Mechanics: Functions

Course Notes

Shaojun Zan

Department of Physics

Imperial College London

Contents

Chapter 1

Functions and Limits

1.1 Definition

If two variables, x and y, follow a rule: 'When x is given, then y is determined as....' Then y is said to be a FUNCTION of x.

We write $y = f(x)$:

x - Independent Variable

y - Dependent Variable

NOTE: writing $x = g(y)$ reverses the roles of the variables.

1.1.1 Domain and Range

For a set of values of x (Domain) there is a corresponding set of y values (Range).

e.g. $A = \pi r^2$:

given domain $0 \leq r \leq 2$, range of A is $0 \leq A \leq 4\pi$.

1.1.2 Common Notation

We often write $y = f(x)$, but sometimes simply $y = y(x)$.

e.g. $f(x) = x + x^2 \Rightarrow y = x + x^2$

1.2 Common Functions

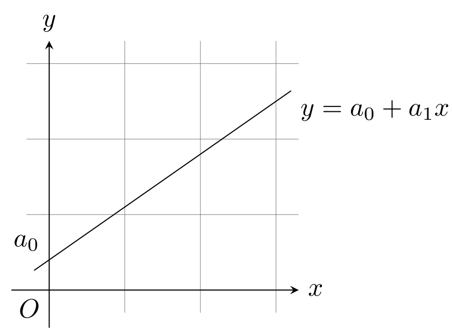
1.2.1 Polynomials

$$
\begin{aligned}
y &= a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n \\
&= \sum_{r=0}^{n} a_r x^r
\end{aligned}
\tag{1.1}
$$

$n \equiv$ DEGREE of polynomial (a finite number)

Linear Functions

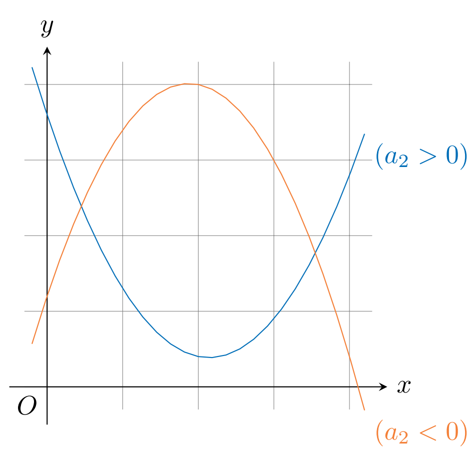
Linear functions are simply polynomials of degree 1. The constant term (a_0) is called the intercept, and the coefficient of x (a_1) is called the slope.



Quadratic Functions

Quadratic functions are simply polynomials of degree 2.

$$
y = a_0 + a_1 x + a_2 x^2 \tag{1.2}
$$



Useful Tip: Completing the Square & Mean Inequality

$$
\begin{aligned}
a_2 > 0 : a_2x^2 + a_1x + a_0 &\equiv \left(\sqrt{a_2}x + \frac{a_1}{2\sqrt{a_2}}\right)^2 - \frac{a_1^2}{4a_2} + a_0 \\
a_2 < 0 : a_2x^2 + a_1x + a_0 &\equiv -\left(|a_2|x^2 - a_1x - a_0\right) = \dots
\end{aligned}
\tag{1.3}
$$

This means the properties of a quadratic function can be found without the use of calculus.

$$x^2 - 2x\sqrt{b} + b \equiv \left(x - \sqrt{b}\right)^2 \geq 0 \quad (b \geq 0)$$

$$\text{(achieves equality only when } x = \sqrt{b}\text{)}$$

Hence,

$$a - 2\sqrt{ab} + b \equiv \left(\sqrt{a} - \sqrt{b}\right)^2 \geq 0 \quad (a, b \geq 0)$$

$$\text{(achieves equality only when } a = b\text{).}$$

By rearranging, we get what is called **The AM-GM Inequality**:

$$
\begin{gathered}
\frac{a+b}{2} \geq \sqrt{ab} \\
\text{Arithmetic Mean (AM)} \geq \text{Geometric Mean (GM)}
\end{gathered}
\tag{1.4}
$$

Apart from the two means mentioned above, there were two other means: harmonic mean and quadratic mean.

$$\text{Harmonic Mean (HM)} = \frac{2}{1/a + 1/b}$$

$$\text{Geometric Mean (GM)} = \sqrt{\frac{a^2 + b^2}{2}}$$

For two positive real numbers $(a, b > 0)$, the inequality for the means should be written as:

$$
\begin{gathered}
\frac{2}{1/a + 1/b} \leq \sqrt{ab} \leq \frac{a+b}{2} \leq \sqrt{\frac{a^2 + b^2}{2}} \\
\text{HM} \leq \text{GM} \leq \text{AM} \leq \text{QM}
\end{gathered}
\tag{1.5}
$$

Proof. First, we can deal with the Harmonic Mean.

$$\text{HM} = \frac{2}{1/a + 1/b} = \frac{2ab}{a+b}$$

According to the AM-GM Inequality, $a+b \geq 2\sqrt{ab}$. Therefore,

$$\text{HM} = \frac{2ab}{a+b} \leq \frac{2ab}{2\sqrt{ab}} = \sqrt{ab} = \text{GM} \Rightarrow \text{HM} \leq \text{GM}.$$

Now look at the Arithmetic Mean.

$$\text{AM} = \frac{a+b}{2}$$

$$\text{AM}^2 = \left(\frac{a+b}{2}\right)^2 = \frac{a^2+b^2+2ab}{4}$$

$$(a-b)^2 = a^2+b^2-2ab \geq 0 \Rightarrow 2ab \leq a^2+b^2$$

Therefore,

$$\text{AM}^2 = \frac{a^2+b^2+2ab}{4} \leq \frac{2(a^2+b^2)}{4} = \frac{a^2+b^2}{2} = \text{QM}^2.$$

As $\text{AM}, \text{GM} > 0$,

$$\text{AM}^2 \leq \text{QM}^2 \Rightarrow \text{AM} \leq \text{QM},$$

and finally,

$$\text{HM} \leq \text{GM} \leq \text{AM} \leq \text{QM}.$$

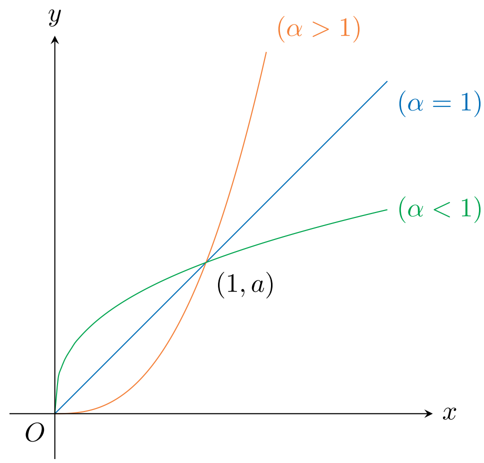
□

This is called **The HM-GM-AM-QM Inequality**. The above means are equal only when $a = b$. This could be generalized into situations including n positive real numbers, with equality achieved only when these n numbers are all equal.

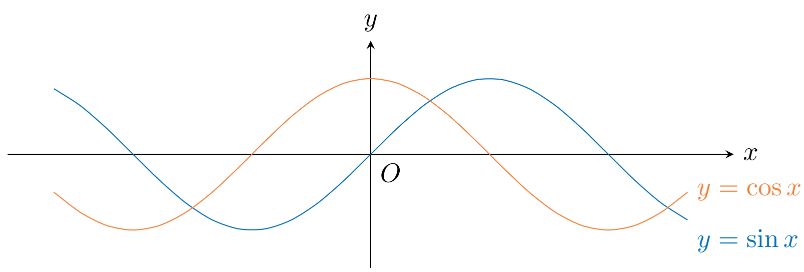
$$\begin{gathered} \frac{n}{\frac{1}{x_1}+\frac{1}{x_2}+\cdots+\frac{1}{x_n}} \leq \sqrt[n]{x_1x_2\ldots x_n} \leq \frac{x_1+x_2+\cdots+x_n}{n} \leq \sqrt{\frac{x_1^2+x_2^2+\cdots+x_n^2}{n}} \\ \frac{n}{\sum_{i=1}^{n}\frac{1}{x_i}} \leq \sqrt[n]{\prod_{i=1}^{n} x_i} \leq \frac{\sum_{i=1}^{n} x_i}{n} \leq \sqrt{\frac{\sum_{i=1}^{n} x_i^2}{n}} \qquad (1.6) \\ \text{HM} \leq \text{GM} \leq \text{AM} \leq \text{QM} \end{gathered}$$

1.2.2 Power Laws

$$\begin{gathered} y = ax^{\alpha} \\ \alpha \equiv \text{power, index, exponent} \qquad (1.7) \\ \text{Differentiate} \Rightarrow \frac{dy}{dx} = \alpha a x^{\alpha-1} \end{gathered}$$

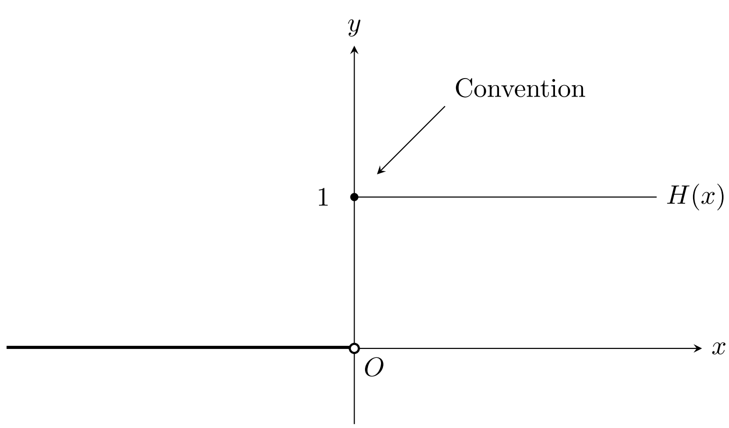


1.2.3 Trigonometric Functions



1.2.4 Heaviside (Step) Function

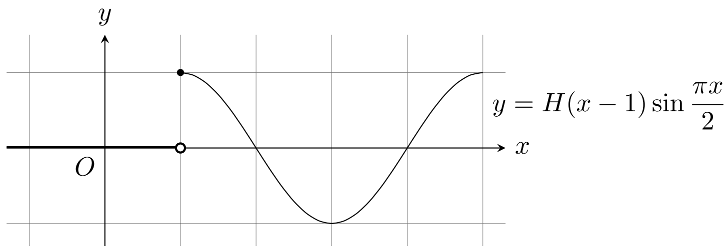
$$H(x) = \begin{cases} 0, & x < 0 \\ 1, & x \geq 0 \end{cases} \tag{1.8}$$



This function is discontinuous at $x = 0$, and

$$\frac{dH}{dx} = 0 \quad (x \neq 0).$$

This is because $H(x)$ is not differentiable at $x = 0$. The Heaviside function could be used to represent switching on or switching off a signal.



1.2.5 Modulus Function

$$|x| = \begin{cases} -x, & x < 0 \\ x, & x > 0 \end{cases} \tag{1.9}$$

The modulus function is also discontinuous at $x = 0$, where $|x| = 0$.

$$\frac{d|x|}{dx} = \begin{cases} -1, & x < 0 \\ 1, & x > 0 \end{cases}$$

1.2.6 Even and Odd Functions

$$f(x) \text{ is } \begin{cases} \text{even} & \text{if } f(x) \equiv f(-x) \text{ for all } x \text{ (reflection over the } y \text{ axis)} \\ \text{odd} & \text{if } f(x) \equiv -f(-x) \text{ for all } x \text{ (reflection through the origin)} \end{cases} \tag{1.10}$$

Notes:

(1) Not all functions are odd or even.

e.g. $f(x) = x + x^2$ is neither odd nor even.

(2) Multiplication of odd and even functions:

$f(x) \times g(x)$		$f(x)$	
		Even	Odd
$g(x)$	Even	Even	Odd
	Odd	Odd	Even

(3) Any function can be expressed as the sum of an even function and an odd function.

e.g. In (1) above, $f(x) \equiv x$ (Odd) $+ x^2$ (Even)

Therefore, a general function $g(x)$ can be written as:

$$\begin{aligned} g(x) &\equiv \frac{1}{2}[g(x) + g(-x)] + \frac{1}{2}[g(x) - g(-x)] \\ &= \text{Even Function} + \text{Odd Function} \end{aligned} \tag{1.11}$$

(4) Integration:

For an even function $f(x)$,

$$\int_{-a}^{a} f(x)\mathrm{d}x = 2\int_{0}^{a} f(x)\mathrm{d}x.$$

For an odd function $f(x)$,

$$\int_{-a}^{a} f(x)\mathrm{d}x = 0.$$

1.2.7 Inverse Functions

A function $y = f(x)$ can sometimes be inverted to get x in terms of y. This is always possible in principle, but not always in practice.

Only the "one-to-one" functions, where every x corresponds to a unique y, have inverse functions.

$$f^{-1}(y_0) = x_0 \text{ if and only if } f(x_0) = y_0 \tag{1.12}$$

The monotonous functions are a true subset of the one-to-one functions. The Monotonous functions are one-to-one, but one-to-one functions are not necessarily monotonous.

$$\text{Monotonous Functions} \subset \text{One-to-one Functions}$$

Properties of Inverse Functions

(1) The graph of $f(x)$ and that of its inverse $f^{-1}(x)$ are symmetric with respect to the x axis.

(2) $f(f^{-1}(x)) = f^{-1}(f(x)) = x$.

Finding Inverse Functions

Finding inverse functions is different from simply expressing x in terms of y.

To find the inverse function $f^{-1}(x)$ of $f(x)$, first write $f(x)$ as $y = y(x)$ (the common notation). Then, interchange x and y in the expression, and solve for y. Ultimately, we are left with the result $y = f^{-1}(x)$.

To express x in terms of y, we simply solve for x from $y = y(x)$, and what we get can be expressed as $x = g(y)$. It is worth noting that this expression stands for exactly the same curve as that of $y = f(x)$. Therefore, we cannot simply solve for x out of $y = y(x)$ to get an inverse of $f(x)$. In addition to that, we need to interchange x and y in the expression $x = g(y)$. Then, what we get is again $y = f^{-1}(x)$.

Notations

The notation $f^{-1}(x)$ should not be confused with

$$\frac{1}{f(x)} \equiv (f(x))^{-1} .$$

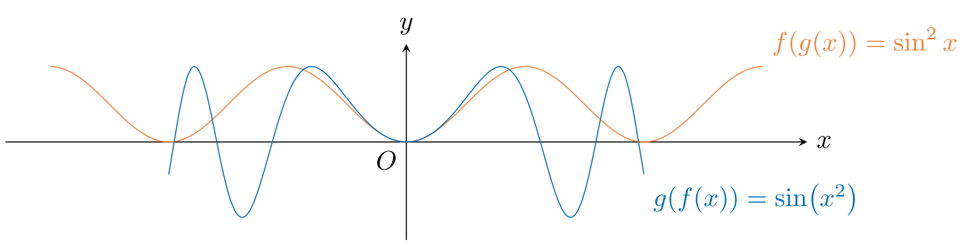
e.g. $\sin^{-1} x \equiv \arcsin x \neq \csc x$

1.2.8 Functions of a Function (Composite Functions)

The function $f(x)$ can be seen as an operator that operates on x. Given two functions (two operators) $f(x)$ and $g(x)$, we can then calculate functions of a function.

e.g. $f(x) = x^2$ and $g(x) = \sin x$:

$$f(g(x)) = (\sin x)^2 \equiv \sin^2 x \text{ and } g(f(x)) = \sin\left(x^2\right).$$



Evidently, the two functions $f(g(x))$ and $g(f(x))$ are not the same, although they are both even. We say that this composition does not commute - order matters!

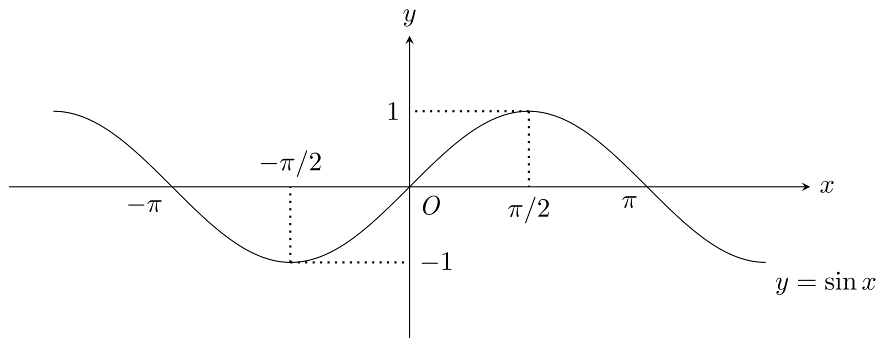
Parity of Composite Functions

The jargon "parity" talks about whether a function is odd or even. For composite functions, if the inner function $g(x)$ is even, then the composite function $f(g(x))$ is even. If the inner function $g(x)$ is odd, then the parity of the composite function $f(g(x))$ is the same as that of the outer function $f(x)$.

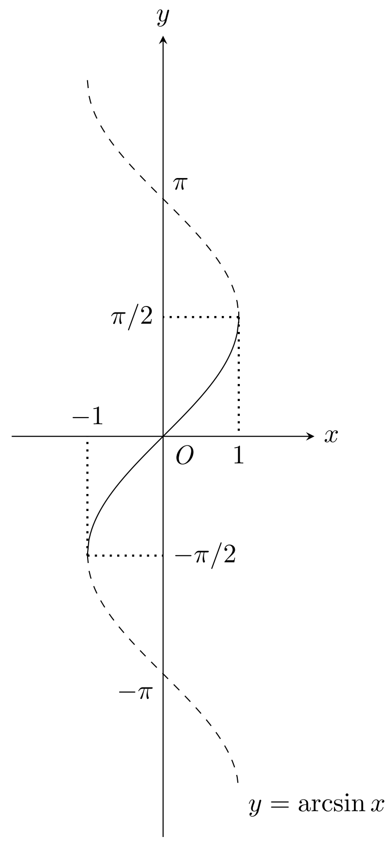
1.2.9 Many-valued Functions

Consider the function $y = \sin x$. For each x, there is only one corresponding y.

$f(g(x))$		$f(x)$	
		Even	Odd
$g(x)$	Even	Even	Even
	Odd	Even	Odd

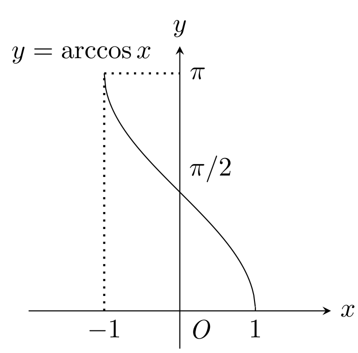


Now, consider the inverse of this function, $y = \arcsin x$. For each value of x (between -1 and 1) there are infinitely many values of y.



Hence, to clarify, we define the **principal values (P.V.)** of $\arcsin x$ to be in the range $-\dfrac{\pi}{2} \leq \arcsin x \leq \dfrac{\pi}{2}$ for $-1 \leq x \leq 1$.

Similarly, we define the P.V. for $\arccos x$ to be in the range $0 \leq \arccos x \leq \pi$ for $-1 \leq x \leq 1$.

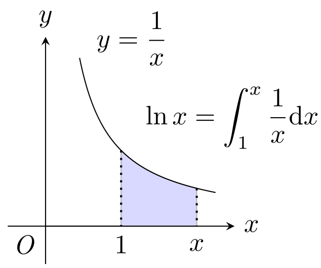


1.3 Logarithmic, Exponential, and Hyperbolic Functions

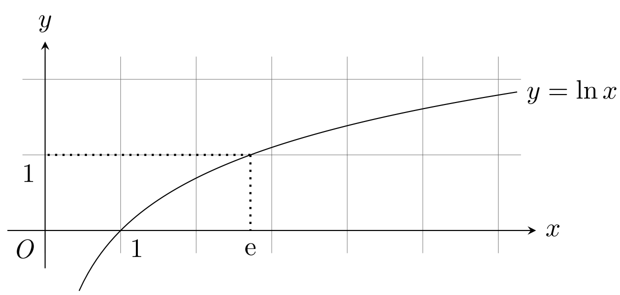
1.3.1 The Logarithm

logarithm is the exponent or power to which a base must be raised to yield a given number. Expressed mathematically, x is the logarithm of n to the base b if $b^x = n$, in which case one writes $x = \log_b n$. The natural logarithm $\ln x$, or $\log_\mathrm{e} x$, is defined as:

$$\ln x = \int_1^x \frac{\mathrm{d}t}{t}. \tag{1.13}$$



The graph of $\ln x$ is as follows.



