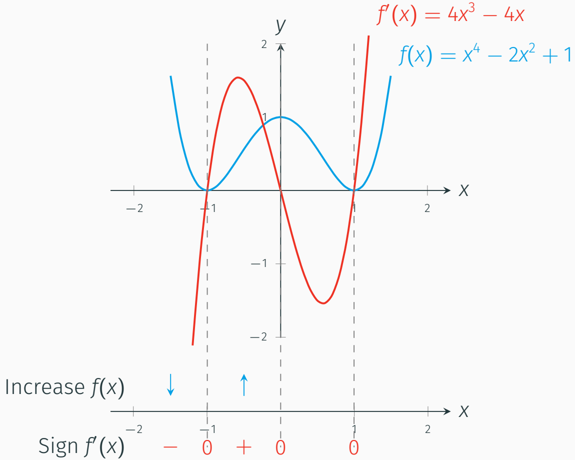
Let us analyze the increase and decrease of the function $f(x) = x^4 - 2x^2 + 1$. Its first derivative is $f'(x) = 4x^3 - 4x$.



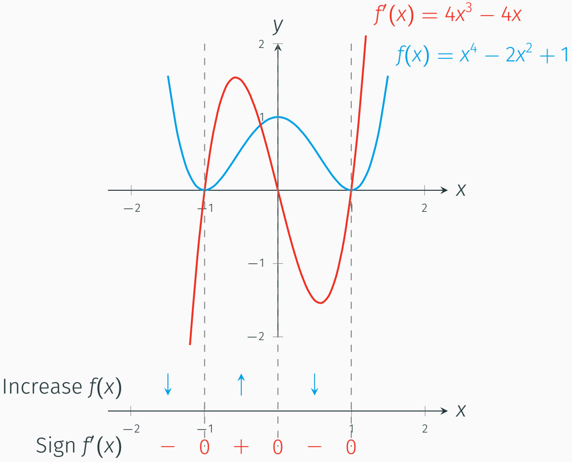
Let us analyze the increase and decrease of the function $f(x) = x^4 - 2x^2 + 1$. Its first derivative is $f'(x) = 4x^3 - 4x$.



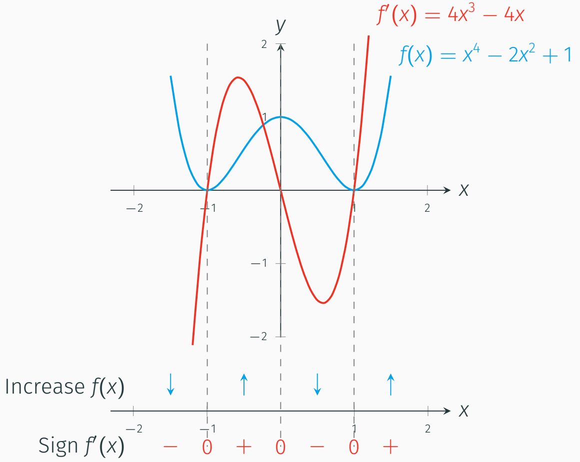
Let us analyze the increase and decrease of the function $f(x) = x^4 - 2x^2 + 1$. Its first derivative is $f'(x) = 4x^3 - 4x$.



Let us analyze the increase and decrease of the function $f(x) = x^4 - 2x^2 + 1$. Its first derivative is $f'(x) = 4x^3 - 4x$.



Let us analyze the increase and decrease of the function $f(x) = x^4 - 2x^2 + 1$. Its first derivative is $f'(x) = 4x^3 - 4x$.



As a consequence of the previous result we can also use the first derivative to determine the relative extrema of a function.

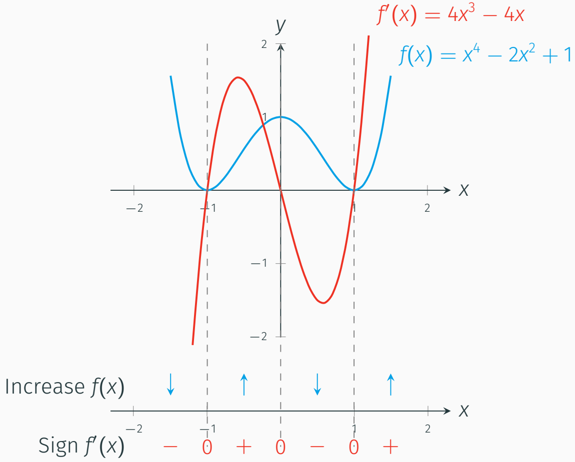
Theorem (First derivative test)

Let $f(x)$ be a function with first derivative in an interval $I \subseteq \mathbb{R}$ and let $a \in I$ be a critical point of f ($f'(a) = 0$).

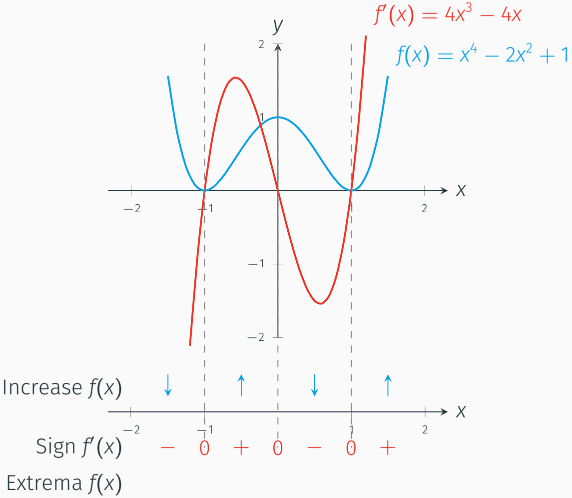
- *If $f'(x) > 0$ on an open interval extending left from a and $f'(x) < 0$ on an open interval extending right from a, then f has a relative maximum at a.*
- *If $f'(x) < 0$ on an open interval extending left from a and $f'(x) > 0$ on an open interval extending right from a, then f has a relative minimum at a.*
- *If $f'(x)$ has the same sign on both an open interval extending left from a and an open interval extending right from a, then f has an inflection point at a.*

Remark *A vanishing derivative is a necessary but not sufficient condition*

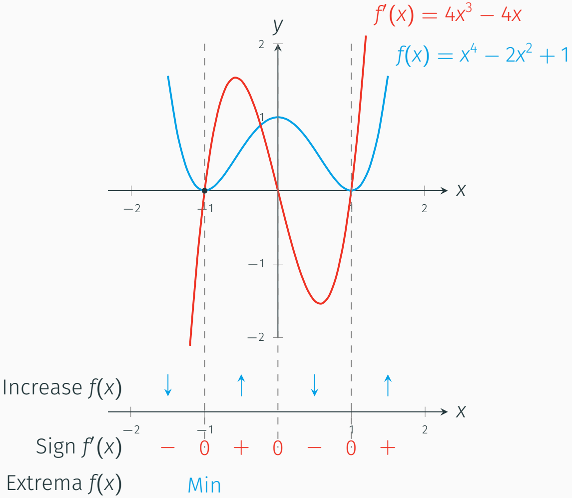
Consider again the function $f(x) = x^4 - 2x^2 + 1$ and let's analyze its relative extrema now. Its first derivative is $f'(x) = 4x^3 - 4x$.



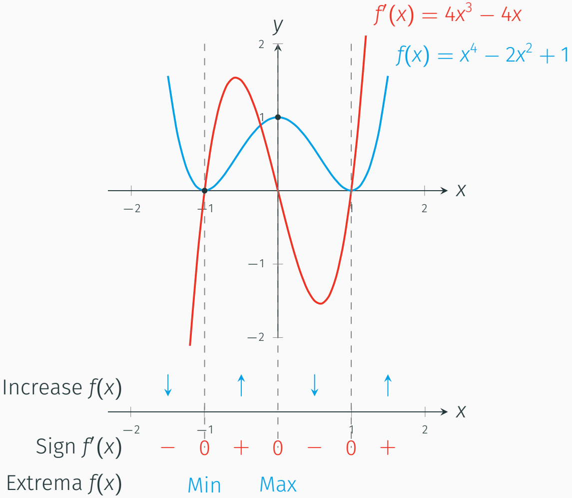
Consider again the function $f(x) = x^4 - 2x^2 + 1$ and let's analyze its relative extrema now. Its first derivative is $f'(x) = 4x^3 - 4x$.



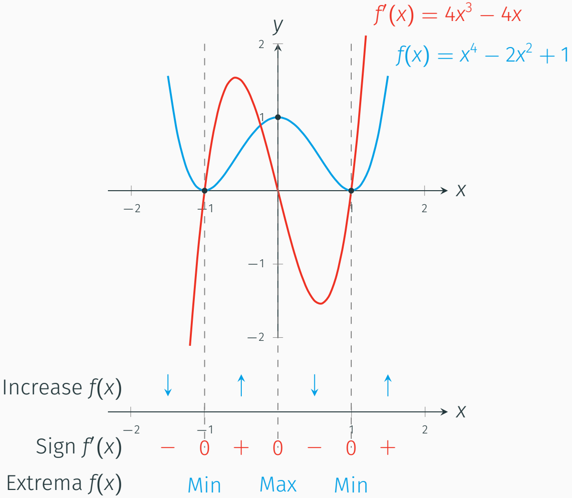
Consider again the function $f(x) = x^4 - 2x^2 + 1$ and let's analyze its relative extrema now. Its first derivative is $f'(x) = 4x^3 - 4x$.



Consider again the function $f(x) = x^4 - 2x^2 + 1$ and let's analyze its relative extrema now. Its first derivative is $f'(x) = 4x^3 - 4x$.



Consider again the function $f(x) = x^4 - 2x^2 + 1$ and let's analyze its relative extrema now. Its first derivative is $f'(x) = 4x^3 - 4x$.



The concavity of a function can be determined by de second derivative.

Theorem

Let $f(x)$ be a function with second derivative in an interval $I \subseteq \mathbb{R}$.

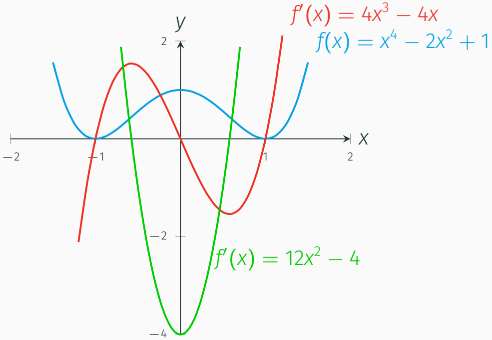
- *If $\forall x \in I$ $f''(x) > 0$ then f is concave up (convex) on I.*
- *If $\forall x \in I$ $f''(x) < 0$ then f is concave down (concave) on I.*

Example The function $f(x) = x^2$ has second derivative $f''(x) = 2 > 0$ $\forall x \in \mathbb{R}$, so it is concave up in all $\mathbb{R}$.

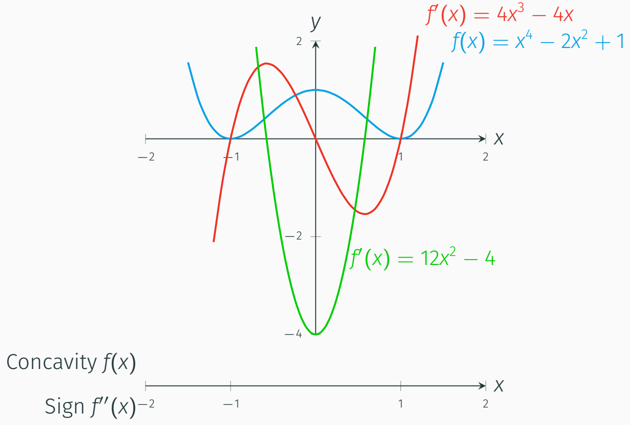
Remark *A function can be concave up or down and not have second derivative.*

Analysis of functions: concavity
Example

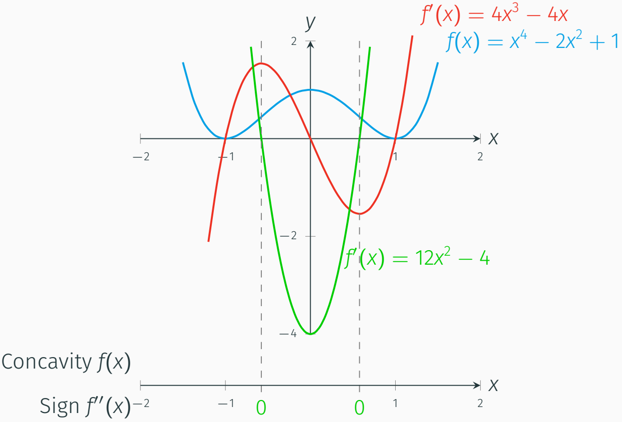
Let us analyze the concavity of the same function of previous examples $f(x) = x^4 - 2x^2 + 1$. Its second derivative is $f''(x) = 12x^2 - 4$.



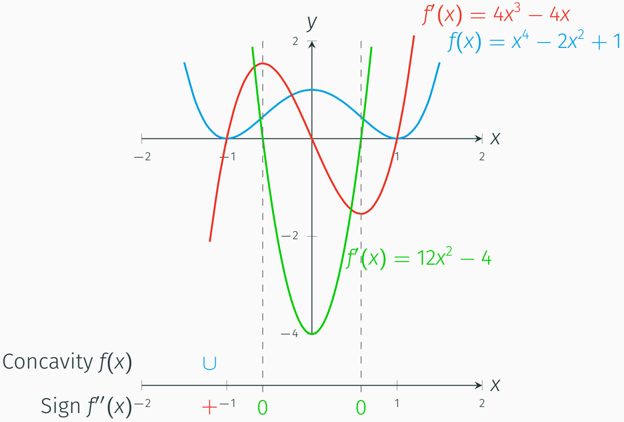
Let us analyze the concavity of the same function of previous examples $f(x) = x^4 - 2x^2 + 1$. Its second derivative is $f''(x) = 12x^2 - 4$.



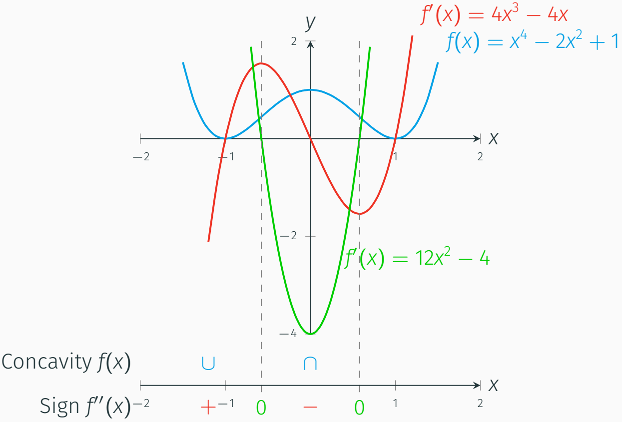
Let us analyze the concavity of the same function of previous examples $f(x) = x^4 - 2x^2 + 1$. Its second derivative is $f''(x) = 12x^2 - 4$.



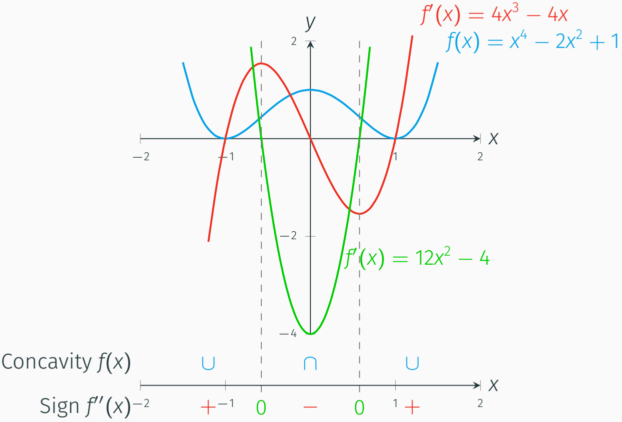
Let us analyze the concavity of the same function of previous examples $f(x) = x^4 - 2x^2 + 1$. Its second derivative is $f''(x) = 12x^2 - 4$.



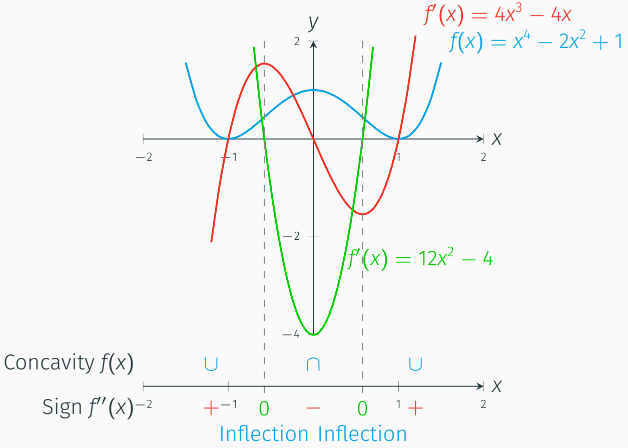
Let us analyze the concavity of the same function of previous examples $f(x) = x^4 - 2x^2 + 1$. Its second derivative is $f''(x) = 12x^2 - 4$.



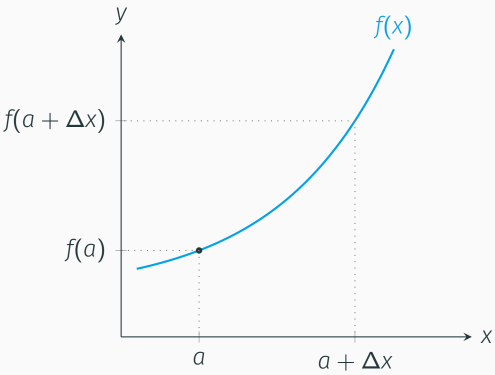
Let us analyze the concavity of the same function of previous examples $f(x) = x^4 - 2x^2 + 1$. Its second derivative is $f''(x) = 12x^2 - 4$.



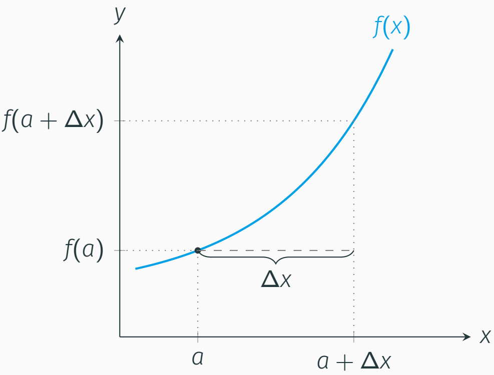
Let us analyze the concavity of the same function of previous examples $f(x) = x^4 - 2x^2 + 1$. Its second derivative is $f''(x) = 12x^2 - 4$.



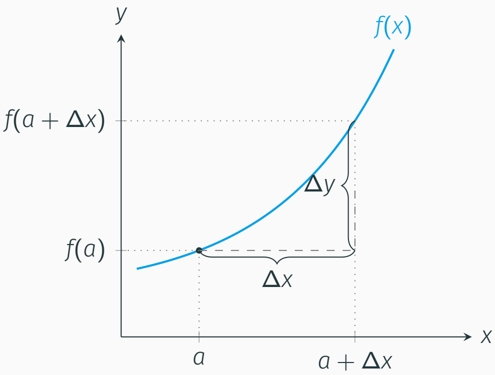
The tangent line to the graph of a function $f(x)$ at $x = a$ can be used to approximate f in a neighbourhood of a.



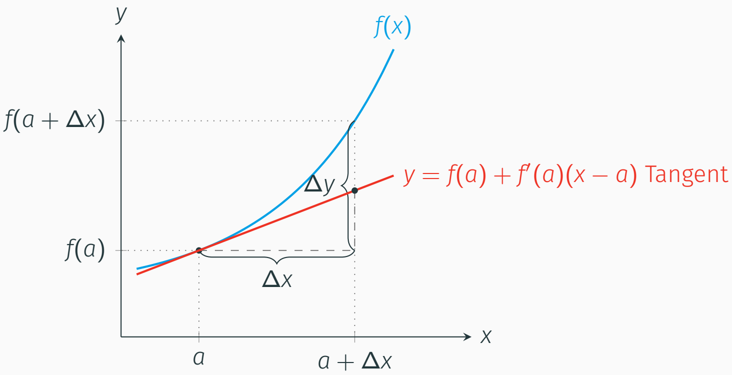
The tangent line to the graph of a function $f(x)$ at $x = a$ can be used to approximate f in a neighbourhood of a.



The tangent line to the graph of a function $f(x)$ at $x = a$ can be used to approximate f in a neighbourhood of a.

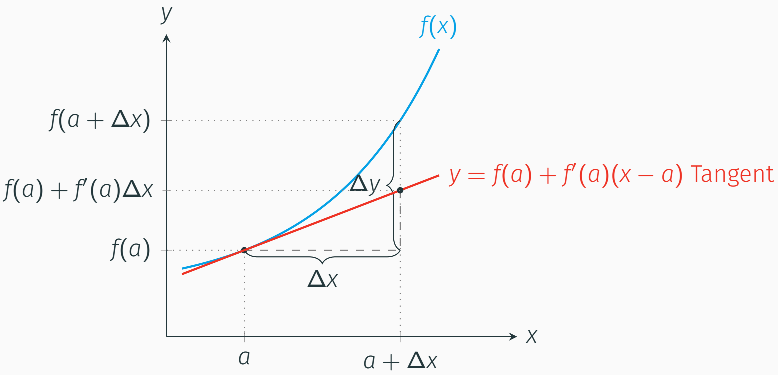


The tangent line to the graph of a function $f(x)$ at $x = a$ can be used to approximate f in a neighbourhood of a.

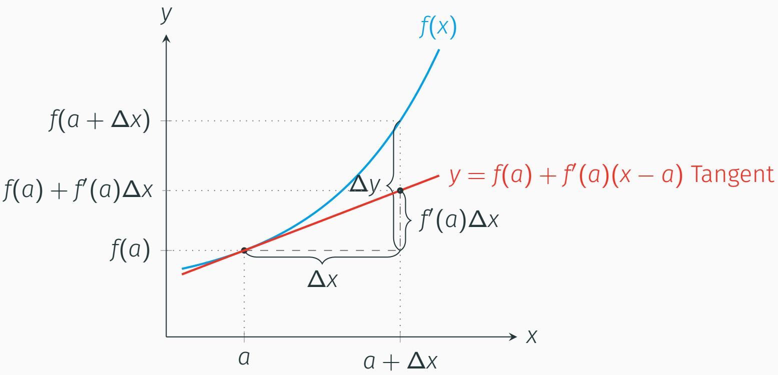


Approximating a function with the derivative

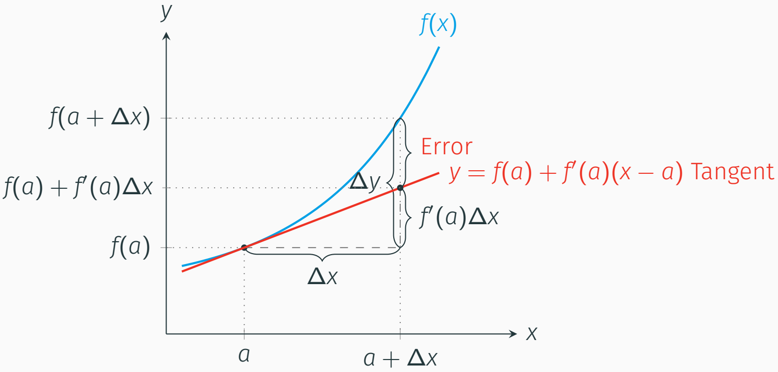
The tangent line to the graph of a function $f(x)$ at $x = a$ can be used to approximate f in a neighbourhood of a.



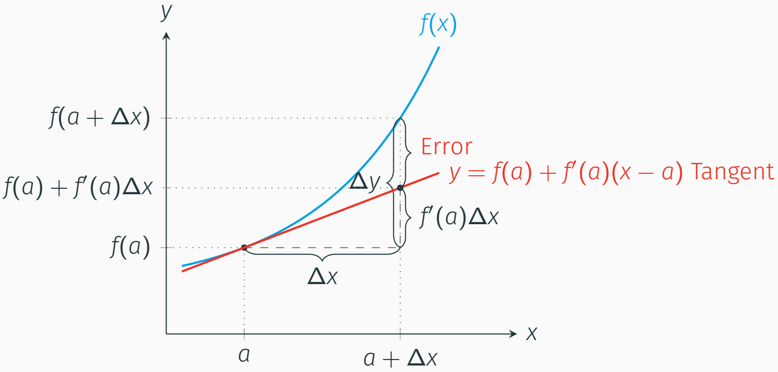
The tangent line to the graph of a function $f(x)$ at $x = a$ can be used to approximate f in a neighbourhood of a.



The tangent line to the graph of a function $f(x)$ at $x = a$ can be used to approximate f in a neighbourhood of a.



The tangent line to the graph of a function $f(x)$ at $x = a$ can be used to approximate f in a neighbourhood of a.



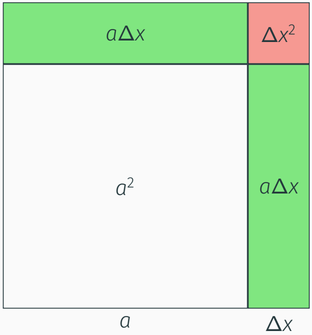
Thus, the increment of a function $f(x)$ in an interval $[a, a + \Delta x]$ can be approximated multiplying the derivative of f at a by the increment of x

$$\Delta y \approx f'(a)\Delta x$$

In the previous example of the function $y = x^2$ that measures the area of a metallic square of side x, if the side of the square is a and we increment it by a quantity Δx, then the increment on the area will be approximately

$$\Delta y \approx f'(a)\Delta x = 2a\Delta x.$$

In the figure below we can see that the error of this approximation is Δx^2, which is smaller than Δx when Δx tends to 0.



Another useful application of the derivative is the approximation of functions by polynomials.

Polynomials are functions easy to calculate (sums and products) with very good properties:

- Defined in all the real numbers.
- Continuous.
- Differentiable of all orders with continuous derivatives.

Goal

Approximate a function $f(x)$ by a polynomial $p(x)$ near a point $x = a$.

Approximating a function by a polynomial of order 0

A polynomial of degree 0 has equation

$$p(x) = c_0,$$

where c_0 is a constant.

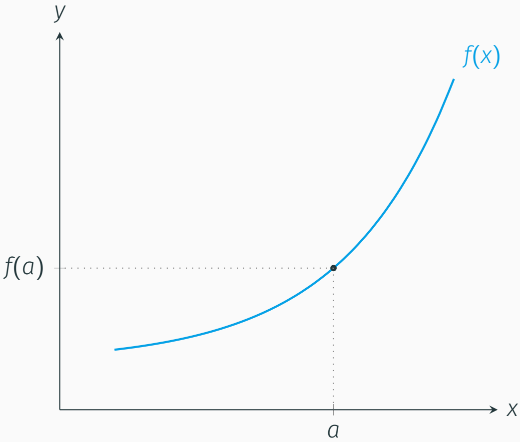
As the polynomial should coincide with the function f at a, it must satisfy

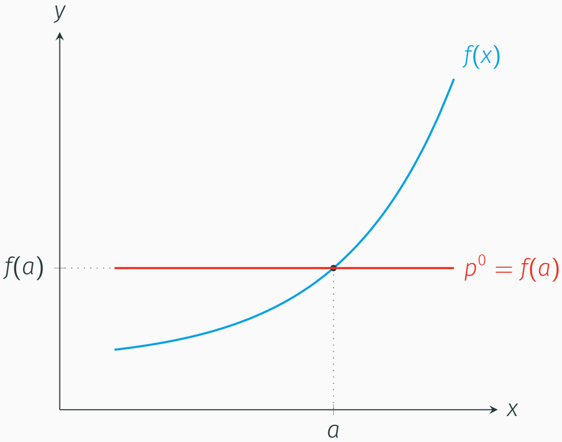
$$p(a) = c_0 = f(a).$$

Therefore, the polynomial of degree 0 that best approximates f near a is

$$p(x) = f(a).$$

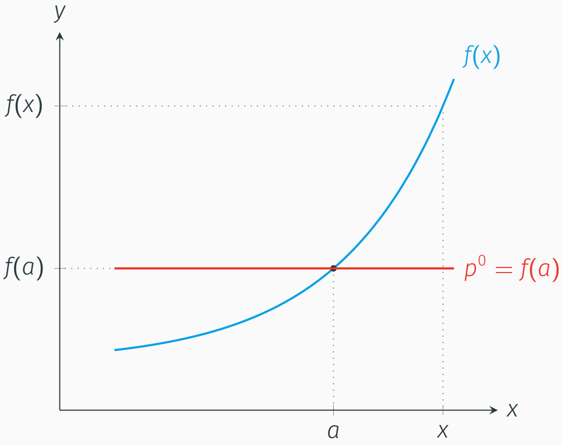
Approximating a function by a polynomial of order 0



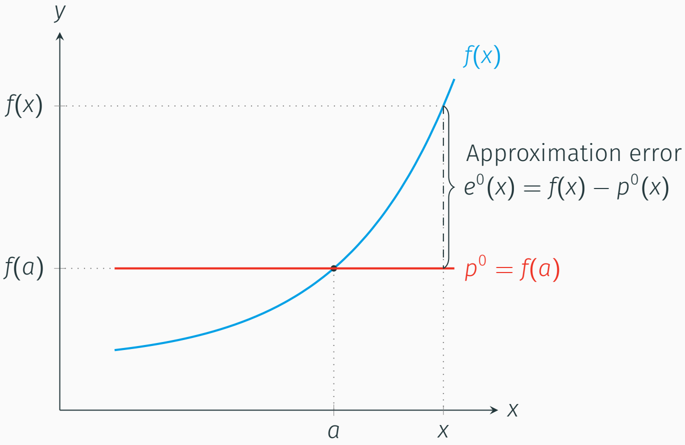


y
$f(x)$
$f(a)$
$p^0 = f(a)$
x
a

Approximating a function by a polynomial of order 0



Approximating a function by a polynomial of order 0



Approximating a function by a polynomial of order 1

A polynomial of degree 1 has equation

$$p(x) = c_0 + c_1 x,$$

but it can also be written as

$$p(x) = c_0 + c_1(x - a).$$

Among all the polynomials of degree 1, the one that best approximates f near a is that which meets the following conditions

1. p and f coincide at a: $p(a) = f(a)$,
2. p and f have the same rate of change at a: $p'(a) = f'(a)$.

The last condition guarantees that p and f have approximately the same tendency, but it requires the function f to be differentiable at a.

Imposing the previous conditions we have

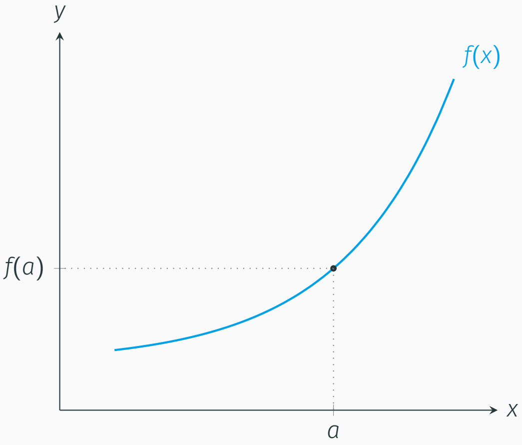
1. $p(x) = c_0 + c_1(x-a) \Rightarrow p(a) = c_0 + c_1(a-a) = c_0 = f(a)$,
2. $p'(x) = c_1 \Rightarrow p'(a) = c_1 = f'(a)$.

Therefore, the polynomial of degree 1 that best approximates f near a is

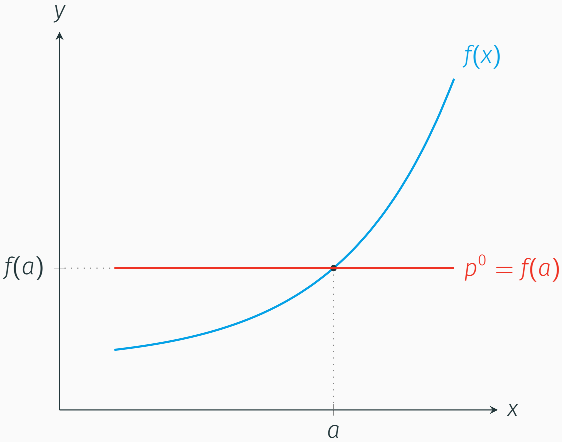
$$p(x) = f(a) + f'(a)(x-a),$$

which turns out to be the tangent line to f at $(a, f(a))$.

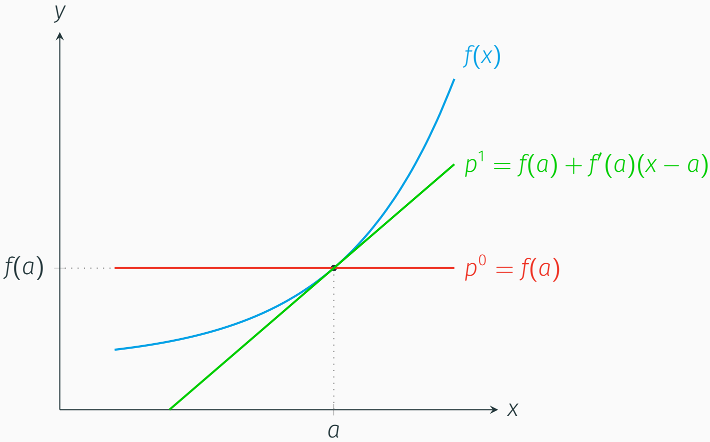
Approximating a function by a polynomial of order 1



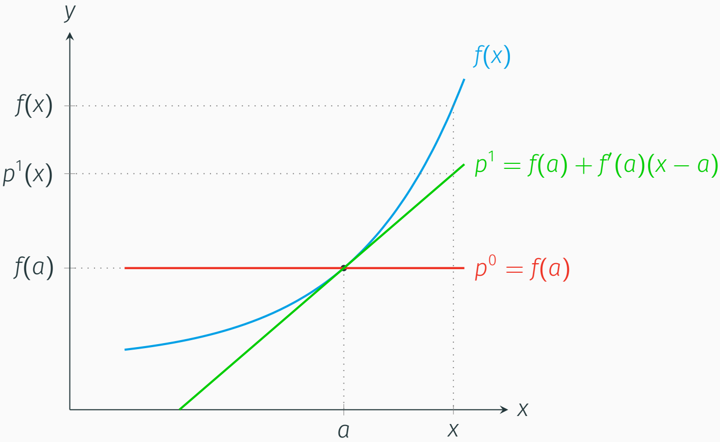
Approximating a function by a polynomial of order 1

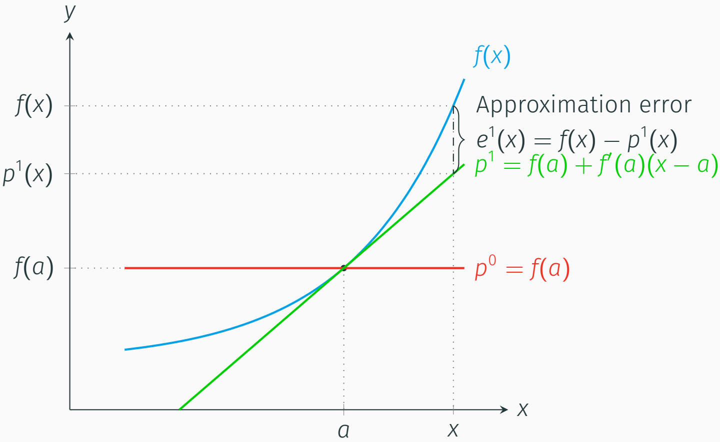


Approximating a function by a polynomial of order 1



Approximating a function by a polynomial of order 1





y
$f(x)$
Approximation error
$e^1(x) = f(x) - p^1(x)$
$p^1 = f(a) + f'(a)(x-a)$
$p^0 = f(a)$
$f(x)$
$p^1(x)$
$f(a)$
a
x
x

Approximating a function by a polynomial of order 2

A polynomial of degree 2 is a parabola with equation

$$p(x) = c_0 + c_1 x + c_2 x^2,$$

but it can also be written as

$$p(x) = c_0 + c_1(x-a) + c_2(x-a)^2.$$

Among all the polynomials of degree 2, the one that best approximate $f(x)$ near a is that which meets the following conditions

1. p and f coincide at a: $p(a) = f(a)$,
2. p and f have the same rate of change at a: $p'(a) = f'(a)$.
3. p and f have the same concavity at a: $p''(a) = f''(a)$.

The last condition requires the function f to be differentiable twice at a.

Best approximating polynomial of order 2

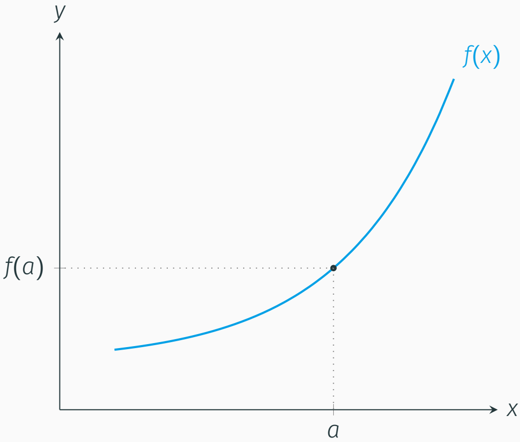
Imposing the previous conditions we have

1. $p(x) = c_0 + c_1(x - a) \Rightarrow p(a) = c_0 + c_1(a - a) = c_0 = f(a)$,
2. $p'(x) = c_1 \Rightarrow p'(a) = c_1 = f'(a)$.
3. $p''(x) = 2c_2 \Rightarrow p''(a) = 2c_2 = f''(a) \Rightarrow c_2 = \frac{f''(a)}{2}$.