Notes

(1) It follows that

$$\frac{\mathrm{d}}{\mathrm{d}x} \ln x = \frac{1}{x}. \tag{1.14}$$

This is the result of the famous **Fundamental Theorem of Calculus**:

$$\int_a^b f(x)\mathrm{d}x = F(b) - F(a), \text{ where } \frac{\mathrm{d}F(x)}{\mathrm{d}x} = f(x).$$

Substituting b by x, we get:

$$\int_a^x f(x)\mathrm{d}x = F(x) - F(a).$$

Taking the derivative of x at both sides gives

$$\frac{\mathrm{d}}{\mathrm{d}x}\int_a^x f(x)\mathrm{d}x = \frac{\mathrm{d}}{\mathrm{d}x}[F(x) - F(a)] = \frac{\mathrm{d}F(x)}{\mathrm{d}x} = f(x),$$

and that is:

$$\frac{\mathrm{d}}{\mathrm{d}x}\int_a^x f(x)\mathrm{d}x = f(x). \tag{1.15}$$

Furthermore, substituting x by $u = g(x)$ gives:

$$\frac{\mathrm{d}}{\mathrm{d}x}\int_a^u f(x)\mathrm{d}x = \frac{\mathrm{d}}{\mathrm{d}x}[F(u) - F(a)] = \frac{\mathrm{d}F(u)}{\mathrm{d}x} = \frac{\mathrm{d}F(u)}{\mathrm{d}u}\cdot\frac{\mathrm{d}u}{\mathrm{d}x}.$$

In general, differentiating an integral that has variable upper and lower limits gives:

$$\frac{\mathrm{d}}{\mathrm{d}x}\int_{a(x)}^{b(x)} = f(b(x))\cdot b'(x) - f(a(x))\cdot a'(x). \tag{1.16}$$

(2) Clearly, $\ln 1 = 0$.

(3) $\ln(x_1x_2) = \ln x_1 + \ln x_2$.

Proof. First, consider $\ln x_1$. From the definition, we can write:

$$\ln x_1 = \int_1^{x_1}\frac{\mathrm{d}t}{t}.$$

Let a dummy variable $s = tx_2$ with x_2 fixed. It follows that $\mathrm{d}s = x_2\mathrm{d}t$. Therefore,

$$\mathrm{d}t = \frac{1}{x_2}\mathrm{d}s \text{ and } t = \frac{s}{x_2}.$$

The above integral could be rewritten as:

$$\begin{aligned}\ln x_1 &= \int_1^{x_1}\frac{\mathrm{d}t}{t} = \int_{t=1}^{t=x_1}\frac{\mathrm{d}s/x_2}{s/x_2}\\ &= \int_{x_2}^{x_1x_2}\frac{\mathrm{d}s}{s}\\ &= \int_1^{x_1x_2}\frac{\mathrm{d}s}{s} - \int_1^{x_2}\frac{\mathrm{d}s}{s}\\ &= \ln(x_1x_2) - \ln x_2.\end{aligned}$$

Rearranging the above expression gives:

$$\ln(x_1x_2) = \ln x_1 + \ln x_2.$$

□

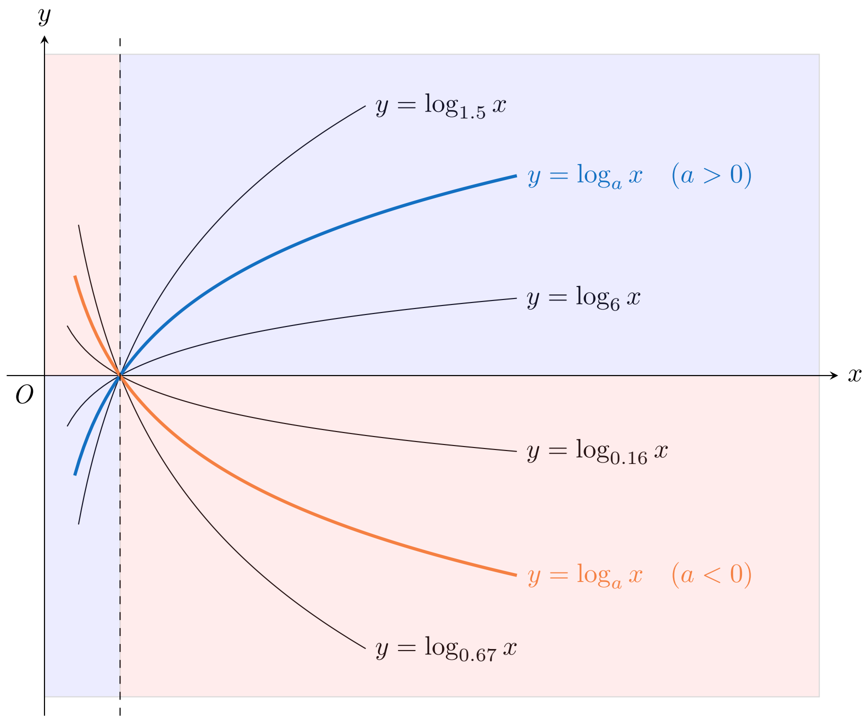
(4) Since $\ln x + \ln\left(\frac{1}{x}\right) = \ln\left(x \cdot \frac{1}{x}\right) = \ln 1 = 0$, we say

$$\ln\left(\frac{1}{x}\right) = -\ln x.$$

(5) Evidently, $\ln\left(x^2\right) = \ln x + \ln x = 2\ln x$, and

$$\ln\left(x^n\right) = n\ln x.$$

(6) Other bases of logarithms:



1.3.2 The Exponential Function

Consider $x = \ln y$. What is the inverse function, $y = f(x)$?

Let $x_1 = \ln y_1$ so that $y_1 = f(x_1)$, and $x_2 = \ln y_2$ so that $y_2 = f(x_2)$. Then,

$$x_1 + x_2 = \ln y_1 + \ln y_2 = \ln(y_1 y_2) \Rightarrow y_1 y_2 = f(x_1 + x_2).$$

This means the function $f(x)$ must satisfy

$$f(x_1 + x_2) = f(x_1) \cdot f(x_2).$$

This implies that $f(x)$ has to be of the form $f(x) = a^x$, since only this satisfies the condition $[a^{x_1} \cdot a^{x_2} = a^{x_1+x_2}]$.

Also, from the condition we can say

$$f(nx) = [f(x)]^n\,, \text{ and } f(n) = [f(1)]^n \text{ for an integer } n, \text{ and } f(0) = 1.$$

From $x = \ln y$, we can get

$$\mathrm{d}x = \frac{1}{y}\mathrm{d}y, \text{ and } \frac{\mathrm{d}x}{\mathrm{d}y} = \frac{1}{y} \Rightarrow \frac{\mathrm{d}y}{\mathrm{d}x} = y.$$

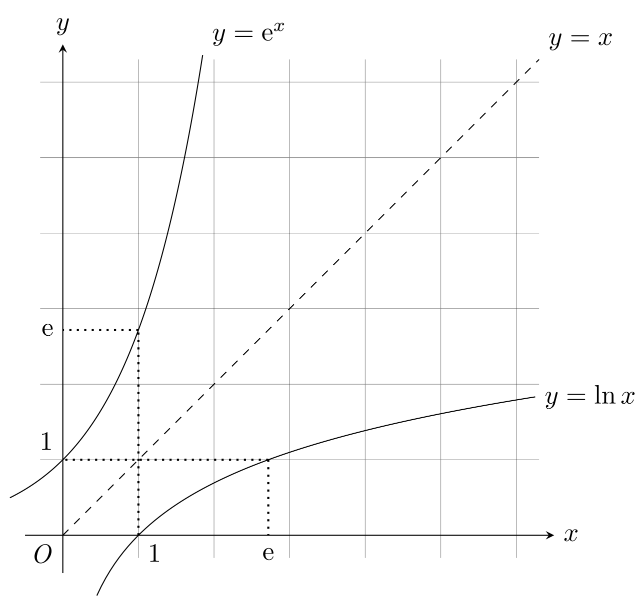
Therefore, finding a is just finding the solution of the equation

$$\frac{\mathrm{d}}{\mathrm{d}x}a^x = a^x.$$

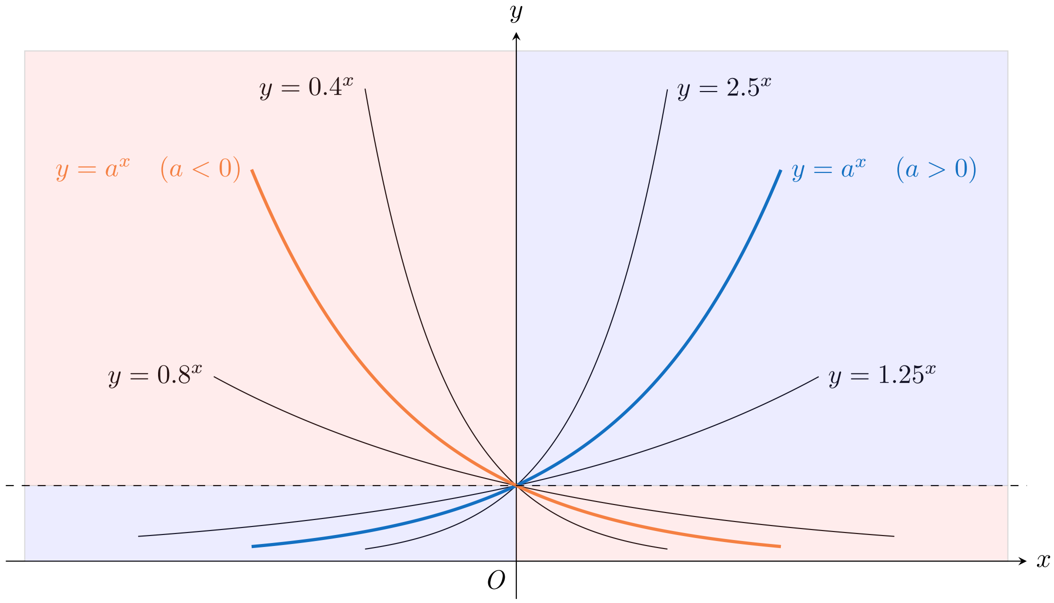
The unique number that satisfies this equation is found to be

$$\begin{aligned} \mathrm{e} &= 1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \cdots + \frac{1}{n!} + \ldots \\ &\approx 2.71828\ldots. \end{aligned}$$

It is important to note that e is an irrational number. And we finally reach the conclusion that $y = \mathrm{e}^x$ is the inverse of $y = \ln x$. This function goes to infinity faster than any power of x.



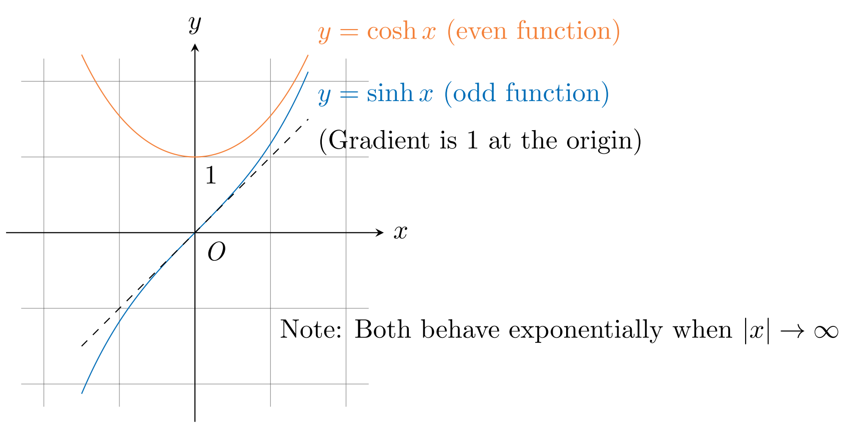
Other Bases of Exponential Functions



1.3.3 Hyperbolic Functions

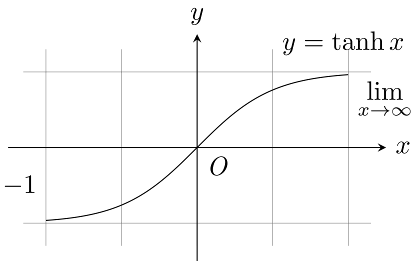
$$\sinh x \equiv \frac{1}{2}\left(\mathrm{e}^x - \mathrm{e}^{-x}\right) \tag{1.17}$$

$$\cosh x \equiv \frac{1}{2}\left(\mathrm{e}^x + \mathrm{e}^{-x}\right) \tag{1.18}$$



Clearly, from the definition and the graph of the two hyperbolic functions, we can see that $\sinh x$ is an odd function, and $\cosh x$ is an even function.

$$\tanh x \equiv \frac{\sinh x}{\cosh x} = \frac{e^x - e^{-x}}{e^x + e^{-x}} \tag{1.19}$$



$\tanh x$ is an odd function.

Notes

(1) From Trigonometric Identities to Hyperbolic Identities: **Osborn's Rule**

- Replace cos by cosh;
- Replace sin by sinh;
- However, replace any product (or implied product) of two sin terms by **minus** the product of two sinh terms.

$$\cos^2 x + \sin^2 x = 1 \Rightarrow \cosh^2 x - \sinh^2 x = 1$$

$$\sin(x_1 + x_2) = \sin x_1 \cos x_2 + \cos x_1 \sin x_2 \Rightarrow \sinh(x_1 + x_2) = \sinh x_1 \cosh x_2 + \cosh x_1 \sinh x_2$$

$$\cos(x_1 + x_2) = \cos x_1 \cos x_2 - \sin x_1 \sin x_2 \Rightarrow \cosh(x_1 + x_2) = \cosh x_1 \cosh x_2 + \sinh x_1 \sinh x_2$$

Note that there is a plus sign instead of a minus sign in the $\cosh(x_1 + x_2)$ term. For the derivatives,

$$\begin{aligned}\frac{d}{dx}\sinh x &= \cosh x\\ \frac{d}{dx}\cosh x &= \sinh x\\ \frac{d}{dx}\tanh x &= \operatorname{sech}^2 x.\end{aligned}$$

The other hyperbolic functions are defined as follows:

$$\begin{aligned}\coth x &= \frac{1}{\tanh x}\\ \operatorname{sech} x &= \frac{1}{\cosh x}\\ \operatorname{csch} x &= \frac{1}{\sinh x}.\end{aligned} \tag{1.20}$$

(2) Inverse functions are related to logarithms.

e.g. $y = \sinh^{-1} x \Rightarrow x = \sinh y$

This gives:

$$x = \frac{1}{2}\left(\mathrm{e}^y - \mathrm{e}^{-y}\right) \quad \Rightarrow \mathrm{e}^{2y} - 2x\mathrm{e}^y - 1 \equiv \left(\mathrm{e}^y\right)^2 - 2x\mathrm{e}^y - 1 = 0 \Rightarrow \mathrm{e}^y = \frac{2x \pm \sqrt{4x^2+4}}{2}.$$

So, by recalling that an real exponential is always greater than 0, we get:

$$\begin{aligned} \mathrm{e}^y &= x + \sqrt{x^2+1} \\ y &= \ln\left(x + \sqrt{x^2+1}\right). \end{aligned}$$

This means

$$\sinh^{-1} x \equiv \ln\left(x + \sqrt{x^2+1}\right). \tag{1.21}$$

Similarly,

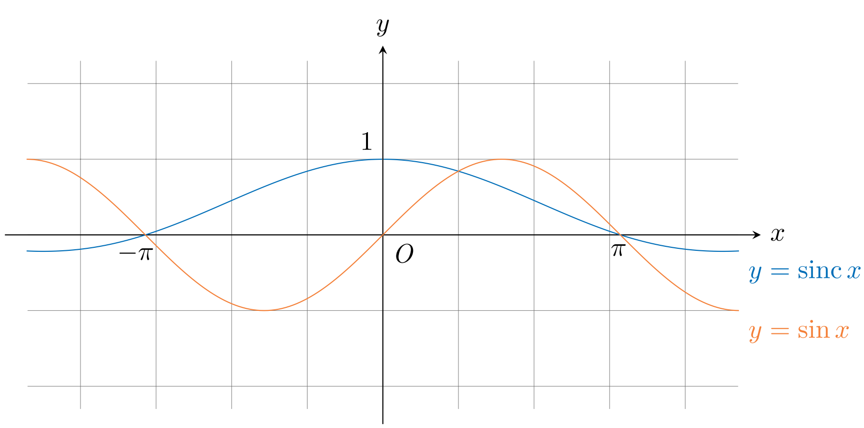
$$\cosh^{-1} x \equiv \ln\left(x + \sqrt{x^2-1}\right) \text{ for } x \geq 1 \text{ only.} \tag{1.22}$$

1.4 Limits of Functions

Consider the function

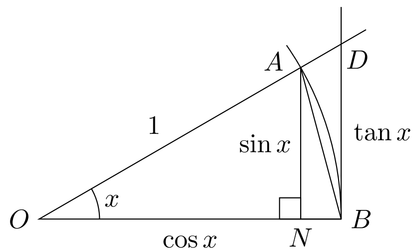
$$f(x) = \operatorname{sinc} x = \frac{\sin x}{x} \text{ for } x \neq 0. \tag{1.23}$$

$f(x)$ is not defined at $x = 0$ since it has the form $\dfrac{0}{0}$. However, plotting the graph of $f(x)$ numerically shows that $f(x)$ gets closer and closer to 1 as x gets closer to 0.



A Geometric 'Proof' of the above Observations

Consider a sector of the unit circle.



$OA = OB = 1$
$AN = \sin x$
$BD = \tan x$
$ON = \cos x$
$BN = 1 - \cos x$

Evidently, we can see that

$$A_{\Delta AOB} < A(Arc\ AOB) < A_{\Delta DOB}$$

$$\frac{1}{2}\sin x \times 1 < \frac{x}{2\pi} \times \pi(1)^2 < \frac{1}{2}\tan x \times 1.$$

Dividing through by $\frac{1}{2}\sin x$ gives:

$$1 < \frac{x}{\sin x} < \frac{1}{\cos x}.$$

Take the reciprocals of these fractions, and we are left with

$$\cos x < \frac{\sin x}{x} < 1.$$

Since $\cos x \to 1$ as $x \to 0$, we have

$$\frac{\sin x}{x} \to 1 \text{ as } x \to 0. \text{ [The Squeeze Principle!]}$$

Notations & Notes

We write

$$\frac{\sin x}{x} \to 1 \text{ as } x \to 0, \text{ or } \lim_{x\to 0} \frac{\sin x}{x} = 1.$$

It is obvious that our 'proof' was for $x > 0$ only. However, the limit process requires that $x \to 0$ through both positive and negative values of x. This geometric proof should at least be modified in order to achieve this.

Mathematically, a more formal approach is needed to turn this 'proof' into proof (via a formal limit definition).

More generally, we write the limit F of a function $f(x)$ at the point $x = x_0$ as

$$\lim_{x\to x_0} f(x) = F.$$

For limits, there is no real issue if our function is "well-behaved" at x_0.

e.g. $f(x) = x^2 + 3 \Rightarrow \lim_{x\to 4} f(x) = 19$ [Of course!]

There are some simple rules about the limits. If $\lim_{x\to x_0} f(x) = F$ and $\lim_{x\to x_0} g(x) = G$, then

$$\begin{aligned}\lim_{x\to x_0} af(x) + bg(x) &= aF + bG \\ \lim_{x\to x_0} f(x)\cdot g(x) &= FG \\ \lim_{x\to x_0} \frac{f(x)}{g(x)} &= \frac{F}{G} \text{ (provided } G \neq 0\text{),}\end{aligned} \tag{1.24}$$

where a and b are constants.