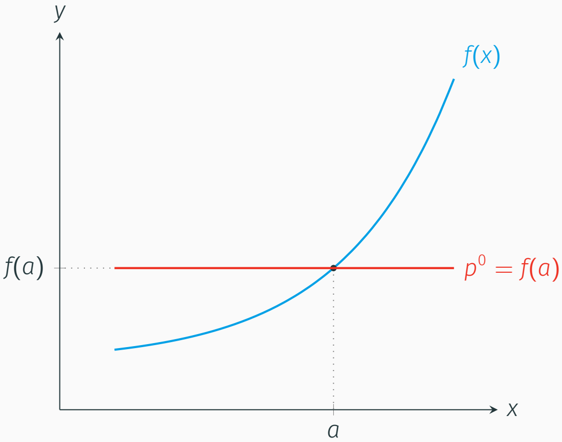
Therefore, the polynomial of degree 2 that best approximates f near a is

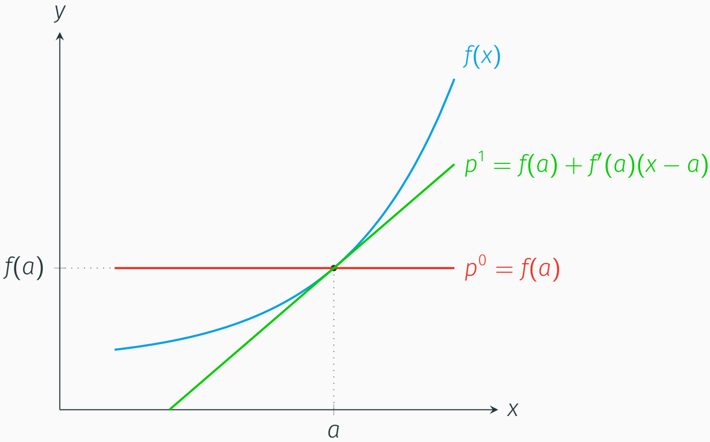
$$p(x) = f(a) + f'(a)(x - a) + \frac{f''(a)}{2}(x - a)^2.$$

Approximating a function by a polynomial of order 2

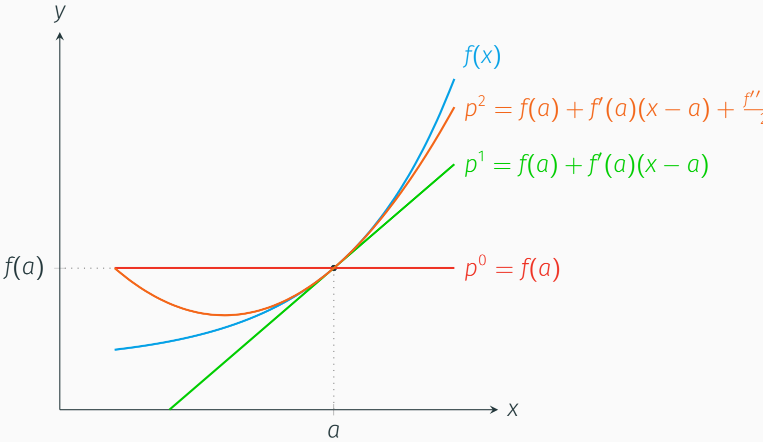


Approximating a function by a polynomial of order 2

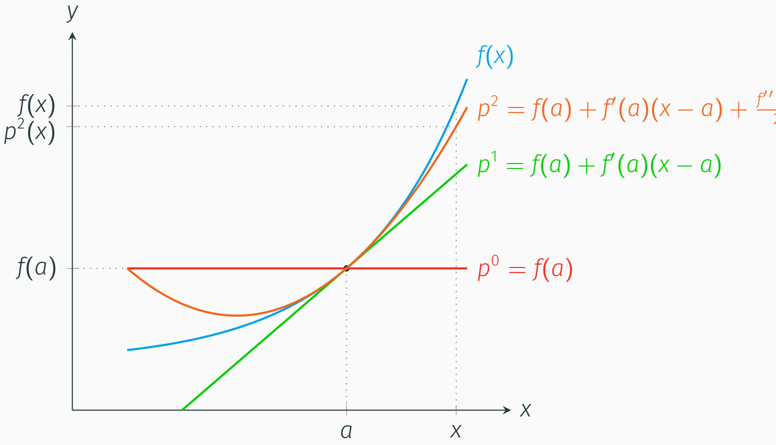



y
$f(x)$
$p^1 = f(a) + f'(a)(x - a)$
$f(a)$
$p^0 = f(a)$
x
a

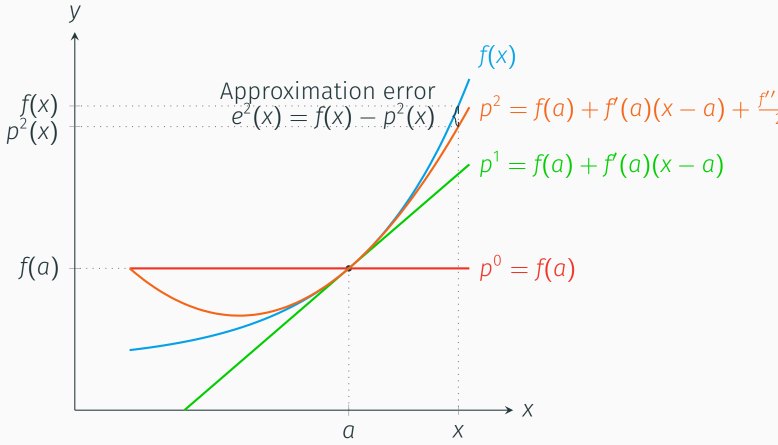
Approximating a function by a polynomial of order 2



Approximating a function by a polynomial of order 2



Approximating a function by a polynomial of order 2



Approximating a function by a polynomial of order n

A polynomial of degree n has equation

$$p(x) = c_0 + c_1 x + c_2 x^2 + \cdots + c_n x^n,$$

but it can also be written as

$$p(x) = c_0 + c_1(x-a) + c_2(x-a)^2 + \cdots + c_n(x-a)^n.$$

Among all the polynomials of degree n, the one that best approximate $f(x)$ near a is that which meets the following $n+1$ conditions

1. $p(a) = f(a)$,
2. $p'(a) = f'(a)$,
3. $p''(a) = f''(a)$,

   $\dots$

n+1. $p^{(n)}(a) = f^{(n)}(a)$.

Observe that these conditions require the function f to be differentiable n times at a.

Coefficients calculation for the best approximating polynomial of order n

The successive derivatives of p are

$$
\begin{aligned}
p(x) &= c_0 + c_1(x-a) + c_2(x-a)^2 + \cdots + c_n(x-a)^n,\\
p'(x) &= c_1 + 2c_2(x-a) + \cdots + nc_n(x-a)^{n-1},\\
p''(x) &= 2c_2 + \cdots + n(n-1)c_n(x-a)^{n-2},\\
&\vdots\\
p^{(n)}(x) &= n(n-1)(n-2)\cdots 1 c_n = n!c_n.
\end{aligned}
$$

Imposing the previous conditions we have

1. $p(a) = c_0 + c_1(a-a) + c_2(a-a)^2 + \cdots + c_n(a-a)^n = c_0 = f(a)$,
2. $p'(a) = c_1 + 2c_2(a-a) + \cdots + nc_n(a-a)^{n-1} = c_1 = f'(a)$,
3. $p''(a) = 2c_2 + \cdots + n(n-1)c_n(a-a)^{n-2} = 2c_2 = f''(a) \Rightarrow c_2 = f''(a)/2$,

   $\ldots$

n+1. $p^{(n)}(a) = n!c_n = f^{(n)}(a) = c_n = \frac{f^{(n)}(a)}{n!}$.

Definition (Taylor polynomial)

Given a function $f(x)$ differentiable n times at a, the *Taylor polynomial* of order n of f at a is the polynomial with equation

$$p_{f,a}^{n}(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2}(x-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n =$$
$$= \sum_{i=0}^{n} \frac{f^{(i)}(a)}{i!}(x-a)^i.$$

The Taylor polynomial of order n of f at a is the nth degree polynomial that best approximates f near a, as is the only one that meets the previous conditions.

Let us approximate the function $f(x) = \log x$ near the value 1 by a polynomial of order 3.

The equation of the Taylor polynomial of order 3 of f at $a = 1$ is

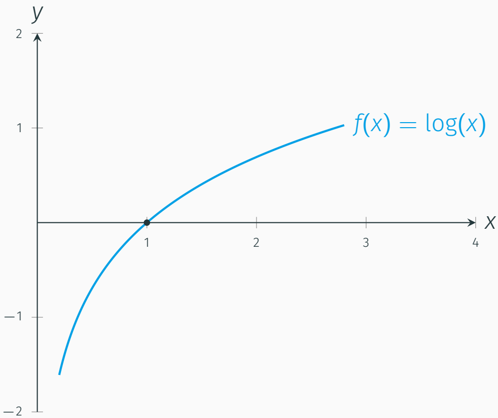
$$p_{f,1}^3(x) = f(1) + f'(1)(x-1) + \frac{f''(1)}{2}(x-1)^2 + \frac{f'''(1)}{3!}(x-1)^3.$$

The derivatives of f at 1 up to order 3 are

$$\begin{aligned}
f(x) &= \log x & f(1) &= \log 1 = 0,\\
f'(x) &= 1/x & f'(1) &= 1/1 = 1,\\
f''(x) &= -1/x^2 & f''(1) &= -1/1^2 = -1,\\
f'''(x) &= 2/x^3 & f'''(1) &= 2/1^3 = 2.
\end{aligned}$$

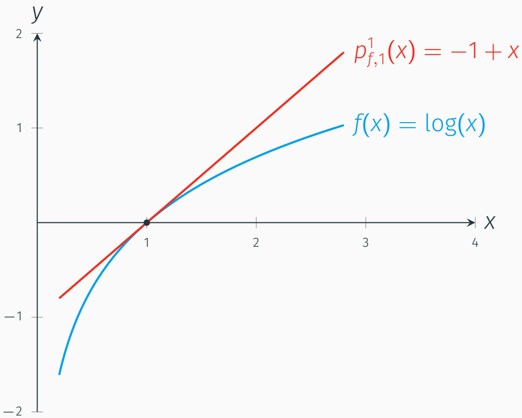
And substituting into the polynomial equation we get

$$p_{f,1}^3(x) = 0 + 1(x-1) + \frac{-1}{2}(x-1)^2 + \frac{2}{3!}(x-1)^3 = \frac{2}{3}x^3 - \frac{3}{2}x^2 + 3x - \frac{11}{6}.$$

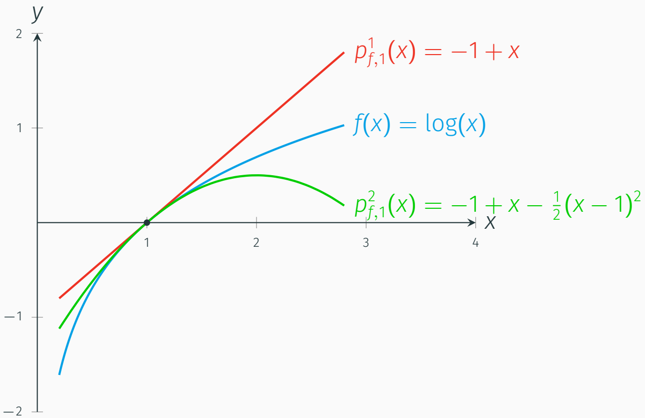


y
2
1
−1
−2
x
1
2
3
4
f(x) = log(x)

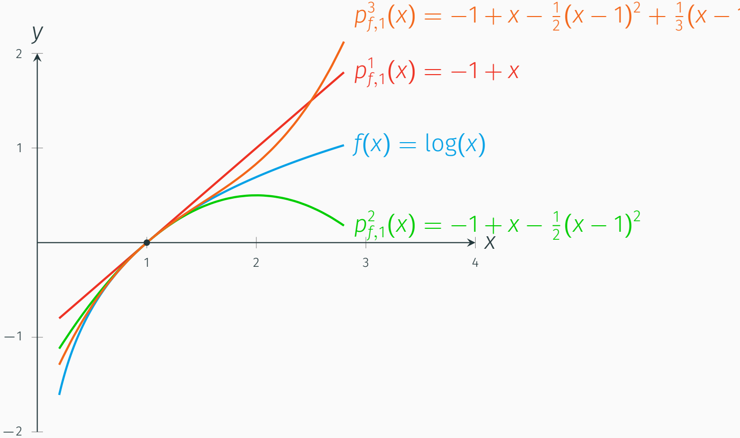
Taylor polynomials of the logarithmic function



Taylor polynomials of the logarithmic function



Taylor polynomials of the logarithmic function



The Taylor polynomial equation has a simpler form when the polynomial is calculated at 0. This special case of Taylor polynomial at 0 is known as the *Maclaurin polynomial*.

Definition (Maclaurin polynomial)

Given a function $f(x)$ differentiable n times at 0, the *Maclaurin polynomial of order n of f* is the polynomial with equation

$$p_{f,0}^{n}(x) = f(0) + f'(0)x + \frac{f''(0)}{2}x^2 + \cdots + \frac{f^{(n)}(0)}{n!}x^n =$$
$$= \sum_{i=0}^{n} \frac{f^{(i)}(0)}{i!}x^i.$$

Let us approximate the function $f(x) = \sin x$ near the value 0 by a polynomial of order 3.

The Maclaurin polynomial equation of order 3 of f is

$$p_{f,0}^3(x) = f(0) + f'(0)x + \frac{f''(0)}{2}x^2 + \frac{f'''(0)}{3!}x^3.$$

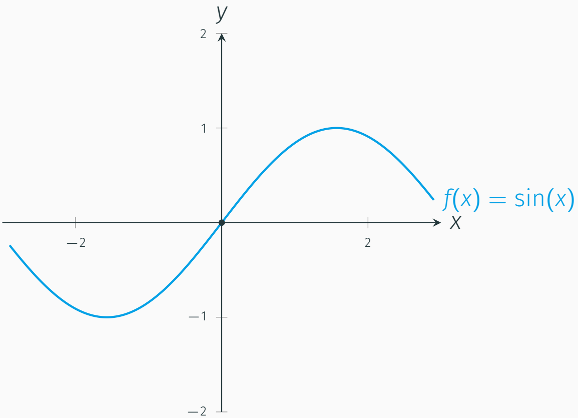
The derivatives of f at 0 up to order 3 are

$$\begin{aligned}
f(x) &= \sin x & f(0) &= \sin 0 = 0,\\
f'(x) &= \cos x & f'(0) &= \cos 0 = 1,\\
f''(x) &= -\sin x & f''(0) &= -\sin 0 = 0,\\
f'''(x) &= -\cos x & f'''(0) &= -\cos 0 = -1.
\end{aligned}$$

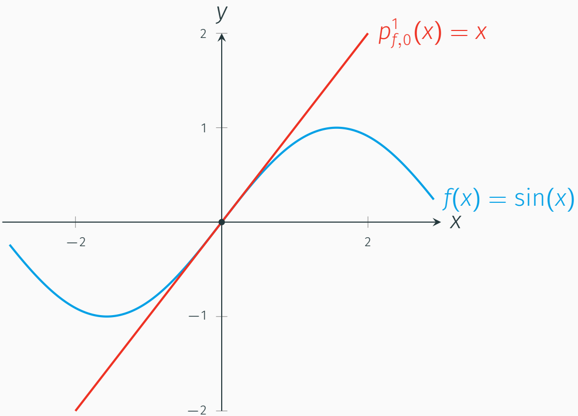
And substituting into the polynomial equation we get

$$p_{f,0}^3(x) = 0 + 1 \cdot x + \frac{0}{2}x^2 + \frac{-1}{3!}x^3 = x - \frac{x^3}{6}.$$

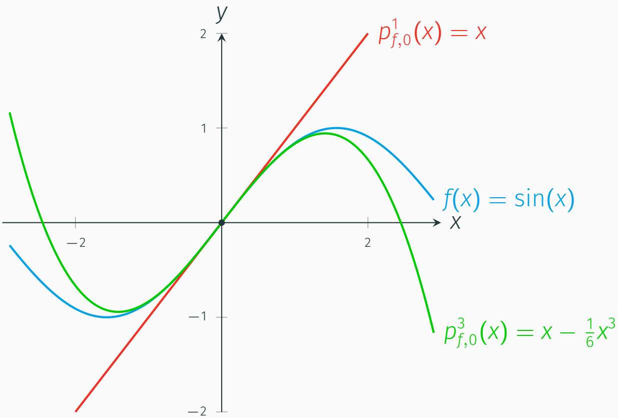
Maclaurin polynomial of the sine function



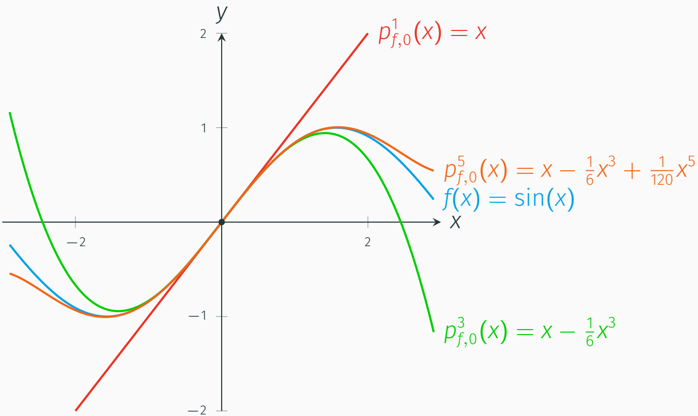
Maclaurin polynomial of the sine function



Maclaurin polynomial of the sine function



Maclaurin polynomial of the sine function



Maclaurin polynomials of elementary functions

$f(x)$	$p^n_{f,0}(x)$
$\sin x$	$x - \dfrac{x^3}{3!} + \dfrac{x^5}{5!} - \cdots + (-1)^k \dfrac{x^{2k-1}}{(2k-1)!}$ if $n = 2k$ or $n = 2k-1$
$\cos x$	$1 - \dfrac{x^2}{2!} + \dfrac{x^4}{4!} - \cdots + (-1)^k \dfrac{x^{2k}}{(2k)!}$ if $n = 2k$ or $n = 2k+1$
$\arctan x$	$x - \dfrac{x^3}{3} + \dfrac{x^5}{5} - \cdots + (-1)^k \dfrac{x^{2k-1}}{(2k-1)}$ if $n = 2k$ or $n = 2k-1$
e^x	$1 + x + \dfrac{x^2}{2!} + \dfrac{x^3}{3!} + \cdots + \dfrac{x^n}{n!}$
$\log(1+x)$	$x - \dfrac{x^2}{2} + \dfrac{x^3}{3} - \cdots + (-1)^{n-1} \dfrac{x^n}{n}$

Taylor polynomials allow to approximate a function in a neighborhood of a value a, but most of the times there is an error in the approximation.

Definition (Taylor remainder)

Given a function $f(x)$ and its Taylor polynomial of order n at a, $p^n_{f,a}(x)$, the *Taylor remainder* of order n of f at a is the difference between the function and the polynomial,

$$r^n_{f,a}(x) = f(x) - p^n_{f,a}(x).$$

The Taylor remainder measures the error in the approximation of $f(x)$ by the Taylor polynomial and allow us to express the function as the Taylor polynomial plus the Taylor remainder

$$f(x) = p^n_{f,a}(x) + r^n_{f,a}(x).$$

This expression is known as the *Taylor formula* of order n or f at a.

It can be proved that

$$\lim_{h\to 0} \frac{r^n_{f,a}(a+h)}{h^n} = 0,$$

Integrals

Definition (Antiderivative of a function)

Given a function $f(x)$, the function $F(X)$ is an *antiderivative* or *primitive function* of f if it satisfies that $F'(x) = f(x)\ \forall x \in \text{Dom}(f)$.

Example The function $F(x) = x^2$ is an antiderivative of the function $f(x) = 2x$ as $F'(x) = 2x$ on $\mathbb{R}$.

Roughly speaking, the calculus of antiderivatives is the reverse process of differentiation, and that is the reason for the name of antiderivative.

As two functions that differs in a constant term have the same derivative, if $F(x)$ is an antiderivative of $f(x)$, so will be any function of the form $F(x) + k$ $\forall k \in \mathbb{R}$. This means that, when a function has an antiderivative, it has an infinite number of antiderivatives.

Definition (Indefinite integral)

The *indefinite integral* of a function $f(x)$ is the set of all its antiderivatives; it is denoted by

$$\int f(x)\,dx = F(x) + C$$

where $F(x)$ is an antiderivative of $f(x)$ and C is a constant.

Example The indefinite integral of the function $f(x) = 2x$ is

$$\int 2x\,dx = x^2 + C.$$

We have seen in a previous chapter that the derivative of a function is the instantaneous rate of change of the function. Thus, if we know the instantaneous rate of change of the function at any point, we can compute the change of the function.