1.5 Non-trivial Limits

The concept of limits becomes important when we encounter combinations like $\frac{0}{0}$, $\frac{\infty}{\infty}$, $\infty\cdot 0$, and $\infty - \infty$. These are called indeterminate forms.

1.5.1 Type $\frac{0}{0}$

For $\frac{0}{0}$ indeterminate form we can use the L Hôpital's Rule (1696).

For $\lim_{x\to x_0} \frac{f(x)}{g(x)}$ where $\lim_{x\to x_0} f(x) = \lim_{x\to x_0} g(x) = 0$, then

$$\lim_{x\to x_0} \frac{f(x)}{g(x)} = \lim_{x\to x_0} \frac{f'(x)}{g'(x)}. \tag{1.25}$$

If both $\lim_{x\to x_0} f'(x)$ and $\lim_{x\to x_0} g'(x)$ are still zero, then differentiate again and

$$\lim_{x\to x_0} \frac{f(x)}{g(x)} = \lim_{x\to x_0} \frac{f''(x)}{g''(x)}.$$

e.g. $\lim_{x\to 0} \frac{\sin x}{x} \Rightarrow \frac{0}{0}$ type:

$$\lim_{x\to 0} \frac{\sin x}{x} = \lim_{x\to 0} \frac{\cos x}{1} = 1.$$

1.5.2 (No L Hôpital in General) Type $\frac{\infty}{\infty}$

For $\lim_{x\to x_0} \frac{f(x)}{g(x)}$ where $\lim_{x\to x_0} f(x) = \lim_{x\to x_0} g(x) = \infty$ and $f(x)$ and $g(x)$ are polynomials, we can use the method of factoring dominant powers. See the example below.

$$\lim_{x\to\infty} \frac{2x^5+2x^2-1}{x^5-x^3+1} \equiv \lim_{x\to\infty} \frac{x^5\left(2+\dfrac{2}{x^3}-\dfrac{1}{x^5}\right)}{x^5\left(1-\dfrac{1}{x^2}+\dfrac{1}{x^5}\right)} = \lim_{x\to\infty} \frac{\left(2+\dfrac{2}{x^3}-\dfrac{1}{x^5}\right)}{\left(1-\dfrac{1}{x^2}+\dfrac{1}{x^5}\right)} = 2.$$

We can see that we should expect a finite answer because the dominant powers in the numerator and the denominator are the same.

1.5.3 Type $\infty \cdot 0$

$$\begin{aligned}
\lim_{x\to\infty} \sqrt{x}\left(\sqrt{x+1}-\sqrt{x-1}\right) &= \lim_{x\to\infty} \sqrt{x}\left(\sqrt{x}\sqrt{1+\frac{1}{x}}-\sqrt{x}\sqrt{1-\frac{1}{x}}\right) \\
&= \lim_{x\to\infty} x\left(\sqrt{1+\frac{1}{x}}-\sqrt{1-\frac{1}{x}}\right) \\
&= \lim_{x\to\infty} x\left[\left(1+\frac{1}{2x}-\frac{1}{8x^2}+\dots\right)-\left(1-\frac{1}{2x}+\frac{1}{8x^2}-\dots\right)\right] \\
&= \lim_{x\to\infty}\left[1+\mathcal{O}\left(\frac{1}{x^2}\right)\right] \\
&= 1.
\end{aligned}$$

The binomial expansion is used in the process of determining this limit. According to **The Binomial Theorem**,

$$\begin{aligned}
(a+b)^n &= \binom{n}{0}a^n + \binom{n}{1}a^{n-1}b + \binom{n}{2}a^{n-2}b^2 + \dots + \binom{n}{n-1}ab^{n-1} + \binom{n}{n}b^n \\
&= \sum_{k=0}^{n}\binom{n}{k}a^k b^{n-k},
\end{aligned} \tag{1.26}$$

where $\dbinom{n}{k} = \dfrac{n!}{k!(n-k)!}$. This expansion seems to be limited to situations in which n is a positive integer, but actually it can be useful after a little modification in situations where n can be any real number. The binomial coefficient, in that case, should be written as

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} \equiv \frac{n(n-1)(n-2)\dots(n-k+1)}{k!}. \tag{1.27}$$

Generally, we can manipulate the expression we are to binomial expand into forms of $(1+x)^n$, and

$$(1+x)^n = 1 + nx + \frac{n(n-1)}{2}x^2 + \dots. \tag{1.28}$$

1.5.4 Type $\infty - \infty$

$$\begin{aligned}
\lim_{x\to\infty} x\sqrt{x^2+2} - x\sqrt{x^2-3} &= \lim_{x\to\infty} x^2\left(\sqrt{1+\frac{2}{x^2}} - \sqrt{1-\frac{3}{x^2}}\right)\\
&= \lim_{x\to\infty} x^2\left[\left(1+\frac{1}{x^2}+\dots\right) - \left(1-\frac{3}{2x^2}+\dots\right)\right]\\
&= \lim_{x\to\infty} x^2\left[\frac{5}{2x^2}+\mathcal{O}\left(\frac{1}{x^4}\right)\right]\\
&= \frac{5}{2}.
\end{aligned}$$

1.5.5 Type 1^∞

Consider the limit $\lim\limits_{x\to 0}(1-x)^{1/x}$.

Take the logarithm and then consider $\lim\limits_{x\to 0}\frac{1}{x}\ln(1-x) = \lim\limits_{x\to 0}\frac{\ln(1-x)}{x}$. This is now a $\frac{0}{0}$ indeterminate form. Using L Hôpital's Rule, we get

$$\lim_{x\to 0}\frac{\ln(1-x)}{x} = \lim_{x\to 0}\frac{-\frac{1}{1-x}}{1} = -1.$$

Therefore, we know that

$$\begin{aligned}
\lim_{x\to 0}(1-x)^{1/x} &= \lim_{x\to 0}\exp\{\ln(1-x)/x\}\\
&= \exp\Big\{\lim_{x\to 0}\ln(1-x)/x\Big\}\\
&= \mathrm{e}^{-1}.
\end{aligned}$$

In slight disguise this is

$$\lim_{n\to\infty}\left(1-\frac{1}{n}\right)^n = \mathrm{e}^{-1}.$$

Also, an alternate definition of the exponential function is:

$$\lim_{n\to\infty}\left(1+\frac{x}{n}\right)^n = \mathrm{e}^{x}. \tag{1.29}$$

1.6 Exercises

(1) Evaluate the limit

$$\lim_{x\to 0}\frac{(1-\cos x)^{3/2}}{x-\sin x}.$$

By L Hôpital's Rule,

$$\frac{\mathrm{d}}{\mathrm{d}x}(1-\cos x)^{3/2} = \frac{3}{2}\sqrt{1-\cos x}\sin x$$
$$\frac{\mathrm{d}}{\mathrm{d}x}(x-\sin x) = 1-\cos x.$$

So,

$$\lim_{x\to 0}\frac{(1-\cos x)^{3/2}}{x-\sin x} = \lim_{x\to 0}\frac{3\sin x\sqrt{1-\cos x}}{2(1-\cos x)}.$$

There are two methods to approach this limit.

First, notice that $\sin x \equiv \sqrt{1-\cos^2 x} \equiv \sqrt{(1-\cos x)(1+\cos x)}$. Then, the above limit becomes:

$$\lim_{x\to 0}\frac{3(1-\cos x)\sqrt{1+\cos x}}{2(1-\cos x)} = \lim_{x\to 0}\frac{3\sqrt{1+\cos x}}{2} = \frac{3\sqrt{2}}{2}.$$

Secondly, write the limit by dividing the common term $\sqrt{1-\cos x}$. The limit becomes:

$$\lim_{x\to 0}\frac{3\sin x}{2\sqrt{1-\cos x}}.$$

By Taylor expansion,

$$\cos x = 1 - \frac{1}{2}x^2 + \frac{1}{4!}x^4 - \dots.$$

So, $1-\cos x = \frac{1}{2}x^2 - \frac{1}{4!}x^4 + \cdots \approx \frac{1}{2}x^2$ when $x \to 0$. Therefore,

$$\lim_{x\to 0}\frac{3\sin x}{2\sqrt{1-\cos x}} = \lim_{x\to 0}\frac{3\sin x}{\sqrt{2}x} = \frac{3\sqrt{2}}{2}.$$

(2) Determine the limit

$$\lim_{x\to\infty} x^{3/2}\left[(x+1)^{1/2} + (x-1)^{1/2} - 2x^{1/2} + x^{1/2}\sin^2\frac{1}{x}\right].$$

Firstly, we can deal with the first three terms:

$$\begin{aligned}
&\lim_{x\to\infty} x^{3/2}\left[\sqrt{x+1}+\sqrt{x-1}-2\sqrt{x}\right] \\
&\qquad = \lim_{x\to\infty} x^{3/2}\left[\sqrt{x}\sqrt{1+\frac{1}{x}} + \sqrt{x}\sqrt{1-\frac{1}{x}} - 2\sqrt{x}\right] \\
&\qquad = \lim_{x\to\infty} x^{2}\left[\sqrt{1+\frac{1}{x}} + \sqrt{1-\frac{1}{x}} - 2\right].
\end{aligned}$$

By binomial expansion,

$$\sqrt{1+\frac{1}{x}} = 1 + \frac{1}{2x} - \frac{1}{8x^2} + \dots$$
$$\sqrt{1-\frac{1}{x}} = 1 - \frac{1}{2x} - \frac{1}{8x^2} - \dots.$$

Therefore,

$$\begin{aligned}
\lim_{x\to\infty} x^{3/2}\left[\sqrt{x+1}+\sqrt{x-1}-2\sqrt{x}\right] &= \lim_{x\to\infty} x^2\left[\left(1+\frac{1}{2x}-\frac{1}{8x^2}+\dots\right)+\left(1-\frac{1}{2x}-\frac{1}{8x^2}-\dots\right)-2\right] \\
&= \lim_{x\to\infty} x^2\left(-\frac{1}{4x^2}\right) \\
&= -\frac{1}{4}.
\end{aligned}$$

Now we can look at the term $\lim_{x\to\infty} x^{3/2}\left[x^{1/2}\sin^2\frac{1}{x}\right] = \lim_{x\to\infty} x^2\sin^2\frac{1}{x}$.
Using Taylor expansion,

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \dots.$$

In this case,

$$\left(\sin\frac{1}{x}\right)^2 = \left(\frac{1}{x} - \frac{1}{3!x^3} + \frac{1}{5!x^5} - \dots\right)^2 = \frac{1}{x^2} - \dots.$$

Therefore,

$$\lim_{x\to\infty} x^2\sin^2\frac{1}{x} = \lim_{x\to\infty} x^2\left(\frac{1}{x^2}-\dots\right) = 1.$$

So, the overall limit is:

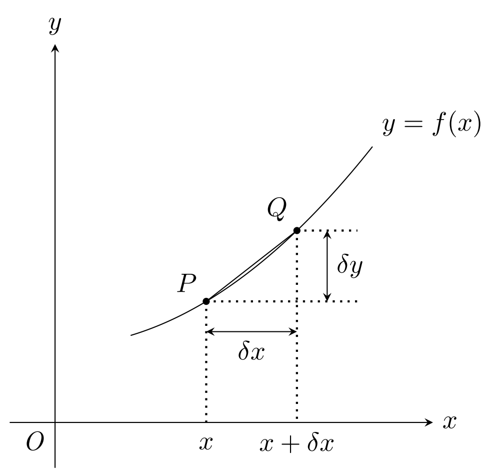
$$\lim_{x\to\infty} x^{3/2}\left[(x+1)^{1/2}+(x-1)^{1/2}-2x^{1/2}+x^{1/2}\sin^2\frac{1}{x}\right] = -\frac{1}{4}+1 = \frac{3}{4}.$$

Chapter 2

Differentiation

2.1 Differentiation from First Principles

Consider the tangent to a curve at point P as the limit of a chord PQ as Q approaches P.



$$\text{Gradient of chord } PQ = \frac{f\,(x+\delta x) - f(x)}{\delta x} \equiv \frac{\delta y}{\delta x}.$$

If P is fixed, i.e. x is fixed, let $\delta x \to 0$ and note that (as in the above) δy also goes to 0.

If the limit exists, we define

$$\lim_{\delta x \to 0} \frac{f\,(x+\delta x) - f(x)}{\delta x} = \frac{\mathrm{d}y}{\mathrm{d}x} \tag{2.1}$$

to be the derivative of $y = f(x)$ at x. We should note that this limit must be the same when we approach it in both $\delta x \to 0^+$ and $\delta x \to 0^-$ directions.

e.g. Find the derivative of $y = \sin x$.

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \lim_{\delta x \to 0} \frac{\sin(x + \delta x) - \sin x}{\delta x}$$

The first method to do this limit is through the formula $\sin(x + \delta x) \equiv \sin x \cos \delta x + \cos x \sin \delta x$. As $\delta x \to 0$, we can see that $\lim_{\delta x \to 0} \sin \delta x = \delta x$, and $\lim_{\delta x \to 0} \cos \delta x = 1$. Therefore,

$$\begin{aligned}
\frac{\mathrm{d}y}{\mathrm{d}x} &= \lim_{\delta x \to 0} \frac{\sin x \cos \delta x + \cos x \sin \delta x - \sin x}{\delta x} \\
&= \lim_{\delta x \to 0} \frac{\sin x + \delta x \cos x - \sin x}{\delta x} \\
&= \lim_{\delta x \to 0} \frac{\delta x \cos x}{\delta x} \\
&= \lim_{\delta x \to 0} \cos x \\
&= \cos x.
\end{aligned}$$

The second method to do this limit is through the formula $\sin A - \sin B = 2 \cos\left(\frac{A+B}{2}\right) \sin\left(\frac{A-B}{2}\right)$. Therefore, we know that $\sin(x + \delta x) - \sin x \equiv 2\cos\left(x + \frac{\delta x}{2}\right) \sin\left(-\frac{\delta x}{2}\right)$. Thus,

$$\begin{aligned}
\frac{\mathrm{d}y}{\mathrm{d}x} &= \lim_{\delta x \to 0} \frac{2\cos\left(x + \frac{\delta x}{2}\right) \sin\left(\frac{\delta x}{2}\right)}{\delta x} \\
&= \lim_{\delta x \to 0} \left[\cos\left(x + \frac{\delta x}{2}\right)\right] \cdot \lim_{\delta x \to 0} \left[\frac{\sin\left(\frac{\delta x}{2}\right)}{\frac{\delta x}{2}}\right] \\
&= \cos x \cdot 1 \\
&= \cos x.
\end{aligned}$$

We need to know how differentiation from first principles works, but actually we can use a stock set of derivatives in practice.

2.2 Product, Quotient, and Chain Rules

2.2.1 Product Rule

$$\frac{\mathrm{d}}{\mathrm{d}x}(fg) = f'g + g'f = \left(\frac{\mathrm{d}f}{\mathrm{d}x}\right) g + \left(\frac{\mathrm{d}g}{\mathrm{d}x}\right) f \tag{2.2}$$

Now apply this rule again to $\frac{\mathrm{d}}{\mathrm{d}x}(fg)$. What we get would be $\frac{\mathrm{d}^2}{\mathrm{d}x^2}(fg)$:

$$\frac{\mathrm{d}^2}{\mathrm{d}x^2}(fg) = \left(\frac{\mathrm{d}^2 f}{\mathrm{d}x^2}\right) g + 2\left(\frac{\mathrm{d}f}{\mathrm{d}x}\right)\left(\frac{\mathrm{d}g}{\mathrm{d}x}\right) + \left(\frac{\mathrm{d}^2 g}{\mathrm{d}x^2}\right) f.$$

This result comes in the same trend as the binomial theorem:

$$(a+b)^2 = a^2 + 2ab + b^2.$$

And actually, it is! According to **The General Leibniz Rule**, if f and g are n-times differentiable functions, then their product fg is also n-times differentiable, and its n-time derivative is given by:

$$(fg)^{(n)} = \sum_{k=0}^{n} \binom{n}{k} f^k g^{(n-k)}. \tag{2.3}$$

2.2.2 Quotient Rule

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{f}{g}\right) = \frac{f'g - g'f}{g^2} \quad \left[\equiv \frac{\mathrm{d}}{\mathrm{d}x}\left(f \cdot \frac{1}{g}\right)\right]. \tag{2.4}$$

2.2.3 Chain Rule

If $y = f(g(x))$, then

$$\frac{\mathrm{d}y}{\mathrm{d}x} = f'(g(x)) \cdot g'(x). \tag{2.5}$$

Alternatively, we can write

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}y}{\mathrm{d}g} \cdot \frac{\mathrm{d}g}{\mathrm{d}x}. \tag{2.6}$$

e.g. Find the derivative of $y = \ln \cos x$.

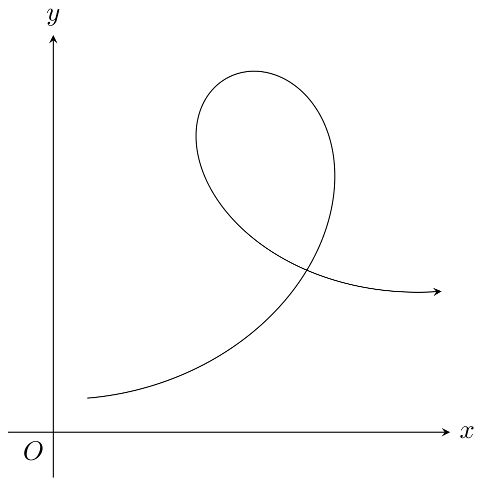
$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}y}{\mathrm{d}g} \cdot \frac{\mathrm{d}g}{\mathrm{d}x} \text{ with } y = \ln g(x), \text{ and } g(x) = \cos x.$$

Therefore,

$$\begin{aligned}\frac{\mathrm{d}y}{\mathrm{d}x} &= \frac{1}{g(x)} \cdot (-\sin x) \\ &= -\frac{\sin x}{\cos x} \\ &= -\tan x.\end{aligned}$$

2.2.4 Parametric Differentiation

Sometimes, x and y can represent the coordinates of a moving point. At this time, we write $x = x(t)$ and $y = y(t)$, and often there is not a clear relationship directly relating x to y.



In this case,

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}y/\mathrm{d}t}{\mathrm{d}x/\mathrm{d}t} = \frac{\dot{y}}{\dot{x}}. \tag{2.7}$$

Similarly,

$$\begin{aligned}\frac{\mathrm{d}^2y}{\mathrm{d}x^2} &= \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)\\ &= \frac{\mathrm{d}t}{\mathrm{d}x}\cdot\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)\\ &= \frac{1}{\mathrm{d}x/\mathrm{d}t}\cdot\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\dot{y}}{\dot{x}}\right)\\ &= \frac{1}{\dot{x}}\cdot\left(\frac{\ddot{y}\dot{x}-\ddot{x}\dot{y}}{\dot{x}^2}\right),\end{aligned}$$

which gives

$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2} = \frac{\ddot{y}\dot{x}-\ddot{x}\dot{y}}{\dot{x}^3}. \tag{2.8}$$

2.2.5 Differentiation of Inverse Functions

Take the two examples to understand differentiation of inverse functions.

(1) $y = \sin^{-1} x \Rightarrow x = \sin y$

Take the derivatives with x on both sides, and

$$1 = \cos y \cdot \frac{\mathrm{d}y}{\mathrm{d}x}.$$

This gives

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{1}{\cos y}.$$

However, as we do not know $\cos y$, we need to express it in terms of x. By recalling the trigonometric identity

$$\sin^2 y + \cos^2 y = 1,$$

we know that

$$\cos y = \sqrt{1 - \sin^2 y}.$$

Note that we do not consider $\cos y = -\sqrt{1 - \sin^2 y}$ here, because the function $y = \sin^{-1} x$ has a range of $-\frac{\pi}{2} \le y \le \frac{\pi}{2}$, in which $\cos y$ is always positive.