Example What is the space covered by an free falling object?

Assume that the only force acting upon an object drop is gravity, with an acceleration of 9.8 m/s^2. As acceleration is the the rate of change of the speed, that is constant at any moment, the antiderivative is the speed of the object,

$$v(t) = 9.8t \text{ m/s}$$

And as the speed is the rate of change of the space covered by object during the fall, the antiderivative of the speed is the space covered by the object,

$$s(t) = \int 9.8t\,dt = 9,8\frac{t^2}{2}.$$

Thus, for instance, after 2 seconds, the covered space is $s(2) = 9.8\frac{2^2}{2} = 19.6$ m.

Given two integrable functions $f(x)$ and $g(x)$ and a constant $k \in \mathbb{R}$, it is satisfied that

1. $\int (f(x) + g(x))\, dx = \int f(x)\, dx + \int g(x)\, dx$,
2. $\int kf(x)\, dx = k \int f(x)\, dx$.

This means that the integral of any linear combination of functions equals the same linear combination of the integrals of the functions.

Elementary integrals

- $\int a\,dx = ax + C$, with a constant.
- $\int x^n\,dx = \dfrac{x^{n+1}}{n+1} + C \quad$ if $n \neq -1$.
- $\int \dfrac{1}{x}\,dx = \ln|x| + C$.
- $\int e^x\,dx = e^x + C$.
- $\int a^x\,dx = \dfrac{a^x}{\ln a} + C$.
- $\int \sin x\,dx = -\cos x + C$.
- $\int \cos x\,dx = \sin x + C$.
- $\int \tan x\,dx = \ln|\sec x| + C$.
- $\int \sec x\,dx = \ln|\sec x + \tan x| + C$.
- $\int \csc x\,dx = \ln|\csc x - \cot x| + C$.
- $\int \cot x\,dx = \ln|\sin x| + C$.
- $\int \sec^2 x\,dx = \tan x + C$.
- $\int \csc^2 x\,dx = -\cot x + C$.
- $\int \sec x \tan x\,dx = \sec x + C$.
- $\int \csc x \cot x\,dx = -\csc x + C$.
- $\int \dfrac{dx}{\sqrt{a^2 - x^2}} = \arcsin \dfrac{x}{a} + C$.
- $\int \dfrac{dx}{a^2 + x^2} = \dfrac{1}{a} \arctan \dfrac{x}{a} + C$.
- $\int \dfrac{dx}{x\sqrt{x^2 - a^2}} = \dfrac{1}{a} \sec^{-1} \dfrac{x}{a} + C$.
- $\int \dfrac{dx}{a^2 - x^2} = \dfrac{1}{2a} \ln \left| \dfrac{x+a}{x-a} \right| + C$.

Unfortunately, unlike differential calculus, the is not a foolproof procedure to compute the antiderivative of a function. However, there are some techniques that allow to integrate some types of functions. The most common methods of integration are

- Integration by parts
- Integration by reduction
- Integration by substitution
- Integration of rational functions
- Integration of trigonometric functions

Given two differentiable functions $u(x)$ and $v(x)$, from the rule for differentiating a product we can get

$$\int u(x)v'(x)\,dx = u(x)v(x) - \int u'(x)v(x)\,dx,$$

or, writing $u'(x)dx = du$ and $v'(x)dx = dv$,

$$\int u\,dv = uv - \int v\,du.$$

To apply this method we have to choose the functions u and dv in a way so that the final integral is easier to compute than the original one.

Example To integrate $\int x \sin x\,dx$ we have to choose $u = x$ and $dv = \sin x\,dx$, so $du = dx$ and $v = -\cos x$, getting

$$\int x \sin x\,dx = -x\cos x - \int(-\cos x)\,dx = -x\cos x + \sin x.$$

If we had chosen $u = \sin x$ and $dv = x\,dx$, we would have got a more difficult integral.

The reduction technique is used when we have to apply the integration by parts several times.

If we want to compute the antiderivative I_n that depends on a natural number n, the reduction formulas allow us to write I_n as a function of I_{n-1}, that is, we have a recurrent relation

$$I_n = f(I_{n-1}, x, n)$$

so by computing the first antiderivative I_0 we should be able to compute the others.

Example To compute $I_n = \int x^n e^x \, dx$ applying integration by parts, we have to choose $u = x^n$ y $dv = e^x \, dx$, so $du = nx^{n-1} \, dx$ and $v = e^x$, getting

$$I_n = \int x^n e^x \, dx = x^n e^x - n \int x^{n-1} e^x \, dx = x^n e^x - nI_{n-1}.$$

Thus, for instance, for $n = 3$ we have

$$\int x^3 e^x \, dx = I_3 = x^3 e^x - 3I_2 = x^3 e^x - 3(x^2 e^x - 2I_1) = x^3 e^x - 3(x^2 e^x - (xe^x - I_0) =$$

$$= x^3 e^x - 3(x^2 e^x - (xe^x - e^x) = e^x(x^3 - 3x^2 + 6x - 6).$$

From the chain rule for differentiating the composition of two functions

$$f(g(x))' = f'(g(x))g'(x),$$

we can make a variable change $u = g(x)$, so $du = g'(x)dx$, and get

$$\int f'(g(x))g'(x)\,dx = \int f'(u)\,du = f(u) + C = f(g(x)) + C.$$

Example To compute the integral of $\int \frac{1}{x \log x}\,dx$ we can make the substitution $u = \log x$, so $du = \frac{1}{x}dx$, and we have

$$\int \frac{dx}{x \log x} = \int \frac{1}{\log x}\frac{1}{x}\,dx = \int \frac{1}{u}\,du = \log|u| + C.$$

Finally, undoing the substitution we get

$$\int \frac{1}{x \log x}\,dx = \log|\log x| + C.$$

A rational function can be written as the sum of a polynomial (with an immediate antiderivative) plus a proper rational function, that is, a rational function in which the degree of the numerator is less than the degree of the denominator.

On the other hand, depending of the factorization of the denominator, a proper rational function can be expressed as a sum of simpler fractions of the following types

- Denominator with a single linear factor: $\dfrac{A}{(x-a)}$
- Denominator with a linear factor repeated n times : $\dfrac{A}{(x-a)^n}$
- Denominator with a single quadratic factor: $\dfrac{Ax+B}{x^2+cx+d}$
- Denominator with a quadratic factor repeated n times: $\dfrac{Ax+B}{(x^2+cx+d)^n}$

Using the linearity of integration, we can compute the antiderivative of a rational function from the antiderivative of these partial fractions

$$\int \frac{A}{x-a}\,dx = A\log|x-a| + C,$$

$$\int \frac{A}{(x-a)^n}\,dx = \frac{-A}{(n-1)(x-a)^{n-1}} + C \text{ si } n \neq 1.$$

$$\int \frac{Ax+B}{x^2+cx+d} = \frac{A}{2}\log|x^2+cx+d| + \frac{2B-Ac}{\sqrt{4d-c^2}}\arctan\frac{2x+c}{\sqrt{4d-c^2}} + C.$$

Consider the function $f(x) = \dfrac{x^2+3x-5}{x^3-3x+2}$.

The factorization of the denominator is $x^3-3x+2 = (x-1)^2(x+2)$; it has a single linear factor $(x+2)$ and a linear factor $(x-1)$, repeated two times. In this case the decomposition in partial fractions is:

$$\begin{aligned}
\frac{x^2+3x-5}{x^3-3x+2} &= \frac{A}{x-1} + \frac{B}{(x-1)^2} + \frac{C}{x+2} = \\
&= \frac{A(x-1)(x+2) + B(x+2) + C(x-1)^2}{(x-1)^2(x+2)} = \\
&= \frac{(A+C)x^2 + (A+B-2C)x + (-2A+2B+C)}{(x-1)^2(x+2)}
\end{aligned}$$

and equating the numerators we get $A = 16/9$, $B = -1/3$ and $C = -7/9$, so

$$\frac{x^2+3x-5}{x^3-3x+2} = \frac{16/9}{x-1} + \frac{-1/3}{(x-1)^2} + \frac{-7/9}{x+2}.$$

Finally, integrating each partial fraction we have

$$\begin{aligned}\int \frac{x^2+3x-5}{x^3-3x+2}\,dx &= \int \frac{16/9}{x-1}\,dx + \int \frac{-1/3}{(x-1)^2}\,dx + \int \frac{-7/9}{x+2}\,dx = \\ &= \frac{16}{9}\int \frac{1}{x-1}\,dx - \frac{1}{3}\int (x-1)^{-2}\,dx - \frac{7}{9}\int \frac{1}{x+2}\,dx = \\ &= \frac{16}{9}\ln|x-1| + \frac{1}{3(x-1)} - \frac{7}{9}\ln|x+2| + C.\end{aligned}$$

Consider the function $f(x) = \dfrac{x+1}{x^2-4x+8}$.

In this case the denominator cannot be factorised as a product of linear factors, but we can write

$$x^2 - 4x + 8 = (x-2)^2 + 4,$$

so

$$\begin{aligned}\int \frac{x+1}{x^2-4x+8}\,dx = \int \frac{x-2+3}{(x-2)^2+4}\,dx &= \\ &= \int \frac{x-2}{(x-2)^2+4}\,dx + \int \frac{3}{(x-2)^2+4}\,dx = \\ &= \frac{1}{2}\ln\left|(x-2)^2+4\right| + \frac{3}{2}\arctan\left(\frac{x-2}{2}\right) + C.\end{aligned}$$

If $f(x) = \sin^n x \cos^m x$ with n or m odd, then we can make the substitution $t = \sin x$ or $t = \cos x$, to convert the function into a polynomial.

Example

$$\int \sin^2 x \cos^3 x \, dx = \int \sin^2 x \cos^2 x \cos x \, dx = \int \sin^2 x (1 - \sin^2 x) \cos x \, dx,$$

and making the substitution $t = \sin x$, so $dt = \cos x dx$, we have

$$\int \sin^2 x (1 - \sin^2 x) \cos x \, dx = \int t^2 (1 - t^2) \, dt = \int t^2 - t^4 \, dt = \frac{t^3}{3} - \frac{t^5}{5} + C.$$

Finally, undoing the substitution we have

$$\int \sin^2 x \cos^3 x \, dx = \frac{\sin^3 x}{3} - \frac{\sin^5 x}{5} + C.$$

If $f(x) = \sin^n x \cos^m x$ with n and m even, then we can make the following substitutions to simplify the integration

$$\sin^2 x = \frac{1}{2}(1 - \cos(2x))$$
$$\cos^2 x = \frac{1}{2}(1 + \cos(2x))$$
$$\sin x \cos x = \frac{1}{2}\sin(2x)$$

Example

$$\int \sin^2 x \cos^4 x\,dx = \int (\sin x \cos x)^2 \cos^2 x\,dx = \int \left(\frac{1}{2}\sin(2x)\right)^2 \frac{1}{2}(1 + \cos(2x))\,dx$$
$$= \frac{1}{8}\int \sin^2(2x)\,dx + \frac{1}{8}\int \sin^2(2x)\cos(2x)\,dx,$$

the first integral is of the same type and the second one of the previous type, so

$$\int \sin^2 x \cos^4 x\,dx = \frac{1}{32}x - \frac{1}{32}\sin(2x) + \frac{1}{24}\sin^3(2x).$$

The equalities

$$\sin x \cos y = \frac{1}{2}(\sin(x-y) + \sin(x+y))$$
$$\sin x \sin y = \frac{1}{2}(\cos(x-y) - \cos(x+y))$$
$$\cos x \cos y = \frac{1}{2}(\cos(x-y) + \cos(x+y))$$

transform products in sums, simplifying the integration

Example

$$\begin{aligned}\int \sin x \cos 2x\, dx &= \int \frac{1}{2}(\sin(x-2x) + \sin(x+2x))\, dx = \\ &= \frac{1}{2}\int \sin(-x)\, dx + \frac{1}{2}\int \sin 3x\, dx = \\ &= \frac{1}{2}\cos(-x) - \frac{1}{6}\cos 3x + C.\end{aligned}$$

If $f(x, y)$ is a rational function then the function $f(\sin x, \cos x)$ can be transformed in an rational function of t with the following substitutions

$$\tan\frac{x}{2} = t \quad \sin x = \frac{2t}{1+t^2} \quad \cos x = \frac{1-t^2}{1+t^2} \quad dx = \frac{2}{1+t^2}dt.$$

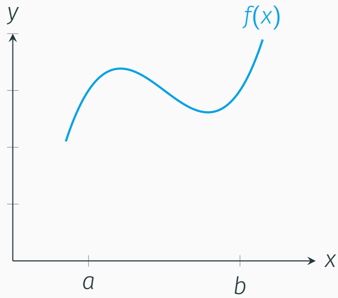
Example

$$\int \frac{1}{\sin x}\,dx = \int \frac{1}{\frac{2t}{1+t^2}}\frac{2}{1+t^2}\,dt = \int \frac{1}{t}\,dt = \log|t| + C = \log|\tan\frac{x}{2}| + C.$$

Definition (Definite integral)

Let $f(x)$ be a function which is continuous on an interval $[a, b]$. Divide this interval into n subintervals of equal width Δx and choose an arbitrary point x_i from each interval. The *definite integral* of f from a to b is defined to be the limit

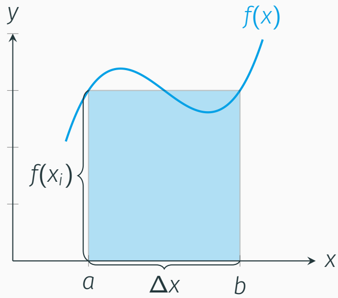
$$\int_a^b f(x)\,dx = \lim_{n\to\infty} \sum_{i=1}^{n} f(x_i)\Delta x.$$



Definition (Definite integral)

Let $f(x)$ be a function which is continuous on an interval $[a, b]$. Divide this interval into n subintervals of equal width Δx and choose an arbitrary point x_i from each interval. The *definite integral* of f from a to b is defined to be the limit

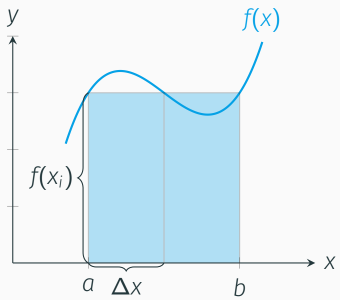
$$\int_a^b f(x)\,dx = \lim_{n\to\infty} \sum_{i=1}^{n} f(x_i)\Delta x.$$



Definition (Definite integral)

Let $f(x)$ be a function which is continuous on an interval $[a, b]$. Divide this interval into n subintervals of equal width Δx and choose an arbitrary point x_i from each interval. The *definite integral* of f from a to b is defined to be the limit

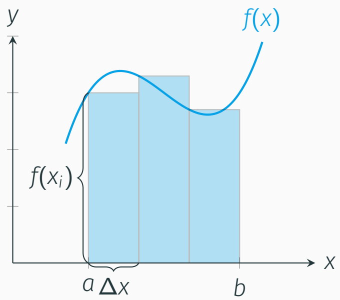
$$\int_a^b f(x)\,dx = \lim_{n\to\infty} \sum_{i=1}^{n} f(x_i)\Delta x.$$



Definition (Definite integral)

Let $f(x)$ be a function which is continuous on an interval $[a, b]$. Divide this interval into n subintervals of equal width Δx and choose an arbitrary point x_i from each interval. The *definite integral* of f from a to b is defined to be the limit

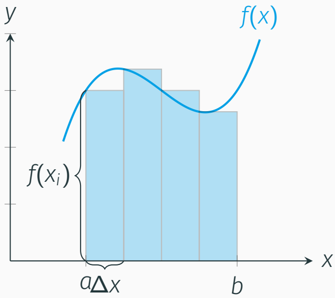
$$\int_a^b f(x)\,dx = \lim_{n\to\infty} \sum_{i=1}^{n} f(x_i)\Delta x.$$



Definition (Definite integral)

Let $f(x)$ be a function which is continuous on an interval $[a, b]$. Divide this interval into n subintervals of equal width Δx and choose an arbitrary point x_i from each interval. The *definite integral* of f from a to b is defined to be the limit

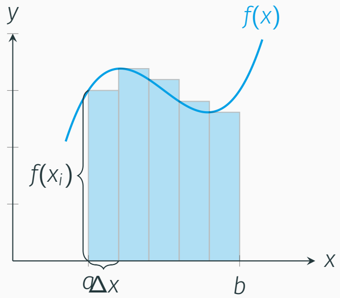
$$\int_a^b f(x)\,dx = \lim_{n\to\infty} \sum_{i=1}^{n} f(x_i)\Delta x.$$



Definition (Definite integral)

Let $f(x)$ be a function which is continuous on an interval $[a, b]$. Divide this interval into n subintervals of equal width Δx and choose an arbitrary point x_i from each interval. The *definite integral* of f from a to b is defined to be the limit

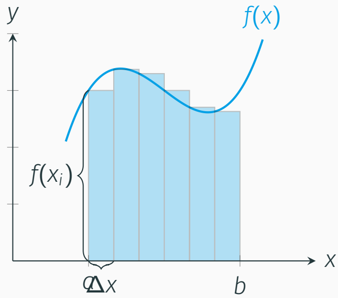
$$\int_a^b f(x)\,dx = \lim_{n\to\infty} \sum_{i=1}^{n} f(x_i)\Delta x.$$



Definition (Definite integral)

Let $f(x)$ be a function which is continuous on an interval $[a, b]$. Divide this interval into n subintervals of equal width Δx and choose an arbitrary point x_i from each interval. The *definite integral* of f from a to b is defined to be the limit

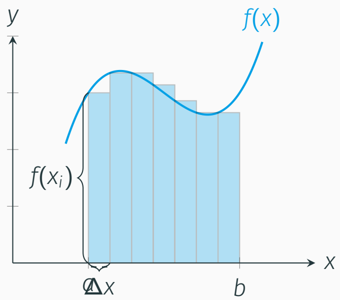
$$\int_a^b f(x)\,dx = \lim_{n\to\infty} \sum_{i=1}^{n} f(x_i)\Delta x.$$



Definition (Definite integral)

Let $f(x)$ be a function which is continuous on an interval $[a, b]$. Divide this interval into n subintervals of equal width Δx and choose an arbitrary point x_i from each interval. The *definite integral* of f from a to b is defined to be the limit

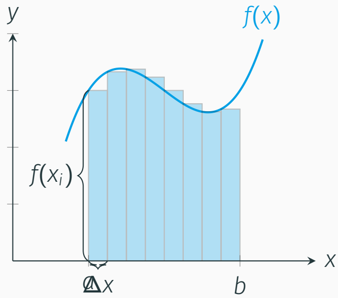
$$\int_a^b f(x)\,dx = \lim_{n\to\infty} \sum_{i=1}^{n} f(x_i)\Delta x.$$



Definition (Definite integral)

Let $f(x)$ be a function which is continuous on an interval $[a, b]$. Divide this interval into n subintervals of equal width Δx and choose an arbitrary point x_i from each interval. The *definite integral* of f from a to b is defined to be the limit

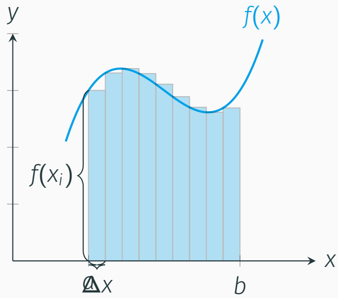
$$\int_a^b f(x)\,dx = \lim_{n\to\infty} \sum_{i=1}^{n} f(x_i)\Delta x.$$



Definition (Definite integral)

Let $f(x)$ be a function which is continuous on an interval $[a, b]$. Divide this interval into n subintervals of equal width Δx and choose an arbitrary point x_i from each interval. The *definite integral* of f from a to b is defined to be the limit