Therefore,

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{1}{\cos y} = \frac{1}{\sqrt{1 - \sin^2 y}} = \frac{1}{\sqrt{1 - x^2}}.$$

(2) $y = \tan^{-1} x \Rightarrow x = \tan y$

Take the derivatives of x on both sides, and

$$1 = \sec^2 y \cdot \frac{\mathrm{d}y}{\mathrm{d}x}.$$

Therefore,

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{1}{\sec^2 y}.$$

Using the trigonometric identity $\sec^2 y = \tan^2 y + 1$, we can get

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{1}{\sec^2 y} = \frac{1}{\tan^2 y + 1} = \frac{1}{x^2 + 1}.$$

We can use similar methods for other inverse functions.

2.2.6 Implicit Functions

Sometimes, the relationship between the variables x and y can be implicit, like

$$F(x, y) = 0, \tag{2.9}$$

with an explicit form $y = f(x)$ not available. However, the product rule of differentiation can still be applied here to yield a derivative.

e.g. $x^2 \sin y + xy = 1$

Take the derivative of x on both sides, and we are left with

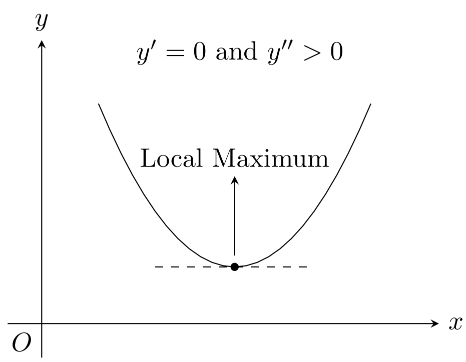
$$2x\sin y + x^2\cos y\frac{\mathrm{d}y}{\mathrm{d}x} + y + x\frac{\mathrm{d}y}{\mathrm{d}x} = 0.$$

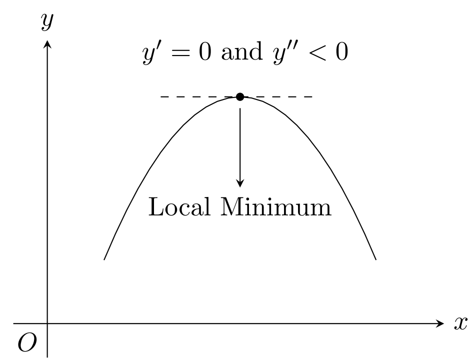
Therefore,

$$\frac{\mathrm{d}y}{\mathrm{d}x} = -\frac{y + 2x\sin y}{x^2\cos y + x}.$$

2.3 Stationary Points and Points of Inflection

A stationary point is a point on a function where $\frac{\mathrm{d}y}{\mathrm{d}x} = 0$. At a stationary point, the monotonicity of the function **may** change, which yields an local extreme value.





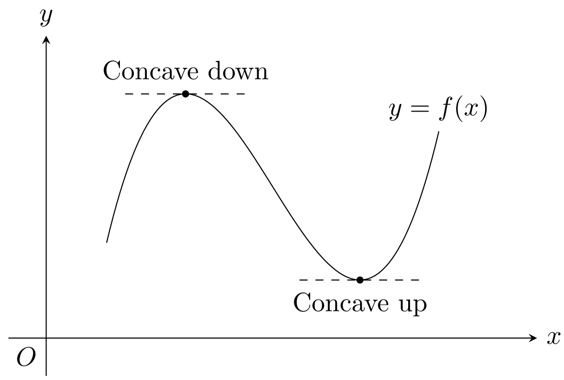
A point of inflection is a point on the graph where $\frac{\mathrm{d}^2y}{\mathrm{d}x^2} = 0$. An inflection point normally indicates the change of sign in the curvature.

$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2} = 0 \begin{cases} \frac{\mathrm{d}y}{\mathrm{d}x} = 0 & \text{Stationary Point of Inflection (inflection with horizontal tangent) or saddle point} \\ \frac{\mathrm{d}y}{\mathrm{d}x} \neq 0 & \text{Non-stationary Point of Inflection} \end{cases}$$

2.3.1 Concavity

Concavity relates to the rate of change of a function's derivative. A function $f(x)$ is concave up (or upwards) where the derivative $f'(x)$ is increasing, which means $\frac{\mathrm{d}^2y}{\mathrm{d}x^2} > 0$. On the other hand, we say the function is concave down when the derivative $f'(x)$ is decreasing, which means $\frac{\mathrm{d}^2y}{\mathrm{d}x^2} < 0$.

$$\begin{cases} y'' < 0 & \text{concave up, the curve lies above the tangent lines} \\ y'' > 0 & \text{concave down, the curve lies below the tangent lines} \end{cases}$$



2.3.2 Finding Extreme Values: Candidate Test

General guidelines to find extreme values:

- Find all potential candidates of extreme values.

$$\text{Candidate Points}\begin{cases}\text{Stationary Points} & \dfrac{\mathrm{d}y}{\mathrm{d}x}=0\\ \text{Indifferentiable Points} & \dfrac{\mathrm{d}y}{\mathrm{d}x}\ \text{DNE}\\ \text{Endpoints of the Interval} & a\le x\le b\Rightarrow x=a\ \&\ x=b\end{cases}$$

- Determine if the candidates are points of extreme values. If so, determine if they are local maxima or local minima.

$$\text{Candidate Points}\begin{cases}\text{Change in Monotonicity}\begin{cases}\text{Local Maximum} & f'(x)\ \text{changes from}+\ \text{to}-\\ \text{Local Minimum} & f'(x)\ \text{changes from}-\ \text{to}+\end{cases}\\ \text{No Change in Monotonicity: Not a Extreme Value Point}\end{cases}$$

- Based on the aforementioned process, find the global maximum and the global minimum.

2.4 Main Principles of Curve Sketching

The general guidelines of curve sketching are given below. These rules come in no particular order, and may not apply to all cases.

- Examine the behaviour of the function when $x \to 0$, $x \to \infty$, and $x \to -\infty$.

- Look for symmetries (even and odd functions).

- If $y = \dfrac{P(x)}{Q(x)}$ with polynomials $P(x)$ and $Q(x)$, then

(1) zeros of P give intersections with x axis;

(2) zeros of Q give infinite discontinuities (or vertical asymptotes).

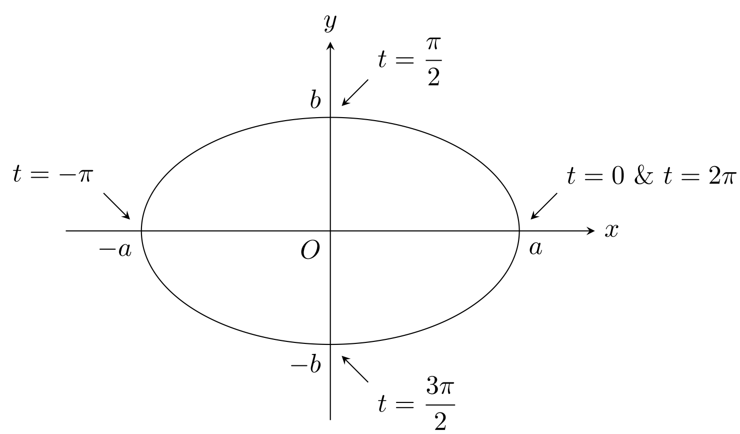
- Look for stationary points and points of inflection. They give important local details.

2.5 Parametric Representation of Curves

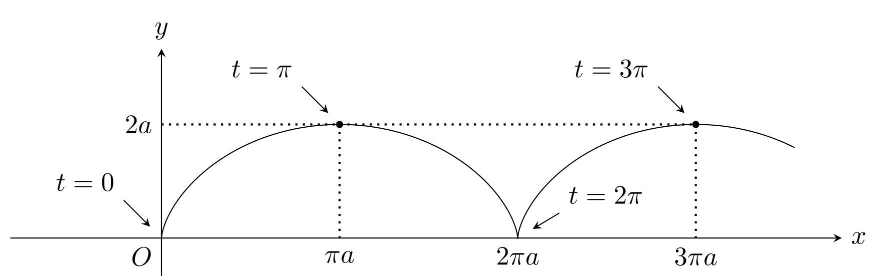
(1) Ellipse: $x = a\cos t$, $y = b\cos t$

In this case, we can eliminate t to get the function that only involves x and y:

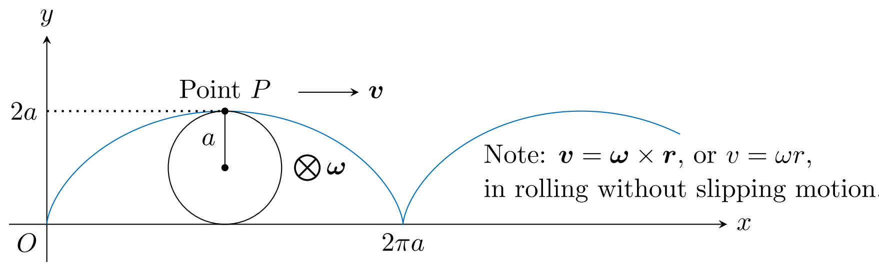
$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1.$$



(2) Cycloid: $x = a(t - \sin t)$, $y = a(1 - \cos t)$

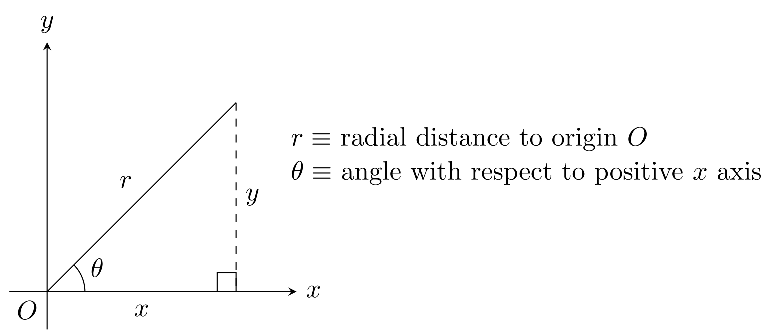


Actually, this curve represents the track of a point P on a sphere of radius a rolling without slipping.



2.6 Polar Coordinates

In two dimensions it is often useful to employ (plane) polar coordinates (r, θ) instead of Cartesian coordinates (x, y).



From Cartesian coordinates to polar coordinates, we can see that

$$\begin{aligned} r &= \sqrt{x^2 + y^2} \\ \theta &= \tan^{-1}\left(\frac{y}{x}\right). \end{aligned} \tag{2.10}$$

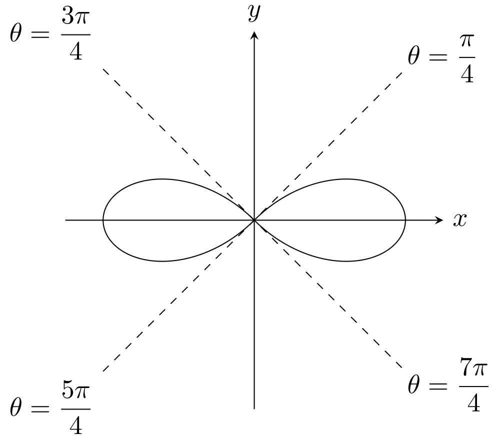
Note that it is defined that $r \geq 0$ and $0 \leq \theta < 2\pi$. Also, from polar coordinates to Cartesian coordinates, we can see that

$$\begin{aligned} x &= r\cos\theta \\ y &= r\sin\theta. \end{aligned} \tag{2.11}$$

(1) Lemniscate function: “figure of eight”

$$r^2 = 2a^2 \cos 2\theta$$

Since r is real and nonnegative, we cannot have $\cos 2\theta < 0$.



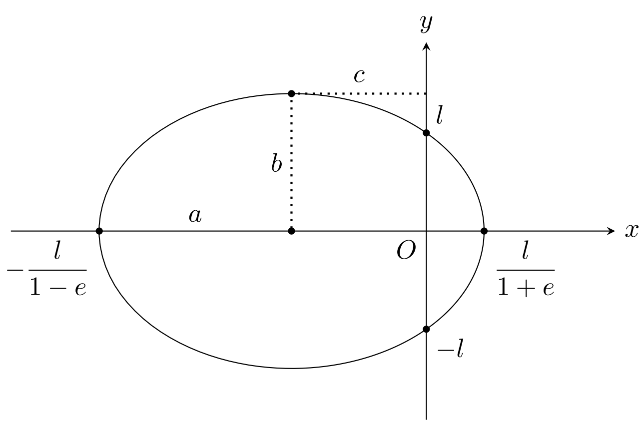
Note that the above graph is plotted only for $-\frac{\pi}{4} \le \theta \le \frac{\pi}{4}$ and $\frac{3\pi}{4} \le \theta \le \frac{5\pi}{4}$. The intervals $\frac{\pi}{4} < \theta < \frac{3\pi}{4}$ and $\frac{5\pi}{4} < \theta < \frac{7\pi}{4}$ are not plotted, and it is shown on the graph that no points are in these two regions. It is also worth noticing that the graph seems to be continuous even if its domain (θ) is not continuous.

(2) Ellipse: important in planetary orbits

$$\frac{l}{r} = 1 + e\cos\theta \tag{2.12}$$

where l is the semi-latus rectum, and e is the eccentricity ($0 < e < 1$).

When using this function to represent the graph of an ellipse, the origin O is at a focus of this ellipse.



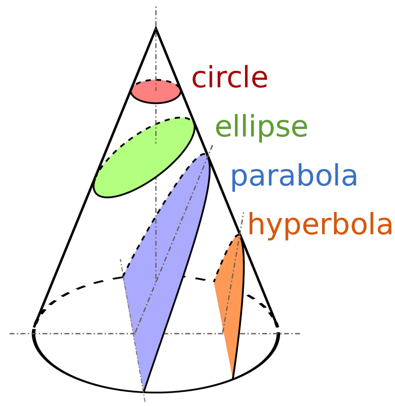
In this case,

$$a = \frac{l}{1-e^2} \qquad b = \frac{l}{\sqrt{1-e^2}}$$

$$l = \frac{b^2}{a} \qquad e = \sqrt{1-\frac{b^2}{a^2}}.$$

2.7 Conics

A conic section is a curve obtained as the intersection of the surface of a cone with a plane. The three types of conic section are the hyperbola, the parabola, and the ellipse. The circle is a special case of the ellipse (historically the fourth type).



Mathematically, the curve of a conic section could be written algebraically as

$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0. \tag{2.13}$$

Alternatively, using matrix notation, we can write:

$$\begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} A & B/2 \\ B/2 & C \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} D & E \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + F = 0. \tag{2.14}$$

More generally,

$$\begin{pmatrix} x & y & 1 \end{pmatrix} \begin{pmatrix} A & B/2 & D/2 \\ B/2 & C & E/2 \\ D/2 & E/2 & F \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = 0. \tag{2.15}$$

2.7.1 Ellipse

Definition

In the plane, the locus of a moving point P, whose sum of distances to two fixed points F_1 and F_2 ($|F_1F_2| = 2c$) is a constant ($|PF_1| + |PF_2| = 2a$), is an ellipse. The points F_1 and F_2 are called the foci of the ellipse, $|F_1F_2| = 2c$ is the focal length of the ellipse, and the constant $2a$ is the length of the major axis.

Standard Equation of an Ellipse

Set coordinates for the three points: $P(x, y)$, $F_1(-c, 0)$, and $F_2(c, 0)$. The condition is $|PF_1|+|PF_2| = 2a$.

Therefore,

$$\begin{aligned}
|PF_1| + |PF_2| &= 2a \\
\sqrt{(x+c)^2 + y^2} + \sqrt{(x-c)^2 + y^2} &= 2a \\
\sqrt{(x+c)^2 + y^2} &= 2a - \sqrt{(x-c)^2 + y^2} \\
(x+c)^2 + y^2 &= 4a^2 - 4a\sqrt{(x-c)^2 + y^2} + (x-c)^2 + y^2 \\
x^2 + 2cx + c^2 + y^2 &= 4a^2 - 4a\sqrt{(x-c)^2 + y^2} + x^2 - 2cx + c^2 + y^2 \\
4cx &= 4a^2 - 4a\sqrt{(x-c)^2 + y^2} \\
a\sqrt{(x-c)^2 + y^2} &= a^2 - cx \\
a^2\left[(x-c)^2 + y^2\right] &= a^4 - 2a^2cx - c^2x^2 \\
a^2(x^2 - 2cx + c^2 + y^2) &= a^4 - 2a^2cx - c^2x^2 \\
a^2(x^2 + c^2 + y^2) &= a^4 + c^2x + 2 \\
a^2x^2 + a^2y^2 - c^2x^2 &= a^4 - a^2c^2 \\
(a^2 - c^2)x^2 + a^2y^2 &= a^2(a^2 - c^2).
\end{aligned}$$

Let $b^2 = a^2 - c^2$, then $b^2x^2 + a^2y^2 = a^2b^2$, and

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1. \tag{2.16}$$

This describes an ellipse centered at the origin, with its foci on the x axis. For an ellipse centered at a random point (x_0, y_0), the equation is:

$$\frac{(x - x_0)^2}{a^2} + \frac{(y - y_0)^2}{b^2} = 1. \tag{2.17}$$

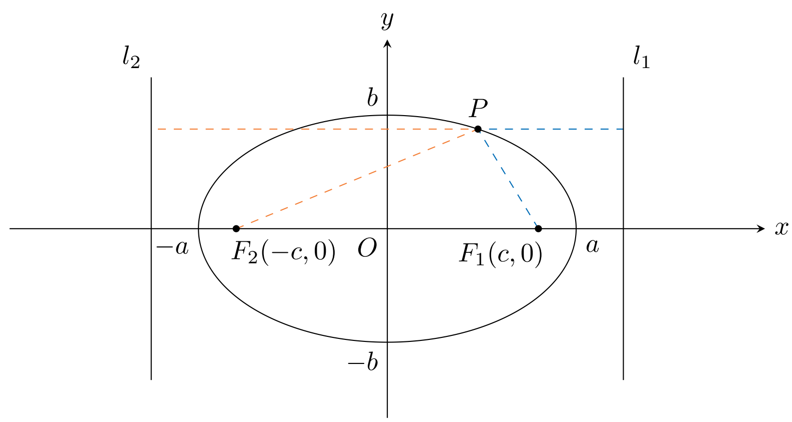
For an ellipse with its foci on the y axis, the equation is:

$$\frac{y^2}{a^2} + \frac{x^2}{b^2} = 1. \tag{2.18}$$

$2a$ here is the length of major axis, $2b$ is the length of minor axis, and $2c$ is the focal length, $a^2 = b^2 + c^2$. Clearly, we can see that a is greater than b and c, but one could not determine generally whether b or c is greater.

Directrix and Eccentricity of an Ellipse

For an arbitrary point P on a conic section, we can find a fixed line which satisfies that the quotient of the distance to one focus and to this fixed line is a constant. This constant is defined as the eccentricity e, and the line is called the directrix. For an ellipse, there are two different directrices for the two foci at different locations.



Let $P(x, y)$ be an arbitrary point on the ellipse. Let $l_1 : x = f$ be the equation for the directrix l_1. Therefore,

$$\begin{aligned}\frac{|PF_1|}{|Pl_1|} &= e \\ \frac{\sqrt{(x-c)^2+y^2}}{|f-x|} &= e \\ \sqrt{(x-c)^2+y^2} &= e|f-x| \\ (x-c)^2+y^2 &= e^2(f-x)^2\end{aligned}$$

$$\Rightarrow (1-e^2)x^2 - 2cx + c^2 + y^2 - e^2f^2 + 2e^2fx = 0.$$

Consider the points $(a, 0)$ and $(-a, 0)$ on the ellipse. By plugging them into the equation, we get

$$\begin{cases}(1-e^2)a^2 - 2ca + c^2 - e^2f^2 + 2e^2fa = 0 \\ (1-e^2)a^2 + 2ca + c^2 - e^2f^2 - 2e^2fa = 0.\end{cases}$$

Subtract one from the other, we get

$$4e^2fa - 4ca = 0.$$

As both points $(a, 0)$ and $(-a, 0)$ are on the ellipse, we know $c = e^2f \Rightarrow f = \dfrac{c}{e^2}$. Thus, looking at point

$(a, 0)$ again,

$$\begin{aligned}
(1-e^2)a^2 + c^2 - cf &= 0 \\
(1-e^2)a^2 + c^2 - \frac{c^2}{e^2} &= 0 \\
(1-e^2)a^2 &= c^2\left(\frac{1}{e^2} - 1\right) \\
(1-e^2)a^2e^2 &= c^2(1-e^2) \\
a^2e^2 &= c^2.
\end{aligned}$$

As e is the ratio of two distances and should be positive, we get that:

$$e = \frac{c}{a}. \tag{2.19}$$

Also, the equation of the directrix on the right is $x = f = \frac{c}{e^2} = \frac{a^2}{c}$:

$$l_1 : x = \frac{a^2}{c}. \tag{2.20}$$

This expression indicates that the eccentricity of an ellipse is the ratio of its half focal length (c) to its half major axis length (a). Also, in the ellipse, the eccentricity is always smaller than 1 ($0 < e < 1$).

It is easy to prove that the equation of the directrix on the left is:

$$l_2 : x = -\frac{a^2}{c}. \tag{2.21}$$

Latus Rectum of an Ellipse

The latus rectum of an ellipse is the focal chord, which is perpendicular to the major axis. The ellipse has two foci and hence it has two latus rectums. The length of semi-latus rectum is denoted as l. For an ellipse centered at the origin, consider a point of intersection.

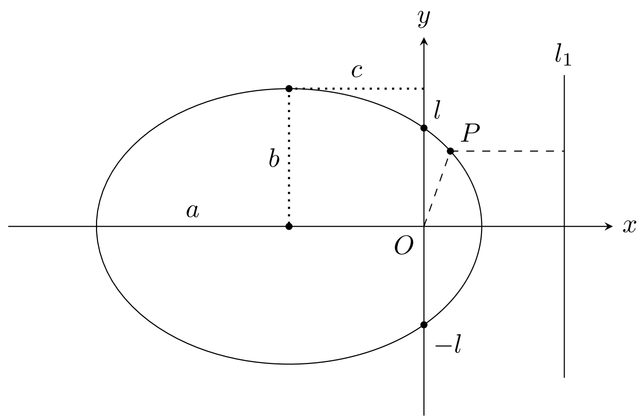
$$\begin{aligned}
\frac{c^2}{a^2} + \frac{l^2}{b^2} &= 1 \\
\frac{l^2}{b^2} &= 1 - e^2 \\
\frac{l^2}{b^2} &= \frac{b^2}{a^2} \\
a^2l^2 &= b^4
\end{aligned}$$

As l is positive, we get that

$$l = \frac{b^2}{a}. \tag{2.22}$$

Equation of an Ellipse in Polar Form

Consider the origin as one of the foci of the ellipse.



From the definition of a directrix, we can then write:

$$\frac{|OP|}{|Pl_1|} = e.$$

As the origin is at the focus, the equation of the right directrix is:

$$x = \frac{a^2}{c} - c.$$

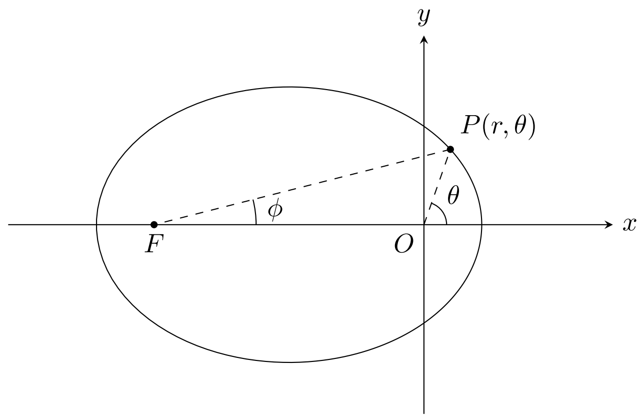
Therefore,

$$\begin{aligned}
\frac{|OP|}{|Pl_1|} &= e \\
\frac{r}{a^2/c - c - r\cos\theta} &= \frac{c}{a} \\
ar &= b^2 - cr\cos\theta \quad (a^2 - c^2 = b^2) \\
1 &= \frac{b^2}{ar} - e\cos\theta \\
1 &= \frac{l}{r} - e\cos\theta \quad \left(l = \frac{b^2}{a}\right) \\
r &= l - re\cos\theta \\
r(1 + e\cos\theta) &= l.
\end{aligned}$$

And this gives the neat polar equation of an ellipse:

$$\frac{l}{r} = 1 + e\cos\theta. \tag{2.23}$$

An alternative method of deriving the polar equation:



Consider the other focus, F, of the ellipse. The angle the line PF makes with the x axis is denoted as ϕ. By recalling the definition of the ellipse, we know that $|OP| + |PF| = 2a$, and $|PF| = 2a - |OP| = 2a - r$. Considering the x and y axes projections of OP and PF, we can write:

$$\begin{cases} x \text{ projection:} & |PF|\cos\phi = |OP|\cos\theta + 2c \\ y \text{ projection:} & |PF|\sin\phi = |OP|\sin\theta. \end{cases}$$

Therefore,

$$\begin{cases} (2a-r)\cos\phi = r\cos\theta + 2c \\ (2a-r)\sin\phi = r\sin\theta. \end{cases}$$

Square these two equations and add them together, we get:

$$\begin{aligned} (2a-r)^2 &= (r\cos\theta + 2c)^2 + r^2\sin^2\theta \\ 4a^2 - 4ar + r^2 &= 4c^2 + 4cr\cos\theta + r^2 \\ a^2 - ar &= c^2 + cr\cos\theta \\ (a + c\cos\theta)r &= b^2 \\ r &= \frac{b^2/a}{1 + c/a\cos\theta} \\ r &= \frac{l}{1 + e\cos\theta}. \end{aligned}$$

Note that we introduced the angle ϕ, but then managed to cancel it out in the process. What we are left is the polar equation of an ellipse: $l/r = 1 + e\cos\theta$.

Parametric Equation of an Ellipse

Firstly, recall the trigonometric identities:

$$\begin{aligned} \sin^2\theta + \cos^2\theta &= 1 \\ \sec^2\theta - \tan^2\theta &= 1 \\ \csc^2\theta - \cot^2\theta &= 1. \end{aligned} \tag{2.24}$$

Using Osborn's Rule,

$$\begin{aligned} \cosh^2\theta - \sinh^2\theta &= 1 \\ \operatorname{sech}^2\theta + \tanh^2\theta &= 1 \\ \coth^2\theta - \operatorname{csch}^2\theta &= 1. \end{aligned} \tag{2.25}$$

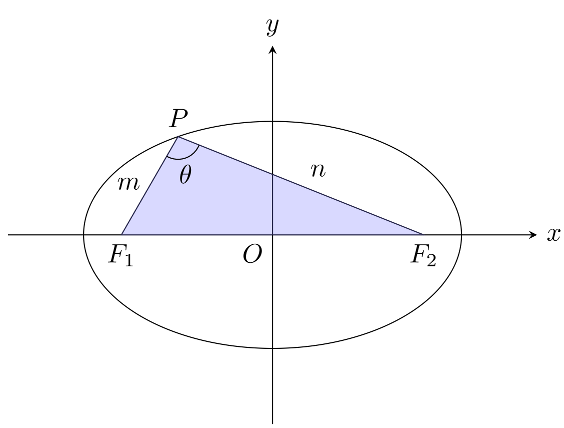
One can see the similarity between the equation $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 1$ and the identities $\cos^2 t + \sin^2 t = 1$ and $\operatorname{sech}^2\theta + \tanh^2\theta = 1$. Therefore, by setting the parameters according to the identities, we get the parametric equation of the ellipse:

$$\begin{cases} x = a\cos t \quad \text{or} \quad a\operatorname{sech} t \\ y = b\sin t \quad \text{or} \quad b\tanh t. \end{cases} \tag{2.26}$$

For an ellipse centered at (x_0, y_0), the parametric equation is:

$$\begin{cases} x = x_0 + a\cos t \quad \text{or} \quad x_0 + a\operatorname{sech} t \\ y = y_0 + b\sin t \quad \text{or} \quad y_0 + a\tanh t. \end{cases} \tag{2.27}$$

Area of The Triangle ΔPF_1F_2



Using Cosine Theorem,

$$
\begin{aligned}
|F_1F_2|^2 &= |PF_1|^2 + |PF_2|^2 - 2|PF_1||PF_2| \\
4c^2 &= m^2 + n^2 - 2mn\cos\theta \\
4c^2 &= (m+n)^2 - 2mn(1+\cos\theta) \\
4c^2 &= 4a^2 - 2mn(1+\cos\theta) \\
mn(1+\cos\theta) &= 2b^2 \\
mn &= \frac{2b^2}{1+\cos\theta}.
\end{aligned}
$$

According to Sine Theorem,

$$
A = \frac{1}{2}mn\sin\theta = \frac{1}{2}\left(\frac{2b^2}{1+\cos\theta}\right)\sin\theta = b^2\frac{\sin\theta}{1+\cos\theta}.
$$

Therefore,

$$
A = b^2\frac{\sin\theta}{1+\cos\theta} = b^2\frac{2\sin\left(\frac{\theta}{2}\right)\cos\left(\frac{\theta}{2}\right)}{1+\left[2\cos^2\left(\frac{\theta}{2}\right)-1\right]} = b^2\frac{\sin\left(\frac{\theta}{2}\right)}{\cos\left(\frac{\theta}{2}\right)},
$$

and the area of the triangle PF_1F_2 is given by

$$
A = b^2\tan\left(\frac{\theta}{2}\right). \tag{2.28}
$$

2.7.2 Hyperbola

Definition

In the plane, the locus of the moving point P, whose absolute value of the difference of the distances to two fixed points F_1 and F_2 ($|F_1F_2| = 2c$) is a constant [$||PF_1| - |PF_2|| = 2a$ ($2a < 2c$)], is a hyperbola.

Standard Equation of a hyperbola

Set the origin at the midpoint of the focal length, and we get $F_1(-c, 0)$ and $F_2(c, 0)$. $P(x, y)$ is an arbitrary point on the hyperbola. Therefore,

$$\begin{aligned}
||PF_1| - |PF_2|| &= 2a \\
|\sqrt{(x+c)^2+y^2} - \sqrt{(x-c)^2+y^2}| &= 2a \\
\sqrt{(x+c)^2+y^2} - \sqrt{(x-c)^2+y^2} &= \pm 2a \\
\sqrt{(x+c)^2+y^2} &= \pm 2a + \sqrt{(x-c)^2+y^2} \\
(x+c)^2+y^2 &= 4a^2 \pm 4a\sqrt{(x-c)^2+y^2} + (x-c)^2+y^2 \\
x^2+2cx+c^2+y^2 &= 4a^2 \pm 4a\sqrt{(x-c)^2+y^2} + x^2-2cx+c^2+y^2 \\
4cx &= 4a^2 \pm 4a\sqrt{(x-c)^2+y^2} \\
cx - a^2 &= \pm a\sqrt{(x-c)^2+y^2} \\
(cx-a^2)^2 &= a^2[(x-c)^2+y^2] \\
c^2x^2 - 2a^2cx + a^4 &= a^2x^2 - 2a^2cx + a^2c^2 + a^2y^2 \\
(c^2-a^2)x^2 - a^2y^2 &= a^2(c^2-a^2).
\end{aligned}$$

Let $b^2 = c^2 - a^2$, then $b^2x^2 - a^2y^2 = a^2b^2$, and

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1. \tag{2.29}$$

This describes a hyperbola centered at the origin, with its foci on the x axis. For a hyperbola centered at a random point (x_0, y_0), the equation is:

$$\frac{(x-x_0)^2}{a^2} - \frac{(y-y_0)^2}{b^2} = 1. \tag{2.30}$$

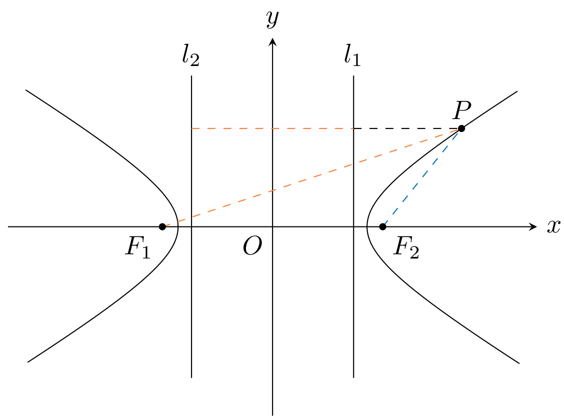
For a hyperbola with its foci on the y axis, the equation is:

$$\frac{y^2}{a^2} - \frac{x^2}{b^2} = 1. \tag{2.31}$$

$2a$ here is the length of the real axis, $2b$ is the length of imaginary axis, and $2c$ is the focal length, $c^2 = a^2 + b^2$. Clearly, we can see that c is greater than a and b, but one could not determine generally whether a or b is greater.

Directrix and Eccentricity of a Hyperbola

There are also two directrices for the two foci in a hyperbola.



Look at the focus F_2 and directrix l_1. Let the equation of the directrix l_1 be $l_1 : x = f$.

$$\begin{aligned}
\frac{|PF_2|}{|Pl_1|} &= e \\
\frac{\sqrt{(x-c)^2+y^2}}{|x-f|} &= e \\
\sqrt{(x-c)^2+y^2} &= e(x-f) \\
(x-c)^2+y^2 &= e^2(x-f)^2 \\
x^2-2cx+c^2+y^2 &= e^2x^2-2e^2fx+e^2f^2 \\
(1-e^2)x^2+2(e^2f-c)x+y^2-e^2f^2+c^2 &= 0
\end{aligned}$$

Consider the two vertices with coordinates $(-a, 0)$ and $(a, 0)$. They give

$$\begin{cases} (1-e^2)a^2+2(e^2f-c)a-e^2f^2+c^2=0 \\ (1-e^2)a^2-2(e^2f-c)a-e^2f^2+c^2=0. \end{cases}$$

Therefore, as these two points both satisfy the equation, $c = e^2 f \Rightarrow f = \dfrac{c}{e^2}$. Therefore, the equation at point $(a, 0)$ becomes

$$
\begin{aligned}
(1-e^2)a^2 - \frac{c^2}{e^2} + c^2 &= 0 \\
(1-e^2)a^2 &= c^2\left(\frac{1}{e^2} - 1\right) \\
(1-e^2)a^2e^2 &= c^2(1-e^2) \\
a^2e^2 &= c^2.
\end{aligned}
$$

As e is always positive, we get that the eccentricity for a hyperbola is the ratio of its half focal length to its semi real axis:

$$e = \frac{c}{a}. \tag{2.32}$$

We should note that, in a hyperbola, $e > 1$ instead of $0 < e < 1$, which is the case of an ellipse. Also, the equation of the directrix is $x = f = \frac{c}{e^2} = \frac{a^2}{c}$:

$$l_1 : x = \frac{a^2}{c}. \tag{2.33}$$

The equation of the directrix on the left is given by:

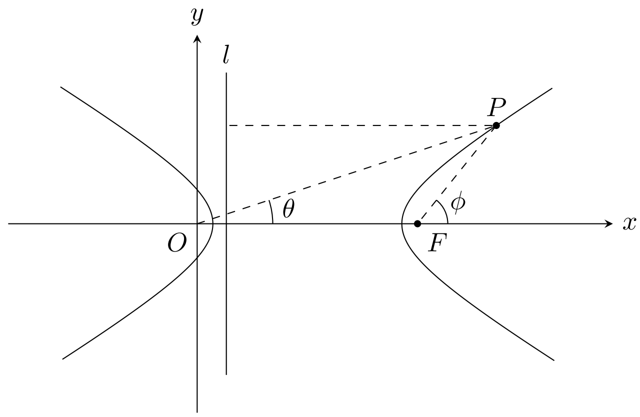
$$l_2 : x = -\frac{a^2}{c}. \tag{2.34}$$

Latus Rectum of a Hyperbola

The latus rectum is the focal chord that comes across the foci and is perpenducular to the real axis. By plugging $x = c$ into the equation $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$, we can see:

$$
\begin{aligned}
\frac{c^2}{a^2} - \frac{l^2}{b^2} &= 1 \\
b^2c^2 - a^2l^2 &= a^2b^2 \\
a^2l^2 &= b^4 \\
l &\equiv \frac{b^2}{a}.
\end{aligned}
$$

Equation of a Hyperbola in Polar Form



We can see from the graph that it seems that a single value of θ corresponds to two values of r, as it does on the initial axis. But actually, this is not the truth. The right branch of the hyperbola is actually when r is negative, which appears on the opposite side of the pole.

By the definition of the directrix,

$$\frac{|OP|}{|Pl|} = e$$

on the right branch,

$$\begin{aligned} \frac{-r}{r\cos\theta - c + a^2/c} &= \frac{c}{a} \\ -ar &= cr\cos\theta + a^2 - c^2 \\ -r(a + c\cos\theta) &= -b^2 \end{aligned}$$

on the left branch,

$$\begin{aligned} \frac{r}{c - a^2/c - r\cos\theta} &= \frac{c}{a} \\ ar &= c^2 - a^2 - cr\cos\theta \\ r(a + c\cos\theta) &= b^2 \end{aligned}$$

$$\begin{aligned} r(1 + e\cos\theta) &= \frac{b^2}{a} \\ \frac{l}{r} &= 1 + e\cos\theta. \end{aligned}$$

Use the projection method again. For the right branch,

$$\begin{cases} x \text{ projection:} & |PF|\cos\phi = |OP|\cos\theta - 2c \\ y \text{ projection:} & |PF|\sin\phi = |OP|\sin\theta. \end{cases}$$

Square the two equations and add them together, we get:

$$\begin{aligned} |PF|^2 &= (r\cos\theta - 2c)^2 + r^2\sin^2\theta \\ &= 4c^2 - 4cr\cos\theta + r^2. \end{aligned}$$

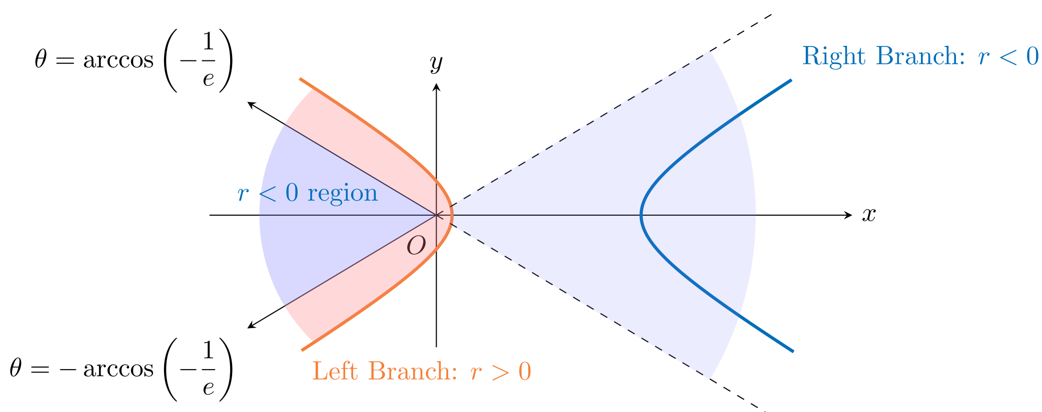
Note that $|OP| = r$, $|PF| = -r-2a$ on the right branch, and $|PF| = 2a+r$ on the left branch. Obviously, the square of $|PF|$ is the same on both branches. Therefore,

$$\begin{aligned} r^2 + 4ar + 4a^2 &= 4c^2 - 4cr\cos\theta + r^2 \\ r(c\cos\theta + a) &= b^2 \\ r &= \frac{b^2}{c\cos\theta + a} \\ \frac{l}{r} &= 1 + e\cos\theta. \end{aligned}$$

Thus, the polar equation of the hyperbola is also

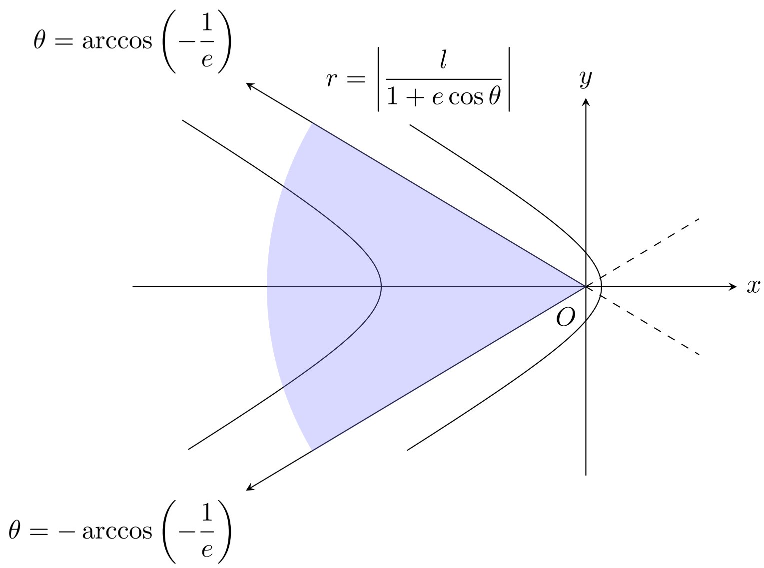
$$\frac{l}{r} = 1 + e\cos\theta, \tag{2.35}$$

but in this case $e > 1$. This means a negative r for some θ, which makes the locus of points appearing on the opposite side of the pole.