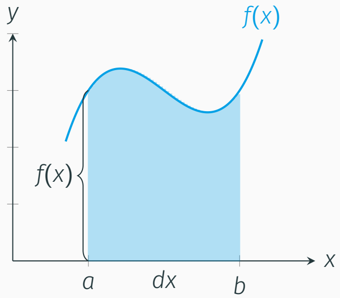
$$\int_a^b f(x)\,dx = \lim_{n\to\infty} \sum_{i=1}^{n} f(x_i)\Delta x.$$



Definition (Definite integral)

Let $f(x)$ be a function which is continuous on an interval $[a, b]$. Divide this interval into n subintervals of equal width Δx and choose an arbitrary point x_i from each interval. The *definite integral* of f from a to b is defined to be the limit

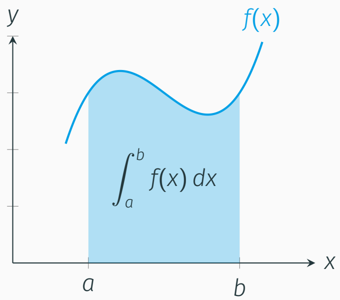
$$\int_a^b f(x)\,dx = \lim_{n\to\infty} \sum_{i=1}^{n} f(x_i)\Delta x.$$



Definition (Definite integral)

Let $f(x)$ be a function which is continuous on an interval $[a, b]$. Divide this interval into n subintervals of equal width Δx and choose an arbitrary point x_i from each interval. The *definite integral* of f from a to b is defined to be the limit

$$\int_a^b f(x)\,dx = \lim_{n\to\infty} \sum_{i=1}^{n} f(x_i)\Delta x.$$



Theorem (First fundamental theorem of Calculus)

If $f(x)$ is continuous on the interval $[a, b]$ and $F(x)$ is an antiderivative of f on $[a, b]$, then

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

Example. Given the function $f(x) = x^2$, we have

$$\int_1^2 x^2\,dx = \left[\frac{x^3}{3}\right]_1^2 = \frac{2^3}{3} - \frac{1^3}{3} = \frac{7}{3}.$$

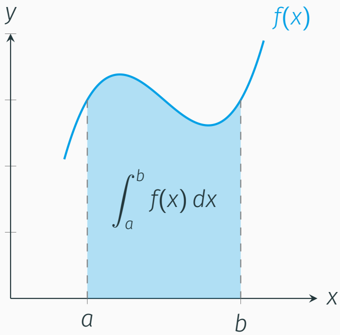
Given two functions $f(x)$ and $g(x)$ integrable on $[a, b]$ and $k \in \mathbb{R}$ the following properties are satisfied:

- $\int_a^b (f(x) + g(x))\, dx = \int_a^b f(x)\, dx + \int_a^b g(x)\, dx$ (linearity)
- $\int_a^b kf(x)\, dx = k \int_a^b f(x)\, dx$ (linearity)
- $\int_a^b f(x)\, dx \leq \int_a^b g(x)\, dx$ si $f(x) \leq g(x)$ $\forall x \in [a, b]$ (monotony)
- $\int_a^b f(x)\, dx = \int_a^c f(x)\, dx + \int_c^b f(x)\, dx$ for any $c \in (a, b)$ (additivity)
- $\int_a^b f(x)\, dx = -\int_b^a f(x)\, dx$

If $f(x)$ is an integrable function on the interval $[a, b]$ and $f(x) \geq 0 \ \forall x \in [a, b]$, then the definite integral

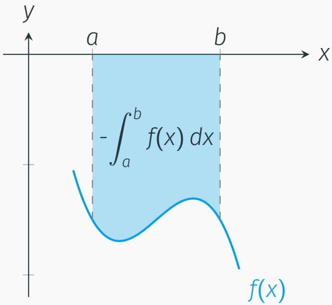
$$\int_a^b f(x)\, dx$$

measures the area between the graph of f and the x axis on the interval $[a, b]$.



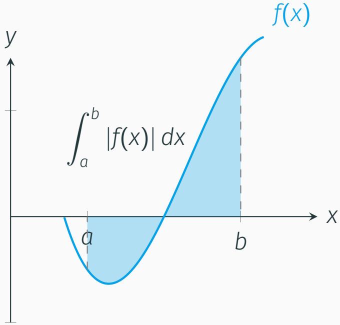
If $f(x)$ is an integrable function on the interval $[a, b]$ and $f(x) \leq 0 \ \forall x \in [a, b]$, then the area between the graph of f and the x axis on the interval $[a, b]$ is

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



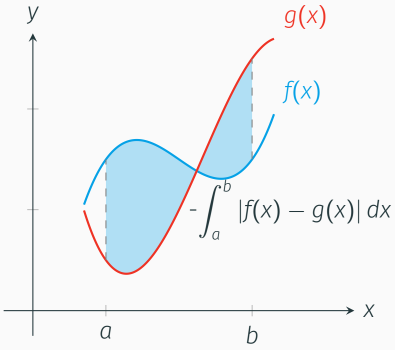
In general, if $f(x)$ is an integrable function on the interval $[a, b]$, no matter the sign of f on $[a, b]$, the area between the graph of f and the x axis on the interval $[a, b]$ is

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



If $f(x)$ and $g(x)$ are two integrable functions on the interval $[a, b]$, then the area between the graph of f and g on the interval $[a, b]$ is

$$\int_a^b |f(x) - g(x)|\, dx.$$



Ordinary Differential Equations

Often in Physics, Chemistry, Biology, Geometry, etc there arise equations that relate a function with its derivative, or successive derivatives.

Definition (Ordinary differential equation)

An *ordinary differential equation* (O.D.E.) is a equation that relates an independent variable x, a function $y(x)$ that depends on x, and the successive derivatives of y, $y', y'', \ldots, y^{(n)}$; it can be written as

$$F(x, y, y', y'', \ldots, y^{(n)}) = 0.$$

The *order* of a differential equation is the greatest order of the derivatives in the equation.

Thus, for instance, the equation $y''' + sen(x)y' = 2x$ is a differential equation of order 3.

Deducing a differential equation

To deduce a differential equation that explains a natural phenomenon is essential to understand what a derivative is and how to interpret it.

Example Newton's law of cooling states

> *"The rate of change of the temperature of a body in a surrounding medium is proportional to the difference between the temperature of the body T and the temperature of the medium T_a."*

The rate of change of the temperature is the derivative of temperature with respect to time dT/dt. Thus, Newton's law of cooling can be explained by the differential equation

$$\frac{dT}{dt} = k(T - T_a),$$

where k is a proportionality constant.

Solution of an ordinary differential equation

Definition (Solution of an ordinary differential equation)

Given an ordinary differential equation $F(x, y, y', y'', \ldots, y^{(n)}) = 0$, the function $y = f(x)$ is a *solution of the ordinary differential equation* if it satisfies the equation, that is, if

$$F(x, f(x), f'(x), f''(x), \ldots, f^{(n}(x)) = 0.$$

The graph of a solution of the ordinary differential equation is known as *integral curve*.

Solving an ordinary differential equations consists on finding all its solutions in a given domain. For integral calculus is required.

The same manner than the indefinite integral is a family of antiderivatives, that differ in a constant term, after integrating an ordinary differential equation we get a family of solutions that differ in a constant. We can get particular solutions by giving values to this constant.

Definition (General solution of an ordinary differential equation)

Given an ordinary differential equation $F(x, y, y', y'', \ldots, y^{(n)}) = 0$ of order n, the *general solution* of the differential equation is a family of functions

$$y = f(x, C_1, \ldots, C_n),$$

depending on n constants, such that for any value of $C_1, \ldots, C_n$ we get a solution of the differential equation.

For every value of the constant we get *particular solution* of the differential equation. Thus, when a differential equation can be solved, it has infinite solutions.

Geometrically, the general solution of a differential equation corresponds to a family of integral curves of the differential equation.

Often, it is common to impose conditions to the solutions of a differential equation to reduce the number of solutions. In many cases, these conditions allow to determine the values of the constants in the general solution to get a particular solution.

In this chapter we are going to study first order differential equations

$$F(x, y, y') = 0.$$

The general solution of a first order differential equation is

$$y = f(x, C),$$

so to get a particular solution from the general one, it is enough to set the value of the constant C, and for that we only need to impose one initial condition.

Definition (Initial value problem)

The group consisting of a first order differential equation and an initial condition is known as *initial value problem*:

$$\begin{cases} F(x, y, y') = 0, & \text{First order differential equation;} \\ y(x_0) = y_0, & \text{Initial condition.} \end{cases}$$

Solving an initial value problem consists in finding a solution of the first

Recall the first order differential equation of the Newton's law of cooling,

$$\frac{dT}{dt} = k(T - T_a),$$

where T is the temperature of the body and T_a is the temperature of the surrounding medium.

It is easy to check that the general solution of this equation is

$$T(t) = Ce^{kt} + T_a.$$

If we impose the initial condition that the temperature of the body at the initial instant is 5 ºC, that is, $T(0) = 5$, we have

$$T(0) = Ce^{k\cdot 0} + T_a = C + T_a = 5,$$

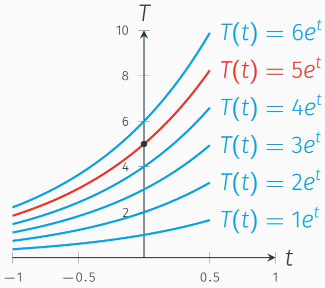
from where we get $C = 5 - T_a$, and this give us the particular solution

$$T(t) = (5 - T_a)e^{kt} + T_a.$$

If we assume in the previous example that the temperature of the surrounding medium is $T_a = 0$ ºC and the cooling constant of the body is $k = 1$, the general solution of the differential equation is

$$T(t) = Ce^t,$$

that is a family of integral curves. Among all of them, only the one that passes through the point $(0, 5)$ corresponds to the particular solution of the previous initial value problem.



Theorem (Existence and uniqueness of solutions of a first order ODE)

Given an initial value problem

$$\begin{cases} y' = F(x, y); \\ y(x_0) = y_0; \end{cases}$$

if $F(x, y(x))$ is a function continuous on an open interval around the point (x_0, y_0), then a solution of the initial value problem exists. If, in addition, $\frac{\partial F}{\partial y}$ is continuous in an open interval around (x_0, y_0), the solution is unique.

Although this theorem guarantees the existence and uniqueness of a solution of a first order differential equation, it does not provide a method to compute it. In fact, there is not a general method to solve first order differential equations, but we will see how to solve some types:

- Separable differential equations
- Homogeneous differential equations
- Linear differential equations

Definition (Separable differential equation)

A *separable differential equation* is a first order differential equation that can be written as

$$y'g(y) = f(x),$$

or what is the same,

$$g(y)dy = f(x)dx,$$

so the different variables are on different sides of the equality (the variables are separated).

The general solution for a separable differential equation comes after integrating both sides of the equation

$$\int g(y)\,dy = \int f(x)\,dx + C.$$

The differential equation of the Newton's law of cooling

$$\frac{dT}{dt} = k(T - T_a),$$

is a separable differential equation since it can be written as

$$\frac{1}{T - T_a} dT = k\, dt.$$

Integrating both sides of the equation we have

$$\int \frac{1}{T - T_a}\, dT = \int k\, dt \Leftrightarrow \log(T - T_a) = kt + C,$$

and solving for T we get the general solution of the equation

$$T(t) = e^{kt+C} + T_a = e^C e^{kt} + T_a = Ce^{kt} + T_a,$$

rewriting $C = e^C$ as an arbitrary constant.

Definition (Homogeneous function)

A function $f(x, y)$ is *homogeneous* of degree n, if it satisfies

$$f(kx, ky) = k^n f(x, y),$$

for any value $k \in \mathbb{R}$.

In particular, a homogeneous function of degree 0 always satisfies

$$f(kx, ky) = f(x, y).$$

Setting $k = 1/x$ we have

$$f(x, y) = f\left(\frac{1}{x}x, \frac{1}{x}y\right) = f\left(1, \frac{y}{x}\right) = g\left(\frac{y}{x}\right).$$

This way, a homogeneous function of degree 0 always can be written as a function of $u = y/x$:

$$f(x, y) = g\left(\frac{y}{x}\right) = g(u).$$

Definition (Homogeneous differential equation)

A *homogeneous differential equation* is a first order differential equation that can be written as

$$y' = f(x, y),$$

where $f(x, y)$ is a homogeneous function of degree 0.

We can solve a homogeneous differential equation by making the substitution

$$u = \frac{y}{x} \Leftrightarrow y = ux,$$

so the equation becomes

$$u'x + u = f(u),$$

that is a separable differential equation.

Once solved the separable differential equation, the substitution must be undone.

Let us consider the following differential equation

$$4x - 3y + y'(2y - 3x) = 0.$$

Rewriting the equation in this way

$$y' = \frac{3y - 4x}{2y - 3x}$$

we can easily check that it is a homogeneous differential equation.

To solve this equation we have to do the substitution $y = ux$, and we get

$$u'x + u = \frac{3ux - 4x}{2ux - 3x} = \frac{3u - 4}{2u - 3}$$

that is a separable differential equation.

Separating the variables we have

$$u'x = \frac{3u - 4}{2u - 3} - u = \frac{-2u^2 + 6u - 4}{2u - 3} \Leftrightarrow \frac{2u - 3}{-2u^2 + 6u - 4}\, du = \frac{1}{x}\, dx.$$

Now, integrating both sides of the equation we have

$$\int \frac{2u-3}{-2u^2+6u-4}\,du = \int \frac{1}{x}\,dx \Leftrightarrow -\frac{1}{2}\log|u^2-3u+2| = \log|x| + C \Leftrightarrow$$

$$\Leftrightarrow \log|u^2-3u+2| = -2\log|x| - 2C,$$

then, applying the exponential function to both sides and simplifying we get the general solution

$$u^2-3u+2 = e^{-2\log|x|-2C} = \frac{e^{-2C}}{e^{\log|x|^2}} = \frac{C}{x^2},$$

rewriting the constant $K = e^{-2C}$.

Finally, undoing the initial substitution $u = y/x$, we arrive at the general solution of the homogeneous differential equation

$$\left(\frac{y}{x}\right)^2 - 3\frac{y}{x} + 2 = \frac{K}{x^2} \Leftrightarrow y^2 - 3xy + 2x^2 = K.$$

Definition (Linear differential equation)

A *linear differential equation* is a first order differential equation that can be written as

$$y' + g(x)y = h(x).$$

To solve a linear differential equation we try to write the left side of the equation as the derivative of a product. For that we multiply both sides by a function $f(x)$, such that

$$f'(x) = g(x)f(x).$$

Thus, we get

$$y'f(x) + g(x)f(x)y = h(x)f(x)$$
$$\Updownarrow$$
$$y'f(x) + f'(x)y = h(x)f(x)$$
$$\Updownarrow$$
$$\frac{d}{dx}(yf(x)) = h(x)f(x)$$

Integrating both sides of the previous equation we get the solution

$$yf(x) = \int h(x)f(x)\,dx + C.$$

On the other hand, the unique function that satisfies $f'(x) = g(x)f(x)$ is

$$f(x) = e^{\int g(x)\,dx},$$

so, substituting this function in the previous solution we arrive at the solution of the linear differential equation

$$ye^{\int g(x)\,dx} = \int h(x)e^{\int g(x)\,dx}\,dx + C,$$

or what is the same

$$y = e^{-\int g(x)\,dx}\left(\int h(x)e^{\int g(x)\,dx}\,dx + C\right).$$

If in the differential equation of the Newton's law of cooling the temperature of the surrounding medium is a function of time $T_a(t)$, then the differential equation

$$\frac{dT}{dt} = k(T - T_a(t)),$$

is a linear differential equation since it can be written as

$$T' - kT = -kT_a(t),$$

where the independent term is $-kT_a(t)$ and the coefficient of T is $-k$.

Substituting in the formula of the general solution of a linear differential equation we have

$$y = e^{-\int -k\,dt}\left(\int -kT_a(t)e^{\int -k\,dt}\,dt + C\right) = e^{kt}\left(-\int kT_a(t)e^{-kt}\,dt + C\right).$$

In the particular case that $T_a(t) = t$, and the proportionality constant $k = 1$, the general solution of the linear differential equation is

$$y = e^t \left(-\int te^{-kt}\, dt + C \right) = e^t(e^{-t}(t+1) + C) = Ce^t + t + 1.$$

If, in addition, we know that the temperature of the body at time $t = 0$ is 5 ºC, that is, we have the initial condition $T(0) = 5$, then we can compute the value of the constant C,

$$y(0) = Ce^0 + 0 + 1 = 5 \Leftrightarrow C + 1 = 5 \Leftrightarrow C = 4,$$

and we get the particular solution

$$y(t) = 4e^t + t + 1.$$

Several Variables Differentiable Calculus

Definition (Vector function of a single real variable)

A *vector function of a single real variable* or *vector field of a scalar variable* is a function that maps every scalar value $t \in D \subseteq \mathbb{R}$ into a vector $(x_1(t), \ldots, x_n(t))$ in $\mathbb{R}^n$:

$$\begin{array}{rccc} f: & \mathbb{R} & \longrightarrow & \mathbb{R}^n \\ & t & \longrightarrow & (x_1(t), \ldots, x_n(t)) \end{array}$$

where $x_i(t)$, $i = 1, \ldots, n$, are real function of a single real variable known as *coordinate functions.*

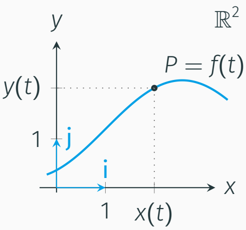
The most common vector field of scalar variable are in the the real plane $\mathbb{R}^2$, where usually they are represented as

$$f(t) = x(t)\mathbf{i} + y(t)\mathbf{j},$$

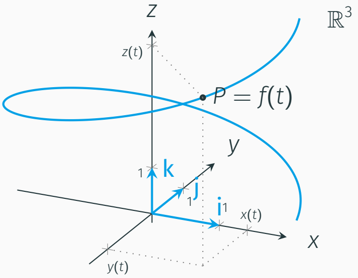
and in the real space $\mathbb{R}^3$, where usually they are represented as

$$f(t) = x(t)\mathbf{i} + y(t)\mathbf{j} + z(t)\mathbf{k},$$

The graphic representation of a vector field in $\mathbb{R}^2$ is a trajectory in the real plane.



The graphic representation of a vector field in $\mathbb{R}^3$ is a trajectory in the real space.



The concept of derivative as the limit of the average rate of change of a function can be extended easily to vector fields.

Definition (Derivative of a vectorial field)

A vectorial field $f(t) = (x_1(t), \ldots, x_n(t))$ is *differentiable* at a point $t = a$ if the limit

$$\lim_{\Delta t \to 0} \frac{f(a + \Delta t) - f(a)}{\Delta t}.$$

exists. In such a case, the value of the limit is known as the *derivative* of the vector field at a, and it is written $f'(a)$.

Derivative of a vector field

Many properties of real functions of a single real variable can be extended to vector fields through its component functions. Thus, for instance, the derivative of a vector field can be computed from the derivatives of its component functions.