If we add an absolute value sign to this function, we may see this plot in a clearer way.

$$r = \left|\frac{l}{1 + e\cos\theta}\right|$$



Parametric Equation of a Hyperbola

We can find the similarity between the equation $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$ and the identities $\sec^2\theta - \tan^2\theta = 1$, $\csc^2\theta - \cot^2\theta = 1$, $\cosh^2\theta - \sinh^2\theta = 1$, and $\coth^2\theta - \operatorname{csch}^2\theta = 1$. These four identities provide four different ways to describe a hyperbolic function in parameters. The commonly used parameters are $\sec t$ with $\tan t$, and $\cosh t$ with $\sinh t$. However, one should note that the positive term should be assigned as $\sec t$ or $\cosh t$ and never $\tan t$ or $\sinh t$, as the negative sign plays an important role here.

The parametric equation of a hyperbola with foci on the x axis is given by:

$$\begin{cases} x = a\sec t \quad \text{or} \quad a\cosh t \\ y = b\tan t \quad \text{or} \quad b\sinh t. \end{cases} \tag{2.36}$$

For a hyperbola centered at (x_0, y_0), the parametric equation is:

$$\begin{cases} x = x_0 + a\sec t \quad \text{or} \quad x_0 + a\cosh t \\ y = y_0 + b\tan t \quad \text{or} \quad y_0 + a\sinh t. \end{cases} \tag{2.37}$$

Asymptotes of Hyperbola

Every hyperbola has two asymptotes. Before diving into the asymptotes of the hyperbola, we can have a review of finding an asymptote.

(1) Horizontal Asymptotes:

We approach a horizontal asymptote by the curve of a function as $x \to \infty$ and $-\infty$.

If a function has horizontal asymptotes, they are given by:

$$\begin{aligned} y &= \lim_{x\to\infty} f(x) \\ y &= \lim_{x\to-\infty} f(x). \end{aligned} \tag{2.38}$$

(2) Vertical Asymptotes:

We approach a vertical asymptote by looking at infinite discontinuities. At an infinite discontinuity point with x coordinate c,

$$\begin{cases} \lim_{x\to c^-} f(x) = \infty \text{ or } -\infty \\ \lim_{x\to c^+} f(x) = \infty \text{ or } -\infty. \end{cases} \tag{2.39}$$

Other discontinuities include jump discontinuities (left and right limits are finite but not equal) and removable discontinuities (left and right limites are equal, but the function is not defined there). A vertical asymptote has the equation

$$x = c, \tag{2.40}$$

where c is a point of infinite discontinuity.

(3) Slant Asymptotes:

We can find a slant asymptote via two steps. A slant asymptote is a straight line with an equation $y = ax + b$. Therefore, the two steps are taken to separately get the slope a and intercept b. To get the slope a,

$$a = \lim_{x\to\infty} \frac{f(x)}{x}. \tag{2.41}$$

To get the intercept b,

$$b = \lim_{x\to\infty} \left(f(x) - ax\right). \tag{2.42}$$

Now we can begin to look at the asymptotes of a hyperbola. For the positive right branch, we can get an explicit function:

$$\begin{aligned} \frac{x^2}{a^2} - \frac{y^2}{b^2} &= 1 \ (x > 0) \\ a^2y^2 &= b^2x^2 - a^2b^2 \\ y &= \sqrt{\frac{b^2}{a^2}x^2 - b^2} \ (x, y > 0). \end{aligned}$$

Therefore, the slope k is

$$\begin{aligned} k &= \lim_{x\to\infty} \frac{f(x)}{x} \\ &= \lim_{x\to\infty} \frac{\sqrt{\frac{b^2}{a^2}x^2 - b^2}}{x} \\ &= \lim_{x\to\infty} \sqrt{\frac{b^2}{a^2} - \frac{b^2}{x^2}} \\ &= \frac{b}{a}. \end{aligned}$$

The intercept m is

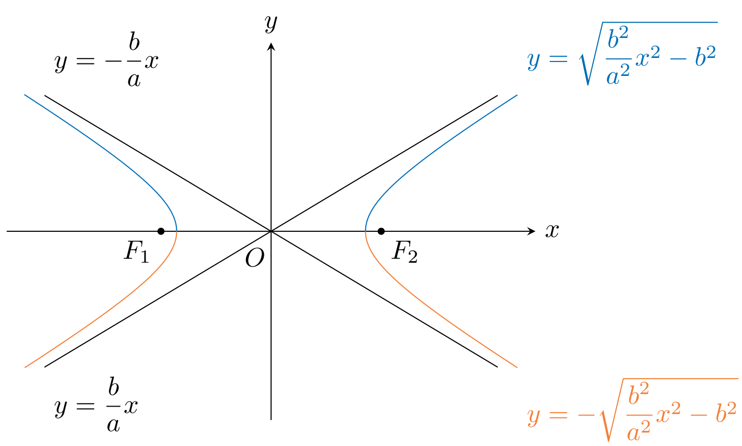
$$\begin{aligned} m &= \lim_{x\to\infty} (f(x) - kx) \\ &= \lim_{x\to\infty} \left(\sqrt{\frac{b^2}{a^2}x^2 - b^2} - \frac{b}{a}x \right) \\ &= 0. \end{aligned}$$

Similarly, we can find the explicit functions for other three sectors of hyperbola:

$$\begin{cases} y = \sqrt{\frac{b^2}{a^2}x^2 - b^2} > 0 & \text{(positive branches)} \\ y = -\sqrt{\frac{b^2}{a^2}x^2 - b^2} < 0 & \text{(negative branches)} \end{cases}$$

By repeating the process of finding the slant asymptotes, it is easy to show that a hyperbola with foci at the x axis has two slant asymptotes with equations

$$y = \pm\frac{b}{a}x. \tag{2.43}$$



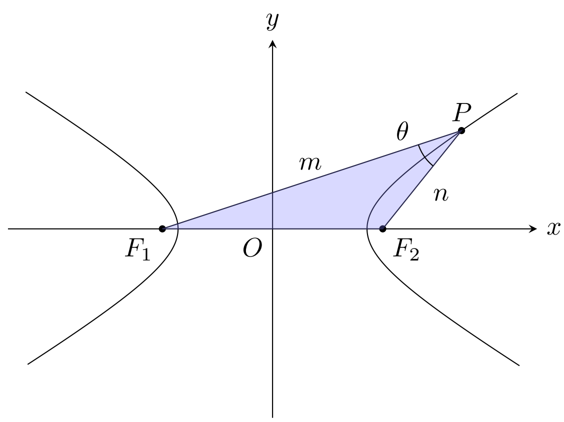
For a hyperbola with foci on the y axis $\left(\frac{y^2}{a^2} - \frac{x^2}{b^2} = 1\right)$, it is not hard to prove that the equation of its asymptotes is:

$$y = \pm \frac{a}{b} x. \tag{2.44}$$

For a hyperbola centered at a random point (x_0, y_0) with real axis parallel to the x axis, the equation of its asymptotes is:

$$y - y_0 = \pm \frac{b}{a}(x - x_0). \tag{2.45}$$

Area of The Triangle ΔPF_1F_2



Using Cosine Theorem,

$$\begin{aligned}
|F_1F_2|^2 &= |PF_1|^2 + |PF_2|^2 - 2|PF_1||PF_2|\cos\theta \\
4c^2 &= m^2 + n^2 - 2mn\cos\theta \\
4c^2 &= (m-n)^2 - 2mn(\cos\theta - 1) \\
4c^2 &= 4a^2 - 2mn(\cos\theta - 1) \\
mn &= \frac{2b^2}{1-\cos\theta}.
\end{aligned}$$

Using Sine Theorem,

$$\begin{aligned}
A &= \frac{1}{2}|PF_1||PF_2|\sin\theta \\
&= \frac{1}{2} mn \sin\theta \\
&= \frac{1}{2}\left(\frac{2b^2}{1-\cos\theta}\right)\sin\theta \\
&= b^2 \frac{\sin\theta}{1-\cos\theta}.
\end{aligned}$$

As we know

$$\frac{\sin\theta}{1-\cos\theta} = \frac{2\sin\left(\frac{\theta}{2}\right)\cos\left(\frac{\theta}{2}\right)}{1-\left(1-2\sin^2\left(\frac{\theta}{2}\right)\right)} = \frac{2\sin\left(\frac{\theta}{2}\right)\cos\left(\frac{\theta}{2}\right)}{2\sin^2\left(\frac{\theta}{2}\right)} = \frac{\cos\left(\frac{\theta}{2}\right)}{\sin\left(\frac{\theta}{2}\right)} = \frac{1}{\tan\left(\frac{\theta}{2}\right)},$$

the area of the triangle PF_1F_2 is given by:

$$A = \frac{b^2}{\tan\frac{\theta}{2}} = b^2\cot\frac{\theta}{2}. \tag{2.46}$$

2.7.3 Parabola

Definition

The locus of a moving point P that is equidistant from a fixed point F and a straight line l in the plane is a parabola. The fixed point F is called the focus, and the straight line l is called the directrix.

Standard Equation of a Parabola

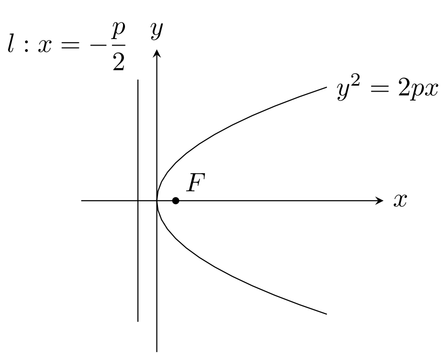
A random point $P(x, y)$ on the parabola is equidistant from the focus $F\left(\frac{p}{2}, 0\right)$ and the directrix $l : x = -\frac{p}{2}$.

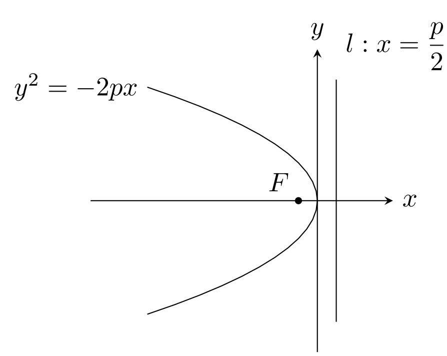
$$
\begin{aligned}
\frac{|PF|}{|Pl|} &= 1 \\
\frac{\sqrt{\left(x - \frac{p}{2}\right)^2 + y^2}}{\left|x + \frac{p}{2}\right|} &= 1 \\
\sqrt{\left(x - \frac{p}{2}\right)^2 + y^2} &= x + \frac{p}{2} \\
\left(x - \frac{p}{2}\right)^2 + y^2 &= \left(x + \frac{p}{2}\right)^2 \\
x^2 - px + \frac{p^2}{4} + y^2 &= x^2 + px + \frac{p^2}{4} \\
y^2 &= 2px.
\end{aligned}
$$

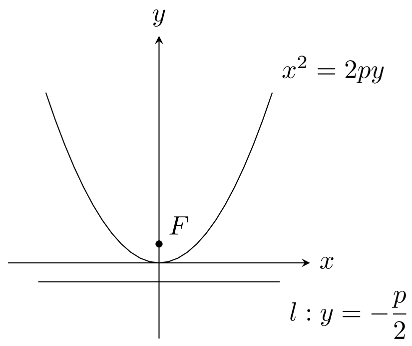
Four Types of Parabolas

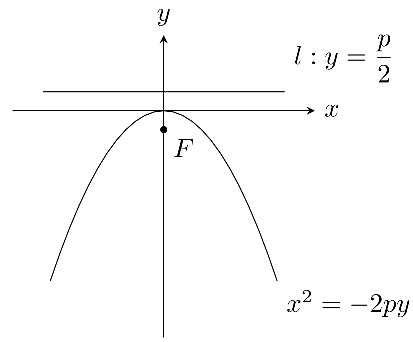
There are a total of four types of parabolas, each opening to a different direction.

$$
\begin{cases}
\text{to the right:} & y^2 = 2px \\
\text{to the left:} & y^2 = -2px \\
\text{upwards:} & x^2 = 2py \\
\text{downwards:} & x^2 = -2py
\end{cases}
\tag{2.47}
$$









For a parabola not centered at the origin, think by translating the vertex of the parabola back to origin. For example, for a parabola with a vertex at (x_0, y_0) opening to the right, we can write $(y-y_0)^2 = 2p(x-x_0)$.

Eccentricity and Latus Rectum

It is very clear that, from the definition, a parabola has an eccentricity of $e = 1$. And as it only has one focus, there is only one directrix.

The latus rectum is the focal chord that comes across the focus and is parallel to the directrix. By plugging $x = \frac{p}{2}$ into the equation $y^2 = 2px$, we can see:

$$l^2 = p^2 \Rightarrow l \equiv p.$$

Equation of a Parabola in Polar Form

Consider the parabola with its focus at the origin. By the definition of the parabola, for one opening to the right,

$$\begin{aligned}
\frac{|OP|}{|Pl|} &= 1 \\
\frac{r}{r\cos\theta + p} &= 1 \\
r &= r\cos\theta + p \\
r(1-\cos\theta) &= p \\
r &= \frac{p}{1-\cos\theta}.
\end{aligned}$$

By recalling that $l \equiv p$, we can write $r = \dfrac{l}{1-\cos\theta}$.

For one opening to the left, consider rotating the above equation by π radians.

$$r = \frac{l}{1-\cos(\theta-\pi)} = \frac{l}{1+\cos\theta}.$$

For one opening upwards, consider a rotation by $\frac{\pi}{2}$ radians.

$$r = \frac{l}{1-\cos(\theta - \pi/2)} = \frac{l}{1-\sin\theta}.$$

For one opening downwards, consider a rotation by $-\frac{\pi}{2}$ radians.

$$r = \frac{l}{1-\cos(\theta + \pi/2)} = \frac{l}{1+\sin\theta}.$$

In conclusion, the equation of a parabola in polar form can be generally expressed as:

$$\frac{l}{r} = 1 + \cos\theta.$$

Parametric Equation of a Parabola

We want the parametric equation to be as simple as possible, but we can see that there is a second order term in any form of parabola, like the y^2 term in $y^2 = 2px$. To simplify the expression, we try to make the RHS, i.e. $2px$, the square of a parametric function.

Clearly, it turns out that $x = 2pt^2$ would be a good choice, as this gives $y = \sqrt{4p^2t^2} = 2pt$. In conclusion, the parametric equation for a parabola opening to the right is:

$$\begin{cases} x = 2pt^2 \\ y = 2pt. \end{cases} \tag{2.48}$$

2.7.4 Summary of Conics

Let F be a fixed point, l be a fixed line, e be a positive constant, and the locus of the point P satisfying $\frac{|PF|}{|Pl|} = e$ is a conic section. F is called the focus, l is called the directrix, and e is called the eccentricity.

	Standard Equation	Eccentricity (e)	Semi-focal Length (c)	Semi-latus Rectum (l)
Circle	$x^2 + y^2 = a^2$	0	0	a
Ellipse	$x^2/a^2 + y^2/b^2 = 1$	$0 < c/a < 1$	$a^2 = b^2 + c^2$	b^2/a
Parabola	$y^2 = 2px$	1	N.A.	p
Hyperbola	$x^2/a^2 - y^2/b^2 = 1$	$c/a > 1$	$c^2 = a^2 + b^2$	b^2/a

The type of conics can be specified with the eccentricity e.

$$\begin{cases} e = 0 & \text{Circle} \\ 0 < e < 1 & \text{Ellipse} \\ e = 1 & \text{Parabola} \\ e > 1 & \text{Hyperbola} \end{cases} \tag{2.49}$$

The polar equation of a conic section can be written as

$$\frac{l}{r} = 1 + e\cos\theta, \tag{2.50}$$

where e is the eccentricity.

It is important to note that sometimes we can see $1 \pm e\cos\theta$, $1 \pm e\sin\theta$, $e\cos\theta \pm 1$, and $e\sin\theta \pm 1$. These equations are also correct, as they just represent a rotation by a multiple of $\frac{\pi}{2}$ radians, and sometimes a reflection with respect of the pole.

2.8 Polynomial and Series Representation of Functions

We can consider local approximations to a smooth function in the neighborhood of a general point $(x_0, f(x_0))$ as a sequence of polynomials.

$$\begin{cases} f(x) \approx f(x_0) \\ f(x) \approx f(x_0) + f'(x_0)(x - x_0) \\ f(x) \approx f(x_0) + f'(x_0)(x - x_0) + \frac{1}{2}f''(x_0)(x - x_0)^2. \end{cases} \tag{2.51}$$

All of these polynomial representations are **local**, with a trade-off to be expected between accuracy over a domain $(x_0 - \epsilon, x_0 + \epsilon)$ and the degree of the polynomial.

If $f(x)$ has successive derivatives at x_0 we can continue to derive a **Taylor Series** expansion at x_0 in the form

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{1}{2}f''(x_0)(x - x_0)^2 + \cdots + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n + \dots. \tag{2.52}$$

For the moment we consider successive truncations to provide approximations to the local behavior near $(x_0, f(x_0))$, with successive derivatives providing more and more information.

2.9 Notes on AM-GM Inequality

We can find out that if we Taylor expand e^x at $x=0$, we get:

$$\mathrm{e}^x = 1 + x + \frac{x^2}{2} + \frac{x^3}{3!} + \cdots + \frac{x^n}{n!} + \ldots. \tag{2.53}$$

So, we can see that $\mathrm{e}^x \geq 1 + x$ with equality only at $x = 0$.

Now, look at the AM-GM inequality. For a set of n numbers, a_1, a_2, $\ldots$, and a_n, we have:

$$\begin{cases} \text{Arithmetic Mean:} \quad \mathrm{AM} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} a_i \\ \text{Geometric Mean:} \quad \mathrm{GM} = \sqrt[n]{a_1 a_2 \ldots a_n}. \end{cases}$$

For each a_i we can write:

$$\exp\left\{\left(\frac{a_i}{\mathrm{AM}} - 1\right)\right\} \geq 1 + \left(\frac{a_i}{\mathrm{AM}} - 1\right) = \frac{a_i}{\mathrm{AM}}.$$

By multiplying this expression with all a_i, we can see that:

$$\exp\left\{\left(\frac{1}{\mathrm{AM}}\sum_{i=1}^{n} a_i - n\right)\right\} \geq \frac{(a_1 a_2 \ldots a_n)}{\mathrm{AM}^n} \equiv \left(\frac{\mathrm{GM}}{\mathrm{AM}}\right)^n.$$

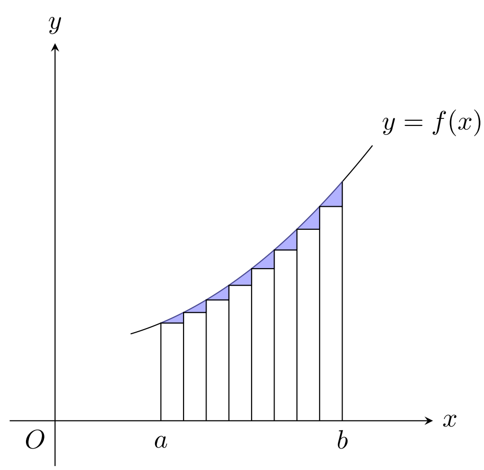
And as $\exp\left\{\left(\frac{1}{\mathrm{AM}}\sum_{i=1}^{n} a_i - n\right)\right\} \equiv 1$, we see that $\mathrm{AM} \geq \mathrm{GM}$, with equality only when $a_1 = a_2 = \cdots = a_n$.