Theorem

Given a vector field $f(t) = (x_1(t), \dots, x_n(t))$, if $x_i(t)$ is differentiable at $t = a$ for all $i = 1, \dots, n$, then f is differentiable at a and its derivative is

$$f'(a) = (x_1'(a), \dots, x_n'(a))$$

The proof for a vectorial field in $\mathbb{R}^2$ is easy.

$$f'(a) = \lim_{\Delta t \to 0} \frac{f(a+\Delta t) - f(a)}{\Delta t} = \lim_{\Delta t \to 0} \frac{(x(a+\Delta t), y(a+\Delta t)) - (x(a), y(a))}{\Delta t} =$$

$$= \lim_{\Delta t \to 0} \left(\frac{x(a+\Delta t) - x(a)}{\Delta t}, \frac{y(a+\Delta t) - y(a)}{\Delta t} \right) =$$

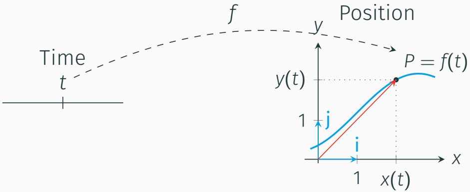
$$= \left(\lim_{\Delta t \to 0} \frac{x(a+\Delta t) - x(a)}{\Delta t}, \lim_{\Delta t \to 0} \frac{y(a+\Delta t) - y(a)}{\Delta t} \right) = (x'(a), y'(a)).$$

The notion of derivative as a velocity along a trajectory in the real line can be generalized to a trajectory in any euclidean space $\mathbb{R}^n$.

In case of a two dimensional space $\mathbb{R}^2$, if $f(t)$ describes the position of a moving object in the real plane at any time t, taking as reference the coordinates origin O and the unitary vectors $\{\mathbf{i} = (1,0), \mathbf{j} = (0,1)\}$, we can represent the position of the moving object P at every moment t with a vector $\vec{OP} = x(t)\mathbf{i} + y(t)\mathbf{j}$, where the coordinates

$$\begin{cases} x = x(t) \\ y = y(t) \end{cases} \qquad t \in \mathrm{Dom}(f)$$

are the *coordinate functions* of f.



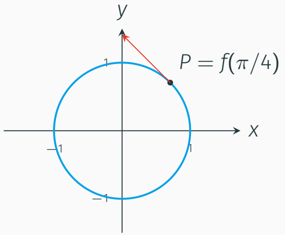
Velocity of a curvilinear motion in the plane

In this context the derivative of a trajectory $f'(a) = (x_1'(a), \ldots, x_n'(a))$ is the *velocity* vector of the trajectory f at moment $t = a$.

Example Given the trajectory $f(t) = (\cos t, \sin t)$, $t \in \mathbb{R}$, whose image is the unit circumference centered in the coordinate origin, its coordinate functions are $x(t) = \cos t$, $y(t) = \sin t$, $t \in \mathbb{R}$, and its velocity is

$$\mathbf{v} = f'(t) = (x'(t), y'(t)) = (-\sin t, \cos t).$$

In the moment $t = \pi/4$, the object is in position $f(\pi/4) = (\cos(\pi/4), \sin(\pi/4)) = (\sqrt{2}/2, \sqrt{2}/2)$ and it is moving with a velocity $\mathbf{v} = f'(\pi/4) = (-\sin(\pi/4), \cos(\pi/4)) = (-\sqrt{2}/2, \sqrt{2}/2)$.



Given a trajectory $f(t)$ in the real plane, the vectors that are parallel to the velocity $\mathbf{v}$ at a moment a are called *tangent vectors* to the trajectory f at the moment a, and the line passing through $P = f(a)$ directed by $\mathbf{v}$ is the tangent line to the graph of f at the moment a.

Definition (Tangent line to a trajectory)

Given a trajectory $f(t)$ in the real plane $\mathbb{R}^2$, the *tangent line* to to the graph of f at a is the line with equation

$$\begin{aligned} l : (x, y) &= f(a) + tf'(a) = (x(a), y(a)) + t(x'(a), y'(a)) \\ &= (x(a) + tx'(a), y(a) + ty'(a)). \end{aligned}$$

We have seen that for the trajectory $f(t) = (\cos t, \sin t)$, $t \in \mathbb{R}$, whose image is the unit circumference centered at the coordinate origin, the object position at the moment $t = \pi/4$ is $f(\pi/4) = (\sqrt{2}/2, \sqrt{2}/2)$ and its velocity $\mathbf{v} = (-\sqrt{2}/2, \sqrt{2}/2)$. Thus the equation of the tangent line to f at that moment is

$$
\begin{aligned}
l : (x, y) = f(\pi/4) + t\mathbf{v} &= \left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right) + t\left(\frac{-\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right) = \\
&= \left(\frac{\sqrt{2}}{2} - t\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2} + t\frac{\sqrt{2}}{2}\right).
\end{aligned}
$$

From the vectorial equation of the tangent to a trajectory $f(t)$ at the moment $t = a$ we can get the coordinate functions

$$\begin{cases} x = x(a) + tx'(a) \\ y = y(a) + ty'(a) \end{cases} \qquad t \in \mathbb{R},$$

and solving for t and equalling both equations we get the *Cartesian equation* of the tangent

$$\frac{x - x(a)}{x'(a)} = \frac{y - y(a)}{y'(a)},$$

if $x'(a) \neq 0$ and $y'(a) \neq 0$.

From this equation it is easy to get the *point-slope equation* of the tangent

$$y - y(a) = \frac{y'(a)}{x'(a)}(x - x(a)).$$

Using the vectorial equation of the tangent of the previous example

$$l:(x,y)=\left(\frac{\sqrt{2}}{2}-t\frac{\sqrt{2}}{2},\frac{\sqrt{2}}{2}+t\frac{\sqrt{2}}{2}\right),$$

its Cartesian equation is

$$\frac{x-\sqrt{2}/2}{-\sqrt{2}/2}=\frac{y-\sqrt{2}/2}{\sqrt{2}/2}$$

and the point-slope equation is

$$y-\sqrt{2}/2=\frac{-\sqrt{2}/2}{\sqrt{2}/2}(x-\sqrt{2}/2)\Rightarrow y=-x+\sqrt{2}.$$

We have seen that the tangent line to a trajectory $f(t)$ at a is the line passing through the point $P = f(a)$ directed by the velocity vector $\mathbf{v} = f'(a) = (x'(a), y'(a))$. If we take as direction vector a vector orthogonal to $\mathbf{v}$, we get another line that is known as *normal line* to the trajectory.

Definition (Normal line to a trajectory)

Given a trajectory $f(t)$ in the real plane $\mathbb{R}^2$, the *normal line* to the graph of f at moment $t = a$ is the line with equation

$$l : (x, y) = (x(a), y(a)) + t(y'(a), -x'(a)) = (x(a) + ty'(a), y(a) - tx'(a)).$$

The Cartesian equation is

$$\frac{x - x(a)}{y'(a)} = \frac{y - y(a)}{-x'(a)},$$

and the point-slope equation is

$$y - y(a) = \frac{-x'(a)}{y'(a)}(x - x(a)).$$

Considering again the trajectory of the unit circumference $f(t) = (\cos t, \sin t)$, $t \in \mathbb{R}$, the normal line to the graph of f at moment $t = \pi/4$ is

$$l : (x, y) = (\cos(\pi/2), \sin(\pi/2)) + t(\cos(\pi/2), \sin(\pi/2)) =$$
$$= \left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right) + t\left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right) = \left(\frac{\sqrt{2}}{2} + t\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2} + t\frac{\sqrt{2}}{2}\right),$$

the Cartesian equation is

$$\frac{x - \sqrt{2}/2}{\sqrt{2}/2} = \frac{y - \sqrt{2}/2}{\sqrt{2}/2},$$

and the point-slope equation is

$$y - \sqrt{2}/2 = \frac{\sqrt{2}/2}{\sqrt{2}/2}(x - \sqrt{2}/2) \Rightarrow y = x.$$

A particular case of tangent and normal lines to a trajectory are the tangent and normal lines to a function of one real variable. For every function $y = f(x)$, the trajectory that trace its graph is

$$g(x) = (x, f(x)) \quad x \in \mathbb{R},$$

and its velocity is

$$g'(x) = (1, f'(x)),$$

so that the tangent line to g at the moment a is

$$\frac{x-a}{1} = \frac{y-f(a)}{f'(a)} \Rightarrow y - f(a) = f'(a)(x-a),$$

and the normal line is

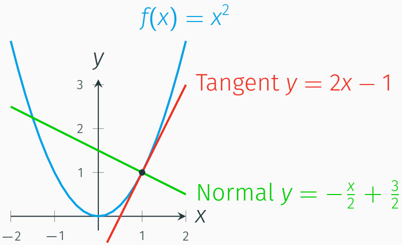
$$\frac{x-a}{f'(a)} = \frac{y-f(a)}{-1} \Rightarrow y - f(a) = \frac{-1}{f'(a)}(x-a),$$

Given the function $y = x^2$, the trajectory that traces its graph is $g(x) = (x, x^2)$ and its velocity is $g'(x) = (1, 2x)$. At the moment $x = 1$ the trajectory passes through the point $(1, 1)$ with a velocity $(1, 2)$. Thus, the tangent line at that moment is

$$\frac{x-1}{1} = \frac{y-1}{2} \Rightarrow y - 1 = 2(x-1) \Rightarrow y = 2x - 1,$$

and the normal line is

$$\frac{x-1}{2} = \frac{y-1}{-1} \Rightarrow y - 1 = \frac{-1}{2}(x-1) \Rightarrow y = \frac{-x}{2} + \frac{3}{2}.$$



Tangent line to a trajectory in the space

The concept of tangent line to a trajectory can be easily extended from the real plane to the three-dimensional space $\mathbb{R}^3$.

If $f(t) = (x(t), y(t), z(t))$, $t \in \mathbb{R}$, is a trajectory in the real space $\mathbb{R}^3$, then at the moment a, the moving object that follows this trajectory will be at the position $P = (x(a), y(a), z(a))$ with a velocity $\mathbf{v} = f'(t) = (x'(t), y'(t), z'(t))$. Thus, the tangent line to f at this moment have the following vectorial equation

$$\begin{aligned} l : (x, y, z) &= (x(a), y(a), z(a)) + t(x'(a), y'(a), z'(a)) = \\ &= (x(a) + tx'(a), y(a) + ty'(a), z(a) + tz'(a)), \end{aligned}$$

and the Cartesian equations are

$$\frac{x - x(a)}{x'(a)} = \frac{y - y(a)}{y'(a)} = \frac{z - z(a)}{z'(a)},$$

provided that $x'(a) \neq 0$, $y'(a) \neq 0$ y $z'(a) \neq 0$.

Given the trajectory $f(t) = (\cos t, \sin t, t)$, $t \in \mathbb{R}$ in the real space, at the moment $t = \pi/2$ the trajectory passes through the point

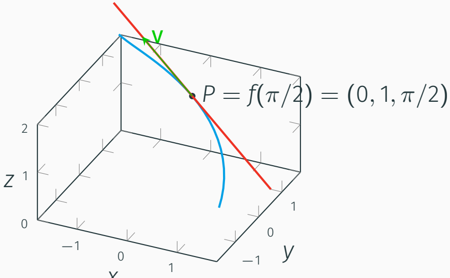
$$f(\pi/2) = (\cos(\pi/2), \sin(\pi/2), \pi/2) = (0, 1, \pi/2),$$

with velocity

$$\mathrm{v} = f'(\pi/2) = (-\sin(\pi/2), \cos(\pi/2), 1) = (-1, 0, 1),$$

and the tangent line to the graph of f at that moment is

$$l : (x, y, z) = (0, 1, \pi/2) + t(-1, 0, 1) = (-t, 1, t + \pi/2).$$



Normal plane to a trajectory in the space

In the three-dimensional space $\mathbb{R}^3$, the normal line to a trajectory is not unique. There are an infinite number of normal lines and all of them are in the normal plane.

If $f(t) = (x(t), y(t), z(t))$, $t \in \mathbb{R}$, is a trajectory in the real space $\mathbb{R}^3$, then at the moment a, the moving object that follows this trajectory will be at the position $P = (x(a), y(a), z(a))$ with a velocity $\mathbf{v} = f'(t) = (x'(t), y'(t), z'(t))$. Thus, using the velocity vector as normal vector the normal plane to f at this moment have the following vectorial equation

$$\begin{aligned}\Pi : &(x - x(a), y - y(a), z - z(a))(x'(a), y'(a), z'(a)) = 0\\ &= x'(a)(x - x(a)) + y'(a)(y - y(a)) + z'(a)(z - z(a)) = 0.\end{aligned}$$

For the trajectory of the previous example $f(t) = (\cos t, \sin t, t)$, $t \in \mathbb{R}$, at the moment $t = \pi/2$ the trajectory passes through the point

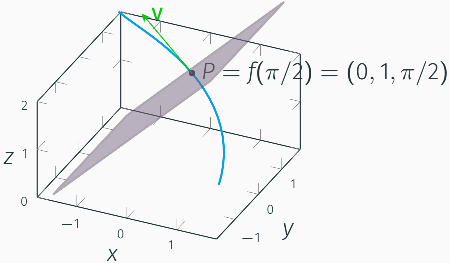
$$f(\pi/2) = (\cos(\pi/2), \sin(\pi/2), \pi/2) = (0, 1, \pi/2),$$

with velocity

$$\mathbf{v} = f'(\pi/2) = (-\sin(\pi/2), \cos(\pi/2), 1) = (-1, 0, 1),$$

and normal plane to the graph of f at that moment is

$$\Pi : \left(x - 0, y - 1, z - \frac{\pi}{2}\right)(-1, 0, 1) = 0 \Leftrightarrow -x + z - \frac{\pi}{2} = 0.$$



A lot of problems in Geometry, Physics, Chemistry, Biology, etc. involve a variable that depend on two or more variables:

- The area of a triangle depends on two variables that are the base and height lengths.
- The volume of a perfect gas depends on two variables that are the pressure and the temperature.
- The way travelled by an object free falling depends on a lot of variables: the time, the area of the cross section of the object, the latitude and longitude of the object, the height above the sea level, the air pressure, the air temperature, the speed of wind, etc.

These dependencies are expressed with functions of several variables.

Definition (Functions of several real variables)

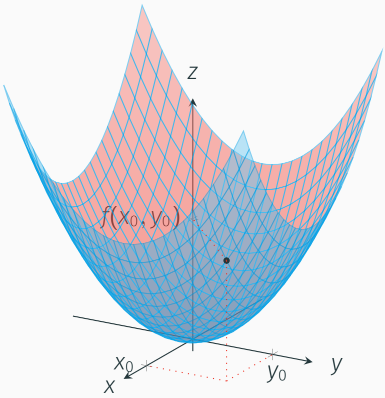
A *function of n real variables* or a *scalar field* from a set $A_1 \times \cdots \times A_n \subseteq \mathbb{R}^n$ in a set $B \subseteq \mathbb{R}$, is a relation that maps any tuple $(a_1, \ldots, a_n) \in A_1 \times \cdots \times A_n$ into a unique element of B, denoted by $f(a_1, \ldots, a_n)$, that is knwon as the *image* of $(a_1, \ldots, a_n)$ by f.

$$\begin{array}{rccc} f: & A_1 \times \cdots \times A_n & \longrightarrow & B \\ & (a_1, \ldots, a_n) & \longrightarrow & f(a_1, \ldots, a_n) \end{array}$$

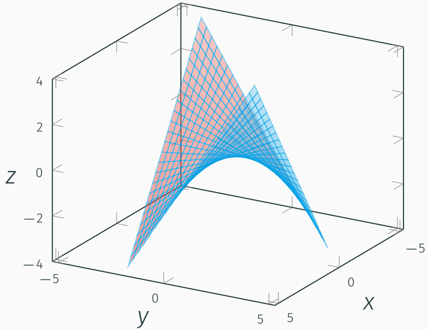
Example

- The area of a triangle is a real function of two real variables
$$f(x, y) = \frac{xy}{2}.$$
- The volume of a perfect gas is a real function of two real variables
$$v = f(t, p) = \frac{nRt}{p}, \quad \text{with } n \text{ and } R \text{ constants.}$$

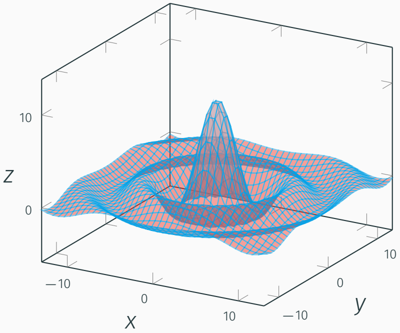
The graph of a function of two variables $f(x, y)$ is a surface in the real space $\mathbb{R}^3$ where every point of the surface has coordinates (x, y, z), with $z = f(x, y)$.



The function $f(x, y) = \dfrac{xy}{2}$ that measures the area of a triangle of base x and height y has the graph below.



The function $f(x, y) = \dfrac{\sin(x^2 + y^2)}{\sqrt{x^2 + y^2}}$ has the peculiar graph below.



Definition (Level set)

Given a scalar field $f : \mathbb{R}^n \to \mathbb{R}$, the *level set* c of f is the set

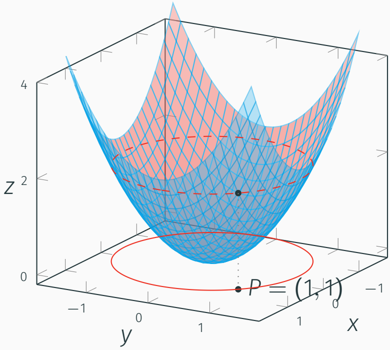
$$C_{f,c} = \{(x_1, \ldots, x_n) : f(x_1, \ldots, x_n) = c\},$$

that is, a set where the function takes on the constant value c.

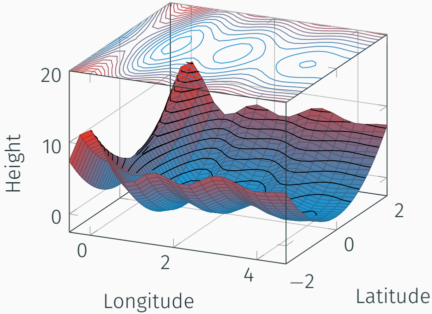
Given the scalar field $f(x, y) = x^2 + y^2$ and the point $P = (1, 1)$, the level set of f that includes P is

$$C_{f,2} = \{(x, y) : f(x, y) = f(1, 1) = 2\} = \{(x, y) : x^2 + y^2 = 2\},$$

that is the circumference of radius $\sqrt{2}$ centered at the origin.

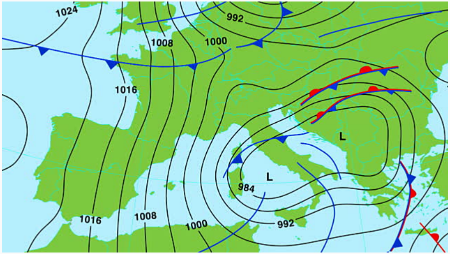


Topographic maps



Level curves correspond to points with the height above the sea level.

Weather maps (Isobars)



Level curves correspond to points with the same atmospheric pressure.

Definition (Partial function)

Given a scalar field $f : \mathbb{R}^n \to \mathbb{R}$, an i-th *partial function* of f is any function $f_i : \mathbb{R} \to \mathbb{R}$ that results of substituting all the variables of f by constants, except the i-th variable, that is:

$$f_i(x) = f(c_1, \ldots, c_{i-1}, x, c_{i+1}, \ldots, c_n),$$

with c_j $(j = 1, \ldots, n,\ j \neq i)$ constants.

Example If we take the function that measures the area of a triangle

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

and set the value of the base to $x = c$, then we the area of the triangle depends only of the height, and f becomes a function of one variable, that is the partial function

$$f_1(y) = f(c, y) = \frac{cy}{2}, \quad \text{with } c \text{ constant.}$$

We can measure the variation of a scalar field with respect to each of its variables in the same way that we measured the variation of a one-variable function.