Chapter 3

Integration

3.1 Riemann's Definition

Integration arose from intuitive ideas about area and volume in geometry. Consider the area under a curve $f(x)$ between coordinates $x = a$ and $x = b$.



To Calculate the area, imagine a large number of rectangular strips located at $x_0, x_1, \ldots, x_n$ with $x_0 = a$ and $x_n = b$. Therefore,

$$\text{Area of Strips} = S_n = \sum_{i=0}^{n-1} f(x_i)\delta x_i,$$

where $f(x_i)$ represent the height of the strips, $\delta x_i = x_{i+1} - x_i$ is the strip width.

Intuition leads us to expect that $S_n \to A$ as the number of strips $n \to \infty$. We expect the that the error (shaded pieces) vanishes in the limit.

Riemann generalised the above to show that the intuition is correct:

(1) He used upper and lower sums (i.e. 'external' and 'internal' rectangles) which has the same limit as $n \to \infty$.

(2) He used the height of the strip $f(\xi_i)$, where ξ_i is any point on $x_i \leq \xi_i \leq x_{i+1}$.

So, Riemann's definition is the limit of the sum over strips:

$$S_n^* = \sum_{i=0}^{n-1} f(\xi_i)\delta x_i \quad \text{with } \xi_i \text{ as above.}$$

He showed that

$$S_n^* \to S_n \to A \text{ as } n \to \infty, \text{ and}$$

$$\int_a^b f(x)\mathrm{d}x = \lim_{\substack{n\to\infty \\ \delta x_i \to 0}} \sum_{i=0}^{n-1} f(\xi_i)\delta x_i \tag{3.1}$$

is the **integral** (the definite integral) of f between $x = a$ and $x = b$.

Notes:

(1) $\begin{cases} f(x) \equiv \text{integrand} \\ a \equiv \text{lower limit of integration} \\ b \equiv \text{upper limit of integration} \end{cases}$

x here is a dummy variable (any symbol will do the same).

$$\int_a^b f(x)\mathrm{d}x = \int_a^b f(s)\mathrm{d}s = \int_a^b f(t)\mathrm{d}t$$

(2) We also define

$$\int_a^b f(x)\mathrm{d}x = -\int_b^a f(x)\mathrm{d}x, \tag{3.2}$$

and also $\displaystyle\int_a^a f(x)\mathrm{d}x = 0$.

(3) $\displaystyle\int_a^c f(x)\mathrm{d}x = \int_a^b f(x)\mathrm{d}x + \int_b^c f(x)\mathrm{d}x$.

(4) We will not cover **Fundamental Theorem of Calculus** here this time, as we have done that in Chapter 1.

3.2 Infinite and Improper Integrals

Infinite integrals have a $+\infty$ or $-\infty$ in its limit. To decide if an infinite integral $\displaystyle\int_a^{+\infty} f(x)\mathrm{d}x$ is meaningful, we write

$$I(N) = \int_a^N f(x)\mathrm{d}x.$$

If this has a finite limit as $N \to \infty$, then the infinite integral exists.

$$\begin{aligned}\int_a^\infty \mathrm{e}^{-x}\,\mathrm{d}x &= \lim_{N\to\infty}\int_a^N \mathrm{e}^{-x}\,\mathrm{d}x\\ &= \lim_{N\to\infty}(\mathrm{e}^{-a}-\mathrm{e}^{-N})\\ &= \mathrm{e}^{-a}\,.\end{aligned}$$

$$\begin{aligned}\int_a^\infty \frac{\mathrm{d}x}{x} &= \lim_{N\to\infty}\int_a^N \frac{\mathrm{d}x}{x}\\ &= \lim_{N\to\infty}(\ln N - \ln a) \quad \Rightarrow \text{DNE.}\end{aligned}$$

In a similar fashion, improper integrals involve a **singularity** of the integrand on the range of integration. Of course we need to spot whether this might be at an endpoint or within the range of integration.

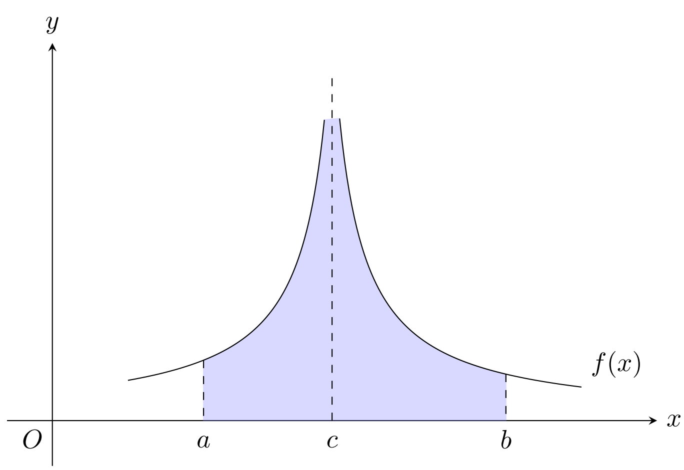
Take $\int_0^1 \frac{1}{\sqrt{x}}\mathrm{d}x$ as an example. There is a potential problem here, as $\frac{1}{\sqrt{x}}$ is infinite at the point $x = 0$. To decide the issue we can integrate from ϵ to 1, where $0 < \epsilon \ll 1$, and then take $\epsilon \to 0$.

$$I(\epsilon) = \lim_{\epsilon\to 0^+}\int_\epsilon^1 \frac{1}{\sqrt{x}}\mathrm{d}x = \lim_{\epsilon\to 0^+}\left[2\sqrt{x}\right]_\epsilon^1 = \lim_{\epsilon\to 0^+} 2 - 2\sqrt{\epsilon} = 2.$$

This means that this integral exists. However, sometimes there can be some integrals that do not exist.

$$\int_0^1 \frac{\mathrm{d}x}{x^2} = \lim_{\epsilon\to 0^+}\int_\epsilon^1 \frac{1}{x^2}\mathrm{d}x = \lim_{\epsilon\to 0^+}\left[-\frac{1}{x}\right]_\epsilon^1 = \lim_{\epsilon\to 0^+}\left(\frac{1}{\epsilon}-1\right) \Rightarrow \text{DNE.}$$

We should pay attention that sometimes the integral has upper and lower limits where the integrand is finite, but there are points of singularity between the upper and lower limits.



To do this, we write:

$$\int_a^b f(x)\mathrm{d}x = \lim_{k\to c^-}\int_a^k f(x)\mathrm{d}x + \lim_{m\to c^+}\int_m^b f(x)\mathrm{d}x. \tag{3.3}$$

3.3 Integration Techniques

3.3.1 Partial Fractions

$$\begin{aligned}
\int \frac{\mathrm{d}x}{x(x+1)} &= \int \left(\frac{1}{x} - \frac{1}{x+1}\right) \mathrm{d}x \\
&= \int \frac{1}{x}\mathrm{d}x - \int \frac{1}{x+1}\mathrm{d}x \\
&= \ln|x| - \ln|x+1| + C \\
&= \ln\left|\frac{x}{x+1}\right| + C.
\end{aligned}$$

For a more general and complicated case, we can say $\int \frac{P(x)}{Q(x)}\mathrm{d}x$. When integrating with the integrand being a rational function, we often integrate after splitting the rational function into simple fractions.

e.g. $\int \frac{x^2-x+4}{x^3-3x^2+2x}\mathrm{d}x$

It is obvious that $x^3-3x^2+2x = x(x-1)(x-2)$; therefore, we can write:

$$\frac{x^2-x+4}{x^3-3x^2+2x} = \frac{x^2-x+4}{x(x-1)(x-2)} \equiv \frac{A}{x} + \frac{B}{x-1} + \frac{C}{x-2}.$$

Therefore,

$$\begin{aligned}
x^2-x+4 &= A(x-1)(x-2) + Bx(x-2) + Cx(x-1) \\
&= A(x^2-3x+2) + B(x^2-2x) + C(x^2-x) \\
&= Ax^2 - 3Ax + 2A + Bx^2 - 2Bx + Cx^2 - Cx \\
&= (A+B+C)x^2 - (3A+2B+C)x + 2A.
\end{aligned}$$

$$\begin{cases} A+B+C=1 \\ 3A+2B+C=1 \\ 2A=4 \end{cases} \quad \Rightarrow \quad \begin{cases} A=2 \\ B=-4 \\ C=3. \end{cases}$$

Thus, $\frac{x^2-x+4}{x(x-1)(x-2)} \equiv \frac{2}{x} - \frac{4}{x-1} + \frac{3}{x-2}$.

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

When we encounter factors like $\dfrac{1}{(x-a)^2}$, we should write not only $\dfrac{1}{(x-a)^2}$, but also $\dfrac{1}{x-a}$ when decomposing the rational function. Also, when we encounter factors like $\dfrac{1}{ax^2+bx+c}$, which could not be decomposed anymore, we should write $\dfrac{Ax+B}{ax^2+bx+c}$ when decomposing. Now see a more complicated example.

e.g. $\displaystyle\int \frac{s^4+81}{s(x^2+9)^2}\mathrm{d}s$

$$\frac{s^4+81}{s(s^2+9)^2} \equiv \frac{A}{s} + \frac{Bs+C}{s^2+9} + \frac{Ds+E}{(s^2+9)^2}$$

Therefore,

$$\begin{aligned} s^4+81 &= A(s^2+9)^2 + (Bs+C)s(s^2+9) + (Ds+E)s \\ &= (A+B)s^4 + Cs^3 + (18A+9B+D)s^2 + (9C+E)s + 81A. \end{aligned}$$

So,

$$\begin{cases} A+B=1 \\ C=0 \\ 18A+9B+D=0 \\ 9C+E=0 \\ 81A=81 \end{cases} \quad \Rightarrow \quad \begin{cases} A=1 \\ B=0 \\ C=0 \\ D=-18 \\ E=0. \end{cases}$$

Thus, $\dfrac{s^4+81}{s(x^2+9)^2} = \dfrac{1}{s} - \dfrac{18s}{(s^2+9)^2}$.

$$\begin{aligned} \int \frac{s^4+81}{s(x^2+9)^2} &= \int \frac{1}{s}\mathrm{d}s - 9\int \frac{2s}{(s^2+9)^2}\mathrm{d}s \\ &= \int \frac{1}{s}\mathrm{d}s - 9\int \frac{1}{(s^2+9)^2}\,\mathrm{d}\big(s^2+9\big) \\ &= \ln|s| + \frac{9}{s^2+9} + C. \end{aligned}$$

3.3.2 Variable Substitution

(1) $\displaystyle\int x\,\mathrm{e}^{-x^2}\,\mathrm{d}x$

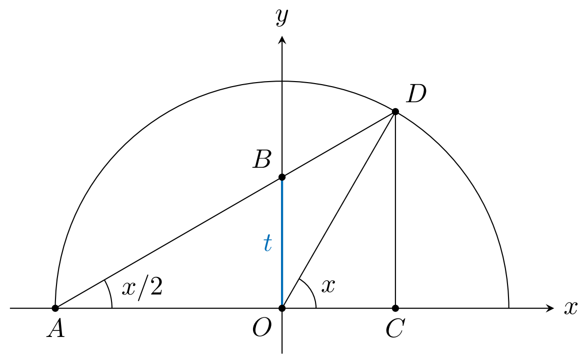
Let $u=x^2 \Rightarrow \mathrm{d}u = 2x\mathrm{d}x$. Therefore,

$$\int x\,\mathrm{e}^{-x^2}\,\mathrm{d}x \equiv \frac{1}{2}\int \mathrm{e}^{-u}\,\mathrm{d}u = -\frac{1}{2}\,\mathrm{e}^{-u} + C \equiv -\frac{1}{2}\,\mathrm{e}^{-x^2} + C.$$

(2) The Weierstrass Substistution: $t = \tan\left(\dfrac{x}{2}\right)$

For trigonometric integrals we can often try $t = \tan\left(\dfrac{x}{2}\right)$.

Consider a unit circle.



From the graph we can see that the substitution $t = \tan\left(\frac{x}{2}\right)$ can also represent the slope of the line AD. So, we can write:

$$y - 0 = t(x - (-1)) \Rightarrow y = tx + t.$$

To find the coordinates of D, use $y = tx + t$ and plug it into the equation of the unit circle. This gives

$$x_D = \frac{1 - t^2}{t^2 + 1}.$$

The x coordinate of D is actually $\cos x$ for a unit circle, so we know that:

$$\cos x = \frac{1 - t^2}{t^2 + 1}. \tag{3.4}$$

This method gives us a way to circumvent the double angle formula. Nevertheless, the double angle formula seems to come in a handy way. As $\tan\left(\frac{x}{2}\right) = \frac{|OB|}{|OA|} = t \equiv \frac{t}{1}$, we can see that $|OB| = 1$, and $|OA| = 1$. Therefore, $|AB| = \sqrt{|OA|^2 + |OB|^2} = \sqrt{1 + t^2}$. This means

$$\begin{cases} \sin\left(\frac{x}{2}\right) = \frac{|OB|}{|AB|} = \frac{t}{\sqrt{1 + t^2}} \\ \cos\left(\frac{x}{2}\right) = \frac{|OA|}{|AB|} = \frac{1}{\sqrt{1 + t^2}}. \end{cases} \tag{3.5}$$

Using the double angle formula,

$$\begin{cases} \sin x = 2\sin\left(\frac{x}{2}\right)\cos\left(\frac{x}{2}\right) = \frac{2t}{1 + t^2} \\ \cos x = 2\cos^2\left(\frac{x}{2}\right) - 1 = \frac{1 - t^2}{1 + t^2} \\ \tan x = \frac{\sin x}{\cos x} = \frac{2t}{1 - t^2}. \end{cases} \tag{3.6}$$

Besides those, we still need to figure out what $\frac{\mathrm{d}x}{\mathrm{d}t}$ is in this substistution. Through the formula $\tan\left(\frac{x}{2}\right) =$

t and differentiate it with respect to t,

$$\begin{aligned}
\sec^2\left(\frac{x}{2}\right)\cdot\frac{1}{2}\frac{\mathrm{d}x}{\mathrm{d}t} &= 1\\
\left(1+\tan^2\left(\frac{x}{2}\right)\right)\cdot\frac{1}{2}\frac{\mathrm{d}x}{\mathrm{d}t} &= 1\\
\left(1+t^2\right)\cdot\frac{\mathrm{d}x}{\mathrm{d}t} &= 2.
\end{aligned}$$

Therefore, we have

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \frac{2}{1+t^2}. \tag{3.7}$$

e.g. $\displaystyle\int\frac{\mathrm{d}x}{2+\cos x}$:

$$\begin{aligned}
\int\frac{\mathrm{d}x}{2+\cos x} &= \int\frac{1}{2+\left(\dfrac{1-t^2}{1+t^2}\right)}\frac{\mathrm{d}x}{\mathrm{d}t}\mathrm{d}t\\
&= \int\frac{2}{1+t^2}\frac{1+t^2}{3+t^2}\mathrm{d}t\\
&= 2\int\frac{\mathrm{d}t}{3+t^2}\\
&= \frac{2}{3}\frac{\mathrm{d}t}{1+t^2/3}\\
&= \frac{2}{\sqrt{3}}\frac{\mathrm{d}\left(t/\sqrt{3}\right)}{1+\left(t/\sqrt{3}\right)^2}\\
&= \frac{2}{\sqrt{3}}\arctan\left(\frac{t}{\sqrt{3}}\right)+C\\
&= \frac{2}{\sqrt{3}}\arctan\left(\frac{1}{\sqrt{3}}\tan\left(\frac{x}{2}\right)\right)+C.
\end{aligned}$$

3.3.3 Integration by Parts

Recall the product rule of differentiation:

$$\frac{\mathrm{d}(uv)}{\mathrm{d}x} = \frac{\mathrm{d}u}{\mathrm{d}x}v+\frac{\mathrm{d}v}{\mathrm{d}x}u.$$

Multiply $\mathrm{d}x$ on both sides, we are left with

$$\mathrm{d}(uv) = v\mathrm{d}u+u\mathrm{d}v.$$

Rearranging gives $u\mathrm{d}v = \mathrm{d}(uv)-v\mathrm{d}u$. Integrate both sides, and we get:

$$\int u\mathrm{d}v = uv-\int v\mathrm{d}u. \tag{3.8}$$

More often, what we see is

$$\int u\frac{\mathrm{d}v}{\mathrm{d}x}\mathrm{d}x = uv - \int v\mathrm{d}u. \tag{3.9}$$

e.g. $\int \ln x\mathrm{d}x$:

$$\begin{aligned}\int \ln x\mathrm{d}x &= x\ln x - \int x\,\mathrm{d}(\ln x)\\ &= x\ln x - \int x\cdot\frac{1}{x}\mathrm{d}x\\ &= x\ln x - x + C.\end{aligned}$$

3.3.4 Special Tricks

(1) Calculate $I = \int_0^{\pi/2}\sin^2\theta\mathrm{d}\theta$.

Consider the substitution $\phi = \frac{\pi}{2} - \theta$, or $\theta = \frac{\pi}{2} - \phi$.

$$\begin{aligned}I &= \int_0^{\pi/2}\sin^2\theta\mathrm{d}\theta\\ &= \int_{\theta=0}^{\theta=\pi/2}\sin^2\left(\frac{\pi}{2}-\phi\right)\mathrm{d}\left(\frac{\pi}{2}-\phi\right)\\ &= -\int_{\pi/2}^{0}\cos^2\phi\mathrm{d}\phi\\ &= \int_0^{\pi/2}\cos^2\phi\mathrm{d}\phi \equiv \int_0^{\pi/2}\cos^2\theta\mathrm{d}\theta \text{ (dummy variable).}\end{aligned}$$

Therefore,

$$2I = \int_0^{\pi/2}(\sin^2\theta + \cos^2\theta)\mathrm{d}\theta = \frac{\pi}{2},$$

and $I = \int_0^{\pi/2}\sin^2\theta\mathrm{d}\theta = \frac{\pi}{4}$.

(2) Calculate $J = \int_0^{\pi/2}\ln(\sin x)\mathrm{d}x$.

Consider the same trick $y = \frac{\pi}{2} - x$, or $x = \frac{\pi}{2} - y$.

$$\begin{aligned}J &= \int_0^{\pi/2}\ln(\sin x)\mathrm{d}x\\ &= \int_{x=0}^{x=\pi/2}\ln\left(\sin\left(\frac{\pi}{2}-y\right)\right)\mathrm{d}\left(\frac{\pi}{2}-y\right)\\ &= -\int_{\pi/2}^{0}\ln(\cos y)\mathrm{d}y\\ &= \int_0^{\pi/2}\ln(\cos y)\mathrm{d}y \equiv \int_0^{\pi/2}\ln(\cos x)\mathrm{d}x\end{aligned}$$

Therefore,

$$\begin{aligned}
2J &= \int_0^{\pi/2} \ln(\sin x)\mathrm{d}x + \int_0^{\pi/2} \ln(\cos x)\mathrm{d}x \\
&= \int_0^{\pi/2} \ln(\sin x \cos x)\mathrm{d}x \\
&= \int_0^{\pi/2} (\ln(\sin 2x) - \ln 2)\,\mathrm{d}x \\
&= \frac{1}{2}\int_0^{\pi/2} \ln(\sin 2x)\,\mathrm{d}(2x) - \frac{\pi}{2}\ln 2.
\end{aligned}$$

Consider another substitution $z = 2x$.

$$\frac{1}{2}\int_0^{\pi/2} \ln(\sin 2x)\,\mathrm{d}(2x) = \frac{1}{2}\int_0^{\pi} \ln(\sin z)\mathrm{d}z,$$

and by the symmetry of $\sin x$ on the domain $[0, \pi]$,

$$\frac{1}{2}\int_0^{\pi} \ln(\sin z)\mathrm{d}z = \frac{1}{2}\cdot 2\int_0^{\pi/2} \ln(\sin z)\mathrm{d}z = \int_0^{\pi/2} \ln(\sin z)\mathrm{d}z \equiv \int_0^{\pi/2} \ln(\sin x)\mathrm{d}x = J.$$

Thus,

$$2J = \frac{1}{2}\int_0^{\pi/2} \ln(\sin 2x)\,\mathrm{d}(2x) - \frac{\pi}{2}\ln 2 = J - \frac{\pi}{2}\ln 2,$$

and this finally gives $J = -\dfrac{\pi}{2}\ln 2$.