Let $z = f(x, y)$ be a scalar field of $\mathbb{R}^2$. If we are at point (x_0, y_0) and we increase the value of x a quantity Δx, then we move in the direction of the x-axis from the point (x_0, y_0) to the point $(x_0 + \Delta x, y_0)$, and the variation of the function is

$$\Delta z = f(x_0 + \Delta x, y_0) - f(x_0, y_0).$$

Thus, the rate of change of the function with respect to x along the interval $[x_0, x_0 + \Delta x]$ is given by the quotient

$$\frac{\Delta z}{\Delta x} = \frac{f(x_0 + \Delta x, y_0) - f(x_0, y_0)}{\Delta x}.$$

Instantaneous rate of change of a scalar field with respect to a variable

If instead o measuring the rate of change in an interval, we measure the rate of change in a point, that is, when Δx approaches 0, then we get the instantaneous rate of change:

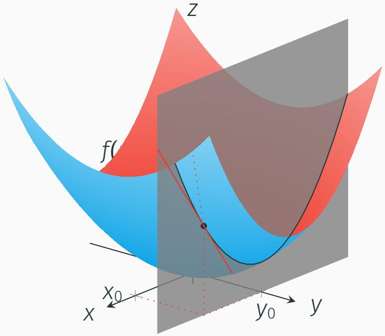
$$\lim_{\Delta x \to 0} \frac{\Delta z}{\Delta x} = \lim_{\Delta x \to 0} \frac{f(x_0 + \Delta x, y_0) - f(x_0, y_0)}{\Delta x}.$$

The value of this limit, if exists, it is known as the *partial derivative* of f with respect to the variable x at the point (x_0, y_0); it is written as

$$\frac{\partial f}{\partial x}(x_0, y_0).$$

This partial derivative measures the instantaneous rate of change of f at the point $P = (x_0, y_0)$ when P moves in the x-axis direction.

Geometrically, a two-variable function $z = f(x, y)$ defines a surface. If we cut this surface with a plane of equation $y = y_0$ (that is, the plane where y is the constant y_0) the intersection is a curve, and the partial derivative of f with respect to to x at (x_0, y_0) is the slope of the tangent line to that curve at $x = x_0$.



The concept of partial derivative can be extended easily from two-variable function to n-variables functions.

Definition (Partial derivative)

Given a n-variables function $f(x_1, \ldots, x_n)$, f is *partially differentiable* with respect to the variable x_i at the point $a = (a_1, \ldots, a_n)$ if exists the limit

$$\lim_{\Delta x_i \to 0} \frac{f(a_1, \ldots, a_{i-1}, a_i + \Delta x_i, a_{i+1}, \ldots, a_n) - f(a_1, \ldots, a_{i-1}, a_i, a_{i+1}, \ldots, a_n)}{h}.$$

In such a case, the value of the limit is known as *partial derivative* of f with respect to x_i at a; it is denoted

$$f'_{x_i}(a) = \frac{\partial f}{\partial x_i}(a).$$

Remark The definition of derivative for one-variable functions is a particular case of this definition for $n = 1$.

When we measure the variation of f with respect to a variable x_i at the point $a = (a_1, \ldots, a_n)$, the other variables remain constant. Thus, if we can consider the i-th partial function

$$f_i(x_i) = f(a_1, \ldots, a_{i-1}, x_i, a_{i+1}, \ldots, a_n),$$

the partial derivative of f with respect to x_i can be computed differentiating this function:

$$\frac{\partial f}{\partial x_i}(a) = f_i'(a_i).$$

Rule

To differentiate partially $f(x_1, \ldots, x_n)$ with respect to the variable x_i, you have to differentiate f as a function of the variable x_i, considering the other variables as constants.

Consider the function that measures the volume of a perfect gas

$$v(t,p) = \frac{nRt}{p},$$

where t is the temperature, p the pressure and n and R are constants.

The instantaneous rate of change of the volume with respect to the pressure is the partial derivative of v with respect to p. To compute this derivative we have to think in t as a constant and differentiate v as if the unique variable was p:

$$\frac{\partial v}{\partial p}(t,p) = \frac{d}{dp}\left(\frac{nRt}{p}\right)_{t\,=\text{cst}} = \frac{-nRt}{p^2}.$$

In the same way, the instantaneous rate of change of the volume with respect to the temperature is the partial derivative of v with respect to t:

$$\frac{\partial v}{\partial t}(t,p) = \frac{d}{dt}\left(\frac{nRt}{p}\right)_{p\,=\text{cst}} = \frac{nR}{p}.$$

Definition (Gradient)

Given a scalar field $f(x_1, \ldots, x_n)$, the *gradient* of f, denoted by ∇f, is a function that maps every point $a = (a_1, \ldots, a_n)$ to a vector with coordinates the partial derivatives of f at a,

$$\nabla f(a) = \left(\frac{\partial f}{\partial x_1}(a), \ldots, \frac{\partial f}{\partial x_n}(a) \right).$$

Later we will show that the gradient in a point is a vector with the magnitude and direction of the maximum rate of change of the function in that point. Thus, *$\nabla f(a)$ shows the direction of maximum increase of f at a*, while $-\nabla f(a)$ show the direction of maximum decrease of f at a.

After heating a surface, the temperature t (in °C) at each point (x, y, z) (in m) of the surface is given by the function

$$t(x, y, z) = \frac{x}{y} + z^2.$$

In what direction will increase the temperature faster at point $(2, 1, 1)$ *of the surface? What magnitude will the maximum increase of temperature have?*

The direction of maximum increase of the temperature is given by the gradient

$$\nabla t(x, y, z) = \left(\frac{\partial t}{\partial x}(x, y, z), \frac{\partial t}{\partial y}(x, y, z), \frac{\partial t}{\partial z}(x, y, z)\right) = \left(\frac{1}{y}, \frac{-x}{y^2}, 2z\right).$$

At point $(2, 1, 1)$ de direction is given by the vector

$$\nabla t(2, 1, 1) = \left(\frac{1}{1}, \frac{-2}{1^2}, 2 \cdot 1\right) = (1, -2, 2),$$

and its magnitude is

$$|\nabla f(2, 1, 1)| = |\sqrt{1^2 + (-2)^2 + 2^2}| = |\sqrt{9}| = 3 \text{ °C/m}.$$

If $f : \mathbb{R}^n \to \mathbb{R}$ is a scalar field and $g : \mathbb{R} \to \mathbb{R}^n$ is a vectorial function, then it is possible to compound g with f, so that $f \circ g : \mathbb{R} \to \mathbb{R}$ is a one-variable function.

Theorem (Chain rule)

If $g(t) = (x_1(t), \ldots, x_n(t))$ is a vectorial function differentiable at t and $f(x_1, \ldots, x_n)$ is a scalar field differentiable at the point $g(t)$, then $f \circ g(t)$ is differentiable at t and

$$(f \circ g)'(t) = \nabla f(g(t)) \cdot g'(t) = \frac{\partial f}{\partial x_1}\frac{dx_1}{dt} + \cdots + \frac{\partial f}{\partial x_n}\frac{dx_n}{dt}$$

Let us consider the scalar field $f(x,y) = x^2 y$ and the vectorial function $g(t) = (\cos t, \sin t)\ t \in [0, 2\pi]$ in the real plane, then

$$\nabla f(x,y) = (2xy, x^2) \quad \text{and} \quad g'(t) = (-\sin t, \cos t),$$

and

$$(f \circ g)'(t) = \nabla f(g(t)) \cdot g'(t) = (2\cos t \sin t, \cos^2 t) \cdot (-\sin t, \cos t) =$$
$$= -2\cos t \sin^2 t + \cos^3 t.$$

We can get the same result differentiating the composed function directly

$$(f \circ g)(t) = f(g(t)) = f(\cos t, \sin t) = \cos^2 t \sin t,$$

and its derivative is

$$(f \circ g)'(t) = 2\cos t(-\sin t)\sin t + \cos^2 t \cos t = -2\cos t \sin^2 t + \cos^3 t.$$

The chain rule for the composition of a vectorial function with a scalar field allow us to get the algebra of derivatives for one-variable functions easily:

$$\begin{aligned}(u+v)' &= u'+v'\\ (uv)' &= u'v+uv'\\ \left(\frac{u}{v}\right)' &= \frac{u'v-uv'}{v^2}\\ (u\circ v)' &= u'(v)v'\end{aligned}$$

To infer the derivative of the sum of two functions u and v, we can take the scalar field $f(x,y)=x+y$ and the vectorial function $g(t)=(u(t),v(t))$. Applying the chain rule we get

$$(u+v)'(t)=(f\circ g)'(t)=\nabla f(g(t))\cdot g'(t)=(1,1)\cdot(u',v')=u'+v'.$$

To infer the derivative of the quotient of two functions u and v, we can take the scalar field $f(x,y)=x/y$ and the vectorial function $g(t)=(u(t),v(t))$.

$$\left(\frac{u}{v}\right)'(t)=(f\circ g)'(t)=\nabla f(g(t))\cdot g'(t)=\left(\frac{1}{v},-\frac{u}{v^2}\right)\cdot(u',v')=\frac{u'v-uv'}{v^2}.$$

For a scalar field $f(x, y)$, we have seen that the partial derivative $\dfrac{\partial f}{\partial x}(x_0, y_0)$ is the instantaneous rate of change of f with respect to x at point $P = (x_0, y_0)$, that is, when we move along the x-axis.

In the same way, $\dfrac{\partial f}{\partial y}(x_0, y_0)$ is the instantaneous rate of change of f with respect to y at the point $P = (x_0, y_0)$, that is, when we move along the y-axis.

But, *what happens if we move along any other direction?*

The instantaneous rate of change of f at the point $P = (x_0, y_0)$ along the direction of a unitary vector u is known as *directional derivative*.

Definition (Directional derivative)

Given a scalar field f of $\mathbb{R}^n$, a point P and a unitary vector $\mathbf{u}$ in that space, we say that f is differentiable at P along the direction of $\mathbf{u}$ if exists the limit

$$f'_{\mathbf{u}}(P) = \lim_{h\to 0} \frac{f(P + h\mathbf{u}) - f(P)}{h}.$$

In such a case, the value of the limit is known as *directional derivative* of f at the point P along the direction of $\mathbf{u}$.

If we consider a unitary vector $\mathbf{u}$, the trajectory that passes through P, following the direction of $\mathbf{u}$, has equation

$$g(t) = P + t\mathbf{u},\ t \in \mathbb{R}.$$

For $t = 0$, this trajectory passes through the point $P = g(0)$ with velocity $\mathbf{u} = g'(0)$.

Thus, the directional derivative of f at the point P along the direction of $\mathbf{u}$ is

$$(f \circ g)'(0) = \nabla f(g(0)) \cdot g'(0) = \nabla f(P) \cdot \mathbf{u}.$$

Given the function $f(x,y) = x^2 + y^2$, its gradient is

$$\nabla f(x,y) = (2x, 2y).$$

The directional derivative of f at the point $P = (1,1)$, along the unit vector $\mathbf{u} = (1/\sqrt{2}, 1/\sqrt{2})$ is

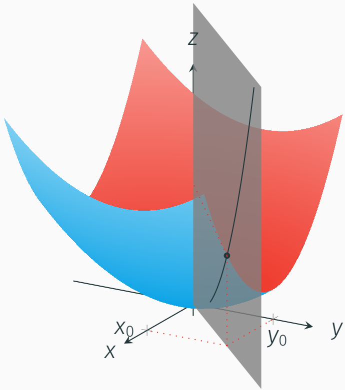
$$f'_{\mathbf{u}}(P) = \nabla f(P) \cdot \mathbf{u} = (2,2) \cdot (1/\sqrt{2}, 1/\sqrt{2}) = \frac{2}{\sqrt{2}} + \frac{2}{\sqrt{2}} = \frac{4}{\sqrt{2}}.$$

To compute the directional derivative along a non-unitary vector $\mathbf{v}$, we have to use the unitary vector that results from normalizing v with the transformation:

$$\mathbf{v}' = \frac{\mathbf{v}}{|\mathbf{v}|}.$$

Geometric interpretation of the directional derivative

Geometrically, a two-variable function $z = f(x, y)$ defines a surface. If we cut this surface with a plane of equation $a(y - y_0) = b(x - x_0)$ (that is, the vertical plane that passes through the point $P = (x_0, y_0)$ with the direction of vector $\mathbf{u} = (a, b)$) the intersection is a curve, and the directional derivative of f at P along the direction of $\mathbf{u}$ is the slope of the tangent line to that curve at point P.



Growth of scalar field along the gradient

We have seen that for any vector $\mathbf{u}$

$$f'_{\mathbf{u}}(P) = \nabla f(P) \cdot \mathbf{u} = |\nabla f(P)| \cos\theta,$$

where θ is the angle between $\mathbf{u}$ and the gradient $\nabla f(P)$.

Taking into account that $-1 \leq \cos\theta \leq 1$, for any vector $\mathbf{u}$ it is satisfied that

$$-|\nabla f(P)| \leq f'_{\mathbf{u}}(P) \leq |\nabla f(P)|.$$

Furthermore, if $\mathbf{u}$ has the same direction and sense than the gradient, we have $f'_{\mathbf{u}}(P) = |\nabla f(P)| \cos 0 = |\nabla f(P)|$. Therefore, *the maximum increase of a scalar field at a point P is along the direction of the gradient at that point*.

In the same manner, if $\mathbf{u}$ has the same direction but opposite sense than the gradient, we have $f'_{\mathbf{u}}(P) = |\nabla f(P)| \cos\pi = -|\nabla f(P)|$. Therefore, *the maximum decrease of a scalar field at a point P is along the opposite direction of the gradient at that point*.

When we have a relation $f(x, y) = 0$, sometimes we can consider y as an *implicit function* of x, at least in a neighborhood of a point (x_0, y_0).

Example The equation $x^2 + y^2 = 25$, whose graph is the circle of radius 5 centered at the origin of coordinates, its not a function, because if we solve the equation for y, we have two images for some values of x,

$$y = \pm\sqrt{25 - x^2}$$

However, near the point $(3, 4)$ we can represent the relation as the function $y = \sqrt{25 - x^2}$, and near the point $(3, -4)$ we can represent the relation as the function $y = -\sqrt{25 - x^2}$.

If an equation $f(x, y) = 0$ defines y as a implicit function of x, $y = h(x)$, in a neighborhood of (x_0, y_0), then we can compute de derivative of y, $h'(x)$, even if we do not know the explicit formula for h.

Theorem (Implicit function (one-variable))

Let $f(x, y) : \mathbb{R}^2 \longrightarrow \mathbb{R}$ a two-variable function and let (x_0, y_0) be a point in $\mathbb{R}^2$ such that $f(x_0, y_0) = 0$. If f has partial derivatives continuous at (x_0, y_0) and $\frac{\partial f}{\partial y}(x_0, y_0) \neq 0$, then there is an open interval $I \subset \mathbb{R}$ with $x_0 \in I$ and a function $h(x) : I \longrightarrow \mathbb{R}$ such that

1. $y_0 = h(x_0)$.
2. $f(x, h(x)) = 0$ *for all* $x \in I$.
3. *h is differentiable on I, and* $y' = h'(x) = \dfrac{-\dfrac{\partial f}{\partial x}}{\dfrac{\partial f}{\partial y}}$

To prove the last result, take the trajectory $g(x) = (x, h(x))$ on the interval I. Then

$$(f \circ g)(x) = f(g(x)) = f(x, h(x)) = 0.$$

Thus, using the chain rule we have

$$(f \circ g)'(x) = \nabla f(g(x)) \cdot g'(x) = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right) \cdot (1, h'(x)) = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y} h'(x) = 0,$$

from where we can deduce

$$y' = h'(x) = \frac{-\dfrac{\partial f}{\partial x}}{\dfrac{\partial f}{\partial y}}$$

This technique that allows us to compute y' in a neighborhood of x_0 without the explicit formula of $y = h(x)$, it is known as *implicit derivation*.

Consider the equation of the circle of radius 5 centered at the origin $x^2 + y^2 = 25$. It can also be written as

$$f(x, y) = x^2 + y^2 - 25 = 0.$$

Take the point $(3, 4)$ that satisfies the equation, $f(3, 4) = 0$.

As f have partial derivatives $\frac{\partial f}{\partial x} = 2x$ and $\frac{\partial f}{\partial y} = 2y$, that are continuous at $(3, 4)$, and $\frac{\partial f}{\partial y}(3, 4) = 8 \neq 0$, then y can be expressed as a function of x in a neighborhood of $(3, 4)$ and its derivative is

$$y' = \frac{-\frac{\partial f}{\partial x}}{\frac{\partial f}{\partial y}} = \frac{-2x}{2y} = \frac{-x}{y} \quad \textit{and} \quad y'(3) = \frac{-3}{4}.$$

In this particular case, that we know the explicit formula of $y = \sqrt{1 - x^2}$, we can get the same result computing the derivative as usual

$$y' = \frac{1}{2\sqrt{1 - x^2}}(-2x) = \frac{-x}{\sqrt{1 - x^2}}.$$

The implicit function theorem can be generalized to functions with several variables.

Theorem (Implicit function)

Let $f(x_1, \ldots, x_n, y) : \mathbb{R}^{n+1} \longrightarrow \mathbb{R}$ a $n+1$-variables function and let $(x_1^0, \ldots, x_2^0, y^0)$ be a point in $\mathbb{R}^{n+1}$ such that $f(x_1^0, \ldots, x_n^0, y^0) = 0$. If f has partial derivatives continuous at $(x_1^0, \ldots, x_n^0, y^0)$ and $\frac{\partial f}{\partial y}(x_1^0, \ldots, x_n^0, y^0) \neq 0$, then there is a region $I \subset \mathbb{R}^n$ with $(x_1^0, \ldots, x_n^0) \in I$ and a function $h(x1, \ldots, xn) : I \longrightarrow \mathbb{R}$ such that

1. $y_0 = h(x_1^0, \ldots, x_n^0)$.
2. $f(x_1, \ldots, x_n, h(x_1, \ldots, x_n)) = 0$ *for all* $(x_1, \ldots, x_n) \in I$.
3. *h is differentiable on I, and* $\dfrac{\partial y}{\partial x_i} = \dfrac{-\dfrac{\partial f}{\partial x_i}}{\dfrac{\partial f}{\partial y}}$

Theorem

Let C be the level set of a scalar field f that includes a point P. If $\mathbf{v}$ is the velocity at P of a trajectory following C, then

$$\nabla f(P) \cdot \mathbf{v} = 0.$$

that is, the gradient of f at P is normal to C at P, provided that the gradient is not zero.

Proof If we take the trajectory $g(t)$ that follows the level set C and passes through P at time $t = t_0$, that is $P = g(t_0)$, so $\mathbf{v} = g'(t_0)$, then

$$(f \circ g)(t) = f(g(t)) = f(P),$$

that is constant at any t. Thus, applying the chain rule we have

$$(f \circ g)'(t) = \nabla f(g(t)) \cdot g'(t) = 0,$$

and, particularly, at $t = t_0$, we have

$$\nabla f(P) \cdot \mathbf{v} = 0.$$

According to the previous result, the normal line to a curve with equation $f(x, y) = 0$ at point $P = (x_0, y_0)$, has equation

$$P + t\nabla f(P) = (x_0, y_0) + t\nabla f(x_0, y_0).$$

Example Given the scalar field $f(x, y) = x^2 + y^2 - 25$, and the point $P = (3, 4)$, the level set of f that passes through P, that satisfies $f(x, y) = f(P) = 0$, is the circle with radius 5 centered at the origin of coordinates. Thus, taking as a normal vector the gradient of f

$$\nabla f(x, y) = (2x, 2y),$$

at the point $P = (3, 4)$ is $\nabla f(3, 4) = (6, 8)$, and the normal line to the circle at P is

$$P + t\nabla f(P) = (3, 4) + t(6, 8) = (3 + 6t, 4 + 8t),$$

On the other hand, the tangent line to the circle at P is

$$((x, y) - P) \cdot \nabla f(P) = ((x, y) - (3, 4)) \cdot (6, 8) = (x - 3, y - 4) \cdot (6, 8) = 6x + 8y = 50.$$

In the same way, if we have a surface with equation $f(x,y,z)=0$, at the point $P=(x_0,y_0,z_0)$ the normal line has equation

$$P+t\nabla f(P)=(x_0,y_0,z_0)+t\nabla f(x_0,y_0,z_0).$$

Example Given the scalar field $f(x,y,z)=x^2+y^2-z$, and the point $P=(1,1,2)$, the level set of f that passes through P, that satisfies $f(x,y)=f(P)=0$, is the paraboloid $z=x^2+y^2$. Thus, taking as a normal vector the gradient of f

$$\nabla f(x,y,z)=(2x,2y,-1),$$

at the point $P=(1,1,2)$ is $\nabla f(1,1,2)=(2,2,-1)$, and the normal line to the paraboloid at P is

$$P+t\nabla f(P)=(1,1,2)+t\nabla f(1,1,2)=(1,1,2)+t(2,2,-1)=(1+2t,1+2t,2-t).$$

On the other hand, the tangent plane to the paraboloid at P is

$$\begin{aligned}((x,y,z)-P)\cdot\nabla f(P)&=((x,y,z)-(1,1,2))(2,2,-1)=(x-1,y-1,z-2)(2,2,-1)\\&=2(x-1)+2(y-1)-(z-2)=2x+2y-z-2=0.\end{aligned}$$

As the partial derivatives of a function are also functions of several variables we can differentiate partially each of them.

If a function $f(x_1,\ldots,x_n)$ has a partial derivative $f'_{x_i}(x_1,\ldots,x_n)$ with respect to the variable x_i in a set A, then we can differentiate partially again f'_{x_i} with respect to the variable x_j. This second derivative, when exists, is known as *second order partial derivative* of f with respect to the variables x_i and x_j; it is written as

$$\frac{\partial^2 f}{\partial x_j \partial x_i} = \frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}\right).$$

In the same way we can define higher order partial derivatives.

The two-variable function

$$f(x, y) = x^y$$

has 4 second order partial derivatives:

$$\frac{\partial^2 f}{\partial x^2}(x, y) = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}(x, y)\right) = \frac{\partial}{\partial x}\left(yx^{y-1}\right) = y(y-1)x^{y-2},$$

$$\frac{\partial^2 f}{\partial y \partial x}(x, y) = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}(x, y)\right) = \frac{\partial}{\partial y}\left(yx^{y-1}\right) = x^{y-1} + yx^{y-1}\log x,$$

$$\frac{\partial^2 f}{\partial x \partial y}(x, y) = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}(x, y)\right) = \frac{\partial}{\partial x}\left(x^y \log x\right) = yx^{y-1}\log x + x^y \frac{1}{x},$$

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

Definition (Hessian matrix)

Given a scalar field $f(x_1, \ldots, x_n)$, with second order partial derivatives at the point $a = (a_1, \ldots, a_n)$, the *Hessian matrix* of f at a, denoted by $\nabla^2 f(a)$, is the matrix

$$\nabla^2 f(a) = \begin{pmatrix} \dfrac{\partial^2 f}{\partial x_1^2}(a) & \dfrac{\partial^2 f}{\partial x_1 \partial x_2}(a) & \cdots & \dfrac{\partial^2 f}{\partial x_1 \partial x_n}(a) \\ \dfrac{\partial^2 f}{\partial x_2 \partial x_1}(a) & \dfrac{\partial^2 f}{\partial x_2^2}(a) & \cdots & \dfrac{\partial^2 f}{\partial x_2 \partial x_n}(a) \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial^2 f}{\partial x_n \partial x_1}(a) & \dfrac{\partial^2 f}{\partial x_n \partial x_2}(a) & \cdots & \dfrac{\partial^2 f}{\partial x_n^2}(a) \end{pmatrix}$$

The determinant of this matrix is known as *Hessian* of f at a; it is denoted $Hf(a) = |\nabla^2 f(a)|$.

Consider again the two-variable function

$$f(x,y) = x^y.$$

Its Hessian matrix is

$$\nabla^2 f(x,y) = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{pmatrix} = \begin{pmatrix} y(y-1)x^{y-2} & x^{y-1}(y \log x + 1) \\ x^{y-1}(y \log x + 1) & x^y (\log x)^2 \end{pmatrix}.$$

At point $(1,2)$ is

$$\nabla^2 f(1,2) = \begin{pmatrix} 2(2-1)1^{2-2} & 1^{2-1}(2\log 1 + 1) \\ 1^{2-1}(2\log 1 + 1) & 1^2(\log 1)^2 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix}.$$

And its Hessian is

$$Hf(1,2) = \begin{vmatrix} 2 & 1 \\ 1 & 0 \end{vmatrix} = 2 \cdot 0 - 1 \cdot 1 = -1.$$

In the previous example we can observe that the *mixed derivatives* of second order $\frac{\partial^2 f}{\partial y \partial x}$ and $\frac{\partial^2 f}{\partial x \partial y}$ are the same. This fact is due to the following result.

Theorem (Symmetry of second partial derivatives)

If $f(x_1, \dots, x_n)$ is a scalar field with second order partial derivatives $\frac{\partial^2 f}{\partial x_i \partial x_j}$ and $\frac{\partial^2 f}{\partial x_j \partial x_i}$ continuous at a point $(a_1, \dots, a_n)$, then

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(a_1, \dots, a_n) = \frac{\partial^2 f}{\partial x_j \partial x_i}(a_1, \dots, a_n).$$

This means that *the order of the partial derivatives does not matter* when computing a second partial derivative.

As a consequence, if the function satisfies the requirements of the theorem for all the second order partial derivatives, the Hessian matrix is symmetric.

Linear approximation of a scalar field

In a previous chapter we saw how to approximate a one-variable function with a Taylor polynomial. This can be generalized to several-variables functions.

If P is a point in the domain of a scalar field f and $\mathbf{v}$ is a vector, the first degree *Taylor formula* of f around P is

$$f(P+\mathbf{v}) = f(P) + \nabla f(P)\cdot\mathbf{v} + R^1_{f,P}(\mathbf{v}),$$

where

$$P^1_{f,P}(\mathbf{v}) = f(P) + \nabla f(P)\mathbf{v}$$

is the first degree *Taylor polynomial* of f at P, and $R^1_{f,P}(\mathbf{v})$ is the *Taylor remainder* for the vector $\mathbf{v}$, that is the error in the approximation.

The remainder satisfies

$$\lim_{|\mathbf{v}|\to 0} \frac{R^1_{f,P}(\mathbf{v})}{|\mathbf{v}|} = 0$$

Observe that the first degree Taylor polynomial for a function of two variables is the tangent plange to the graph of f at P.

If f is a scalar field of two variables $f(x, y)$ and $P = (x_0, y_0)$, as for any point $Q = (x, y)$ we can take the vector $\mathbf{v} = \vec{PQ} = (x - x_0, y - y_0)$, then the first degree Taylor polynomial of f at P, can be written as

$$\begin{aligned} P^1_{f,P}(x, y) &= f(x_0, y_0) + \nabla f(x_0, y_0)(x - x_0, y - y_0) = \\ &= f(x_0, y_0) + \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0). \end{aligned}$$

Given the scalar field $f(x,y) = \log(xy)$, its gradient is

$$\nabla f(x,y) = \left(\frac{1}{x}, \frac{1}{y}\right),$$

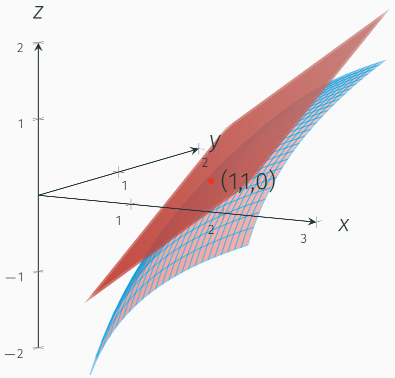
and the first degree Taylor polynomial at the point $P = (1,1)$ is

$$\begin{aligned} P^1_{f,P}(x,y) &= f(1,1) + \nabla f(1,1) \cdot (x-1, y-1) = \\ &= \log 1 + (1,1) \cdot (x-1, y-1) = x - 1 + y - 1 = x + y - 2. \end{aligned}$$

This polynomial approximates f near the point P. For instance,

$$f(1.01, 1.01) \approx P^1_{f,P}(1.01, 1.01) = 1.01 + 1.01 - 2 = 0.02.$$

The graph of the scalar field $f(x, y) = \log(xy)$ and the first degree Taylor polynomial of f at the point $P = (1, 1)$ is below.



If P is a point in the domain of a scalar field f and $\mathbf{v}$ is a vector, the second degree *Taylor formula* of f around P is

$$f(P+\mathbf{v}) = f(P) + \nabla f(P)\cdot \mathbf{v} + \frac{1}{2}\left(\mathbf{v}\nabla^2 f(P)\mathbf{v}\right) + R^2_{f,P}(\mathbf{v}),$$

where

$$P^2_{f,P}(\mathbf{v}) = f(P) + \nabla f(P)\mathbf{v} + \frac{1}{2}\left(\mathbf{v}\nabla^2 f(P)\mathbf{v}\right)$$

is the second degree *Taylor polynomial* of f at the point P, and $R^2_{f,P}(\mathbf{v})$ is the *Taylor remainder* for the vector $\mathbf{v}$, that is the error in the approximation.

The remainder satisfies

$$\lim_{|\mathbf{v}\to 0|} \frac{R^2_{f,P}(\mathbf{v})}{|\mathbf{v}|^2} = 0.$$

This means that the remainder is smaller than the square of the module of $\mathbf{v}$.

Quadratic approximation of a two-variable function

If f is a scalar field of two variables $f(x,y)$ and $P=(x_0,y_0)$, then the second degree Taylor polynomial of f at P, can be written as

$$\begin{aligned}P^2_{f,P}(x,y) &= f(x_0,y_0)+\nabla f(x_0,y_0)(x-x_0,y-y_0)+\\ &+\frac{1}{2}(x-x_0,y-y_0)\nabla^2 f(x_0,y_0)(x-x_0,y-y_0)=\\ &= f(x_0,y_0)+\frac{\partial f}{\partial x}(x_0,y_0)(x-x_0)+\frac{\partial f}{\partial y}(x_0,y_0)(y-y_0)+\end{aligned}$$

$$+\frac{1}{2}\left(\frac{\partial^2 f}{\partial x^2}(x_0,y_0)(x-x_0)^2+2\frac{\partial^2 f}{\partial y\partial x}(x_0,y_0)(x-x_0)(y-y_0)+\frac{\partial^2 f}{\partial y^2}(x_0,y_0)(y-y_0)^2\right)$$

Given the scalar field $f(x,y) = \log(xy)$, its gradient is

$$\nabla f(x,y) = \left(\frac{1}{x}, \frac{1}{y}\right),$$

its Hessian matrix is

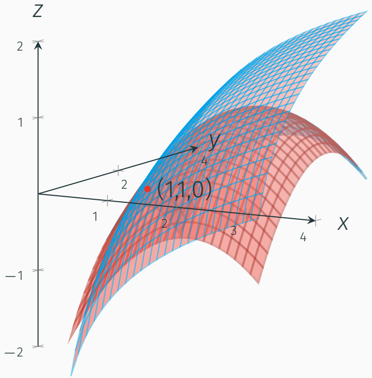
$$Hf(x,y) = \begin{pmatrix} \frac{-1}{x^2} & 0 \\ 0 & \frac{-1}{y^2} \end{pmatrix}$$

and the second degree Taylor polynomial of f at the point $P = (1,1)$ is

$$\begin{aligned}
P^2_{f,P}(x,y) &= f(1,1) + \nabla f(1,1)\cdot(x-1,y-1) + \frac{1}{2}(x-1,y-1)\nabla^2 f(1,1)\cdot(x-1,y-1) \\
&= \log 1 + (1,1)\cdot(x-1,y-1) + \frac{1}{2}(x-1,y-1)\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} x-1 \\ y-1 \end{pmatrix} = \\
&= x-1+y-1+\frac{-x^2-y^2+2x+2y-2}{2} = \frac{-x^2-y^2+4x+4y-6}{2}.
\end{aligned}$$

Thus, $\mathbf{v}f(1.01, 1.01) \approx P^1_{f,P}(1.01, 1.01) =$
$\dfrac{-1.01^2 - 1.01^2 + 4\cdot 1.01 + 4\cdot 1.01 - 6}{2} = 0.0199.$

The graph of the scalar field $f(x, y) = \log(xy)$ and the second degree Taylor polynomial of f at the point $P = (1, 1)$ is below.



Definition (Relative maximum and minimum)

A scalar field f in $\mathbb{R}^n$ has a *relative maximum* at a point P if there is a value $\epsilon > 0$ such that

$$f(P) \geq f(X)\ \forall X, |\vec{PX}| < \epsilon.$$

f has a *relative minimum* at f if there is a value $\epsilon > 0$ such that

$$f(P) \leq f(X)\ \forall X, |\vec{PX}| < \epsilon.$$

Both relative maxima and minima are known as *relative extrema* of f.

Theorem

If a scalar field f in $\mathbb{R}^n$ has a relative maximum or minimum at a point P, then P is a critical or stationary point of f, that is, a point where the gradient vanishes

$$\nabla f(P) = 0.$$

Proof Taking the trajectory that passes through P with the direction of the gradient at that point

$$g(t) = P + t\nabla f(P),$$

the function $h = (f \circ g)(t)$ does not decrease at $t = 0$ since

$$h'(0) = (f \circ g)'(0) = \nabla f(g(0)) \cdot g'(0) = \nabla f(P) \cdot \nabla f(P) = |\nabla f(P)|^2 \geq 0,$$

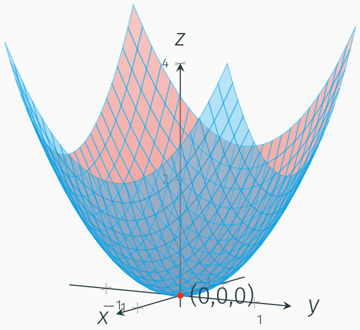
and it only vanishes if $\nabla f(P) = 0$.

Thus, if $\nabla f(P) \neq 0$, f can not have a relative maximum at P since following the trajectory of g from P there are points where f has an image greater than the image at P. In the same way, following the trajectory of g in the opposite direction there are points where f has an image less than the

Given the scalar field $f(x, y) = x^2 + y^2$, it is obvious that f only has a relative minimum at $(0, 0)$ since

$$f(0, 0) = 0 \leq f(x, y) = x^2 + y^2, \ \forall x, y \in \mathbb{R}.$$

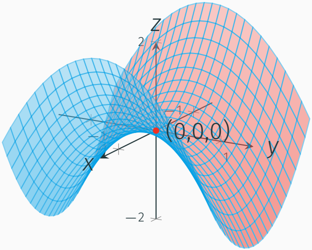
Is easy to check that f has a critical point at $(0, 0)$, that is $\nabla f(0, 0) = 0$.



Not all the critical points of a scalar field are points where the scalar field has relative extrema. If we take, for instance, the scalar field $f(x,y) = x^2 - y^2$, its gradient is

$$\nabla f(x,y) = (2x, -2y),$$

that only vanishes at $(0,0)$. However, this point is not a relative maximum since the points $(x,0)$ in the x-axis have images $f(x,0) = x^2 \geq 0 = f(0,0)$, nor a relative minimum since the points $(0,y)$ in the y-axis have images $f(0,y) = -y^2 \leq 0 = f(0,0)$. This type of critical points that are not relative extrema are known as *saddle points*.



From the second degree Taylor's formula of a scalar field f at a point P we have

$$f(P+\mathbf{v})-f(P)\approx\nabla f(P)\mathbf{v}+\frac{1}{2}\nabla^2 f(P)\mathbf{v}\cdot\mathbf{v}.$$

Thus, if P is a critical point of f, as $\nabla f(P)=0$, we have

$$f(P+\mathbf{v})-f(P)\approx\frac{1}{2}\nabla^2 f(P)\mathbf{v}\cdot\mathbf{v}.$$

Therefore, the sign of the $f(P+\mathbf{v})-f(P)$ is the sign of the second degree term $\nabla^2 f(P)\mathbf{v}\cdot\mathbf{v}$.

There are four possibilities:

- Definite positive: $\nabla^2 f(P)\mathbf{v}\cdot\mathbf{v}>0\ \forall\mathbf{v}\neq 0$.
- Definite negative: $\nabla^2 f(P)\mathbf{v}\cdot\mathbf{v}<0\ \forall\mathbf{v}\neq 0$.
- Indefinite: $\nabla^2 f(P)\mathbf{v}\cdot\mathbf{v}>0$ for some $\mathbf{v}\neq 0$ and $\nabla^2 f(P)\mathbf{u}\cdot\mathbf{u}<0$ for some $\mathbf{u}\neq 0$.
- Semidefinite: In any other case.

Thus, depending on de sign of $\nabla^2 f(P)\mathbf{v}\cdot\mathbf{v}$, we have

Theorem

Given a critical point P of a scalar field f, it holds that

- *If $\nabla^2 f(P)$ is definite positive then f has a relative minimum at P.*
- *If $\nabla^2 f(P)$ is definite negative then f has a relative maximum at P.*
- *If $\nabla^2 f(P)$ is indefinite then f has a saddle point at P.*

When $\nabla^2 f(P)$ is semidefinite we can not draw any conclusion and we need higher order partial derivatives to classify the critical point.

In the particular case of a scalar field of two variables, we have

Theorem

Given a critical point $P = (x_0, y_0)$ of a scalar field $f(x, y)$, it holds that

- *If $Hf(P) > 0$ and $\dfrac{\partial^2 f}{\partial x^2}(x_0, y_0) > 0$ then f has a relative minimum at P.*
- *If $Hf(P) > 0$ and $\dfrac{\partial^2 f}{\partial x^2}(x_0, y_0) < 0$ then f has a relative maximum at P.*
- *IF $Hf(P) < 0$ then f has a saddle point at P.*

Given the scalar field $f(x,y) = \frac{x^3}{3} - \frac{y^3}{3} - x + y$, its gradient is

$$\nabla f(x,y) = (x^2 - 1, -y^2 + 1),$$

and it has critical points at $(1,1)$, $(1,-1)$, $(-1,1)$ and $(-1,-1)$.

The hessian matrix is

$$\nabla^2 f(x,y) = \begin{pmatrix} 2x & 0 \\ 0 & -2y \end{pmatrix}$$

and the hessian is

$$Hf(x,y) = -4xy.$$

Thus, we have

- Point $(1,1)$: $Hf(1,1) = -4 < 0 \Rightarrow$ Saddle point.
- Point $(1,-1)$: $Hf(1,-1) = 4 > 0$ and $\frac{\partial^2}{\partial x^2}(1,-1) = 2 > 0 \Rightarrow$ Relative min.
- Point $(-1,1)$: $Hf(-1,1) = 4 > 0$ and $\frac{\partial^2}{\partial x^2}(-1,1) = -2 < 0 \Rightarrow$ Relative max.
- Point $(-1,-1)$: $Hf(-1,-1) = -4 < 0 \Rightarrow$ Saddle point.

Relative extrema and saddle points of the function $f(x, y) = \dfrac{x^3}{3} - \dfrac{y^3}{3} - x + y$.