(3) Calculate $K_1 = \displaystyle\int_0^{\pi/2} \frac{\mathrm{d}x}{1+\tan^2 x}$.

This integral can be easily solved if we notice that $\dfrac{1}{1+\tan^2 x} \equiv \dfrac{1}{\sec^2 x} \equiv \cos^2 x$.

Therefore, $K_1 = \displaystyle\int_0^{\pi/2} \frac{\mathrm{d}x}{1+\tan^2 x} = \int_0^{\pi/2} \cos^2 x\mathrm{d}x = \frac{\pi}{4}$.

Now, consider $K_2 = \displaystyle\int_0^{\pi/2} \frac{\mathrm{d}x}{1+\tan^{\sqrt{2}} x}$. With the substitution $y = \dfrac{\pi}{2} - x$, or $x = \dfrac{\pi}{2} - y$,

$$\begin{aligned}
K_2 &= \int_0^{\pi/2} \frac{\mathrm{d}x}{1+\tan^{\sqrt{2}} x} \\
&= \int_{x=0}^{x=\pi/2} \frac{1}{1+\tan^{\sqrt{2}}\left(\frac{\pi}{2} - y\right)}\mathrm{d}\left(\frac{\pi}{2} - y\right) \\
&= -\int_{x=0}^{x=\pi/2} \frac{1}{1+\cot^{\sqrt{2}} y}\mathrm{d}y \\
&= -\int_{\pi/2}^{0} \frac{1}{1+1/\left(\tan^{\sqrt{2}} y\right)}\mathrm{d}y \\
&= \int_0^{\pi/2} \frac{\tan^{\sqrt{2}} y}{1+\tan^{\sqrt{2}} y}\mathrm{d}y \equiv \int_0^{\pi/2} \frac{\tan^{\sqrt{2}} x}{1+\tan^{\sqrt{2}} x}\mathrm{d}x.
\end{aligned}$$

Therefore,

$$2K_2 = \int_0^{\pi/2} \left(\frac{1}{1+\tan^{\sqrt{2}} x} + \frac{\tan^{\sqrt{2}} x}{1+\tan^{\sqrt{2}} x} \right) = \int_0^{\pi/2} 1\mathrm{d}x = \frac{\pi}{2},$$

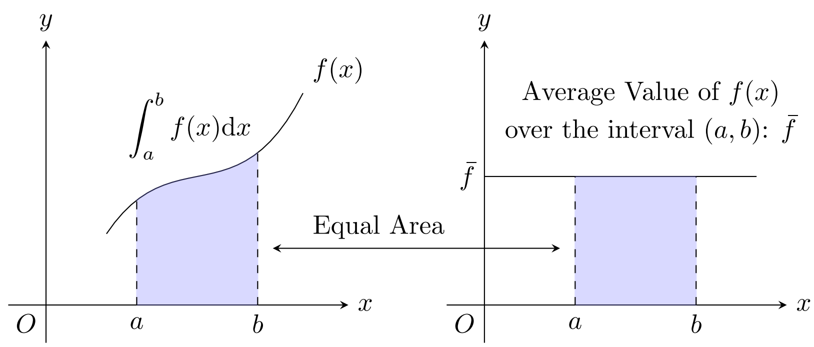
and $K_2 = \dfrac{\pi}{4}$.

Actually, we can replace the $\sqrt{2}$ power of $\tan x$ by any positive constant and still get the same result.

3.4 Mean Value, Area, and Length

3.4.1 Mean Value

Consider a function $f(x)$ and a specified interval $[a, b]$.



From the graph, we can see that

$$\int_a^b f(x)\mathrm{d}x = \bar{f} \cdot (b-a),$$

where $\bar{f}$ is defined to be the mean value of $f(x)$ over the interval (a, b). Therefore,

$$\bar{f} \equiv \langle f \rangle \equiv \frac{1}{b-a} \int_a^b f(x)\mathrm{d}x. \tag{3.10}$$

This $\langle f \rangle$ notation is inspired by Dirac in quantum mechanics.

Another mean value is called the **Root Mean Square Value** (rms value). For $f(x)$ over the interval (a, b),

$$f_{rms} \equiv \langle f^2 \rangle^{1/2} \equiv \sqrt{\frac{1}{b-a} \int_a^b f^2(x)\mathrm{d}x}. \tag{3.11}$$

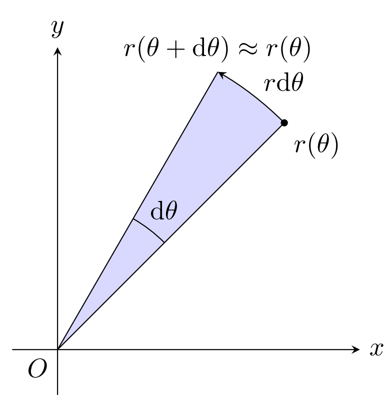
For a sine wave, $y = \sin x$,

$$\begin{aligned}
\langle \sin x \rangle &= \frac{1}{2\pi}\int_0^{2\pi} \sin x \mathrm{d}x = 0 = \langle \cos x \rangle \\
\langle \sin^2 x \rangle &= \frac{1}{2\pi}\int_0^{2\pi} \sin^2 x \mathrm{d}x = \frac{1}{2} \\
(\sin x)_{rms} = \langle \sin^2 x \rangle^{1/2} &= \frac{1}{\sqrt{2}}.
\end{aligned}$$

As we know from waves, the power of a wave is proportional to the square of its amplitude. It is obviously better to use the rms value instead of the common mean value to calculate the power of the wave.

3.4.2 Area in Polar Coordinates

Consider a polar curve rotating by an infinitesimal angle $\mathrm{d}\theta$. This angle $\mathrm{d}\theta$ is so small that we can neglect the change in the modulus r.



Now, look at the area $\mathrm{d}A$ covered in this rotation. As the radius is nearly unchanged, we can treat this curve as an arc. That is to say,

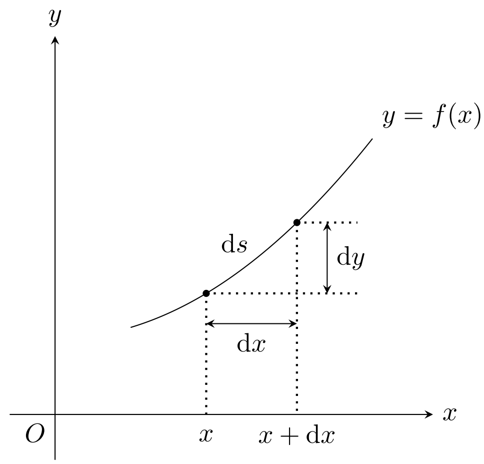
$$\mathrm{d}A = \pi r^2(\theta) \cdot \frac{\mathrm{d}\theta}{2\pi} = \frac{1}{2} r^2(\theta)\mathrm{d}\theta.$$

Therefore, the area under a polar curve from initial angle $\theta = \theta_1$ to final angle $\theta = \theta_2$ is:

$$A = \int \mathrm{d}A = \int_{\theta_1}^{\theta_2} \frac{1}{2} r^2 \mathrm{d}\theta. \tag{3.12}$$

3.4.3 Path Length

Cartesian Coordinates



Take an infinitesimal change in x direction $\mathrm{d}x$. The corresponding change in y is $\mathrm{d}y$, and the path length should be given by $\overset{\frown}{PQ}$. As $\mathrm{d}x \to 0$, the chord PQ approximates the curve $\overset{\frown}{PQ}$. Therefore,

$$\begin{aligned}
\mathrm{d}s^2 &= \mathrm{d}x^2 + \mathrm{d}y^2 \\
\left(\frac{\mathrm{d}s}{\mathrm{d}x}\right)^2 &= 1 + \left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2 \\
\frac{\mathrm{d}s}{\mathrm{d}x} &= \sqrt{1 + \left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2} \\
\mathrm{d}s &= \sqrt{1 + \left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2}\,\mathrm{d}x
\end{aligned}$$

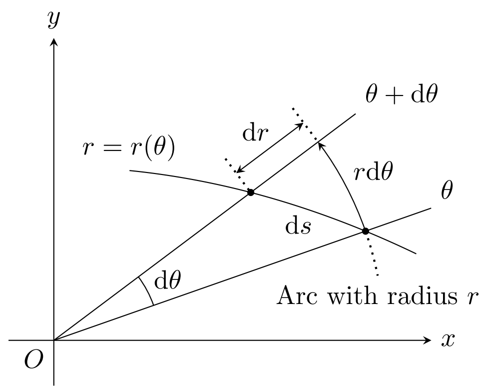
$$s = \int \mathrm{d}s = \int \sqrt{1 + \left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2}\,\mathrm{d}x. \tag{3.13}$$

Parametric Equations

$$\begin{aligned}
\mathrm{d}s^2 &= \mathrm{d}x^2 + \mathrm{d}y^2 \\
\left(\frac{\mathrm{d}s}{\mathrm{d}t}\right)^2 &= \left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2 + \left(\frac{\mathrm{d}y}{\mathrm{d}t}\right)^2 \\
\frac{\mathrm{d}s}{\mathrm{d}t} &= \sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2 + \left(\frac{\mathrm{d}y}{\mathrm{d}t}\right)^2} \\
\mathrm{d}s &= \sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2 + \left(\frac{\mathrm{d}y}{\mathrm{d}t}\right)^2}\,\mathrm{d}t.
\end{aligned}$$

$$s = \int \mathrm{d}s = \int \sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2 + \left(\frac{\mathrm{d}y}{\mathrm{d}t}\right)^2}\,\mathrm{d}t. \tag{3.14}$$

Polar Coordinates



Consider an infinitesimal change in the angle, $\mathrm{d}\theta$. The modulus changes by $\mathrm{d}r$, and the change perpendicular to $\mathrm{d}r$ can be approximated by an arc with radius r. When $\mathrm{d}\theta \to 0$, the arc can be seen as a straight line segment, so

$$\begin{aligned}
\mathrm{d}s^2 &= \mathrm{d}r^2 + (r\mathrm{d}\theta)^2 \\
\left(\frac{\mathrm{d}s}{\mathrm{d}\theta}\right)^2 &= \left(\frac{\mathrm{d}r}{\mathrm{d}\theta}\right)^2 + r^2 \\
\frac{\mathrm{d}s}{\mathrm{d}\theta} &= \sqrt{\left(\frac{\mathrm{d}r}{\mathrm{d}\theta}\right)^2 + r^2} \\
\mathrm{d}s &= \sqrt{r^2 + \left(\frac{\mathrm{d}r}{\mathrm{d}\theta}\right)^2}\,\mathrm{d}\theta,
\end{aligned}$$

$$s = \int \mathrm{d}s = \int_{\theta_1}^{\theta_2} \sqrt{r^2 + \left(\frac{\mathrm{d}r}{\mathrm{d}\theta}\right)^2}\,\mathrm{d}\theta. \tag{3.15}$$

3.5 Centre of Mass

Say we have N masses m_i $(i = 1, 2, \ldots, N)$ at positions (x_i, y_i). The centre of mass $G(\bar{x}, \bar{y})$ is defined by:

$$\begin{aligned}
\bar{x} &= \frac{\sum_{i=1}^{N} m_i x_i}{M} \\
\bar{y} &= \frac{\sum_{i=1}^{N} m_i y_i}{M},
\end{aligned} \tag{3.16}$$

where M is the total mass, $M = \sum_{i=1}^{N} m_i$.

The numerators of these expressions are often called the **first moments of mass**.

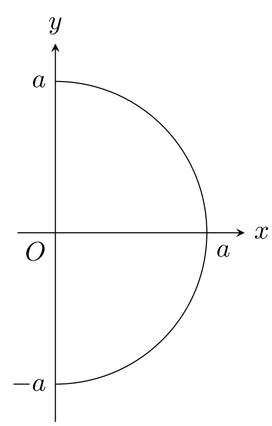
Now, we can generalise these ideas into continuous distributions of mass. Consider a very thin plate on a closed area D, and the surface density at the point (x, y) is given by $\rho(x, y)$. Now look at a closed area $\mathrm{d}\sigma$ so small that we can assume it has constant surface density. Therefore, the first moments of this area element is:

$$x\mathrm{d}m = x\rho(x,y)\mathrm{d}\sigma = x\rho(x,y)\mathrm{d}x\mathrm{d}y$$
$$y\mathrm{d}m = y\rho(x,y)\mathrm{d}\sigma = y\rho(x,y)\mathrm{d}x\mathrm{d}y.$$

And as the mass of the thin plate is given by $M = \int \mathrm{d}m = \iint_D \rho(x,y)\mathrm{d}x\mathrm{d}y$, we get:

$$\begin{aligned} \bar{x} &= \frac{\iint_D x\mathrm{d}m}{\int \mathrm{d}m} = \frac{\iint_D x\rho(x,y)\mathrm{d}x\mathrm{d}y}{\iint_D \rho(x,y)\mathrm{d}x\mathrm{d}y} \\ \bar{y} &= \frac{\iint_D y\mathrm{d}m}{\int \mathrm{d}m} = \frac{\iint_D y\rho(x,y)\mathrm{d}x\mathrm{d}y}{\iint_D \rho(x,y)\mathrm{d}x\mathrm{d}y}. \end{aligned} \tag{3.17}$$

Sometimes we can use polar coordinates to calculate centre of mass. Consider a uniform wire of a semicircle with radius a centered at the origin.



In this case, assume the linear density is λ.

$$\mathrm{d}m = \lambda\mathrm{d}l = \lambda \cdot a\mathrm{d}\theta$$

$$M = \int \mathrm{d}m = \lambda a \int_{-\pi/2}^{\pi/2} \mathrm{d}\theta = \lambda a\pi.$$

Therefore,

$$\begin{aligned}\bar{x} &= \frac{1}{M}\int x\mathrm{d}m \\ &= \frac{1}{M}\int_{-\pi/2}^{\pi/2}(a\cos\theta)(\lambda a\mathrm{d}\theta) \\ &= \frac{\lambda a^2}{M}\cdot\sin\theta\Big|_{-\pi/2}^{\pi/2} \\ &= \frac{2\lambda a^2}{\lambda a\pi} \\ &= \frac{2a}{\pi}.\end{aligned}$$

Similarly,

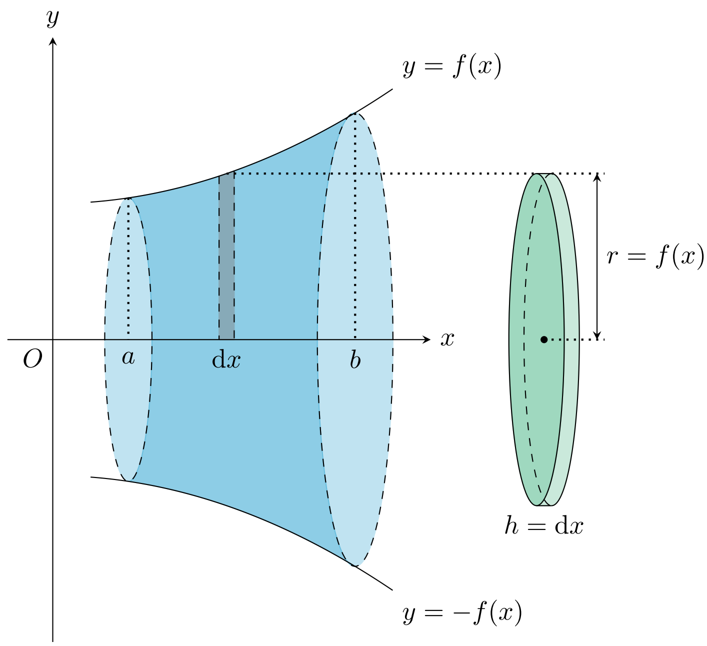
$$\begin{aligned}\bar{y} &= \frac{1}{M}\int y\mathrm{d}m \\ &= \frac{1}{M}\int_{-\pi/2}^{\pi/2}(a\sin\theta)(\lambda a\mathrm{d}\theta) \\ &= \frac{\lambda a^2}{M}\int_{-\pi/2}^{\pi/2}\sin\theta\mathrm{d}\theta \\ &= \frac{\lambda a^2}{M}(-\cos\theta)\Big|_{-\pi/2}^{\pi/2} \\ &= 0.\end{aligned}$$

So, the coordinates of the centre of mass are $\left(\frac{2a}{\pi}, 0\right)$.

3.6 Volume of Revolution

3.6.1 Disks

Take function $y = f(x)$ in the 2-dimensional plane and rotate it about the x axis. what we get is a 3-dimensional shape.



Consider a small strip with x coordinate x and strip width $\mathrm{d}x$. The width is so small that we can assume the height $f(x)$ is unchanged with this change of $\mathrm{d}x$. When rotated about the x axis, this strip generates a thin disk volume of $\mathrm{d}V$:

$$\begin{aligned}\mathrm{d}V &= (\text{area of disk}) \cdot (\text{height of disk}) \\ &= \pi y^2 \mathrm{d}x.\end{aligned}$$

Therefore,

$$V = \int \mathrm{d}V = \int_a^b \pi y^2 \mathrm{d}x \tag{3.18}$$

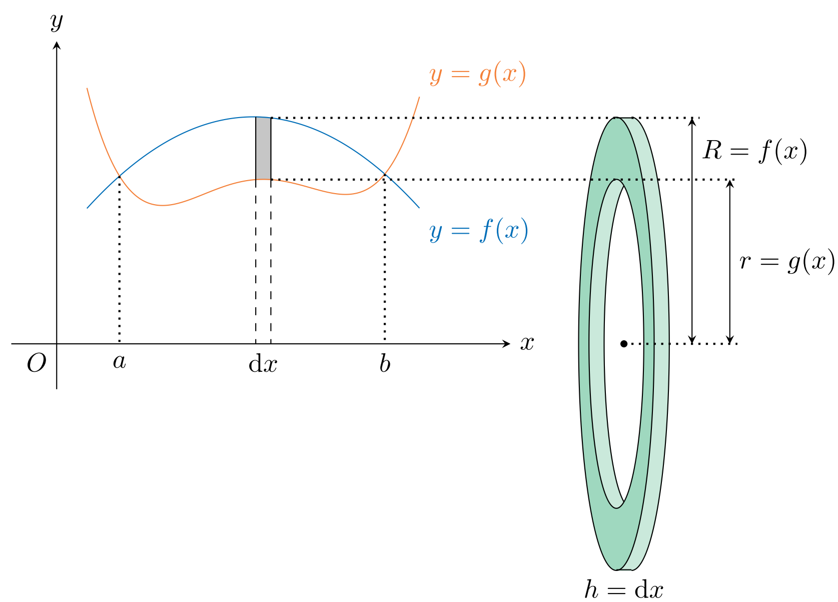
is the volume of revolution when a curve of $y = f(x)$ is rotated about the x axis.

When a curve is rotated about the y axis, we write:

$$V = \int \mathrm{d}V = \int_a^b \pi x^2 \mathrm{d}y. \tag{3.19}$$

3.6.2 Washers

The method of washers is used to deal with volume of revolution between two curves.



Consider a small rectangle bounded between $y = f(x)$ and $y = g(x)$ with width $\mathrm{d}x$. When rotated about the x axis, the corresponding volume $\mathrm{d}V$ is:

$$\begin{aligned}
\mathrm{d}V &= (\text{area of washer}) \cdot (\text{height of washer}) \\
&= \pi \left(R^2 - r^2\right) \mathrm{d}x \\
&= \pi \left([f(x)]^2 - [g(x)]^2\right) \mathrm{d}x.
\end{aligned}$$

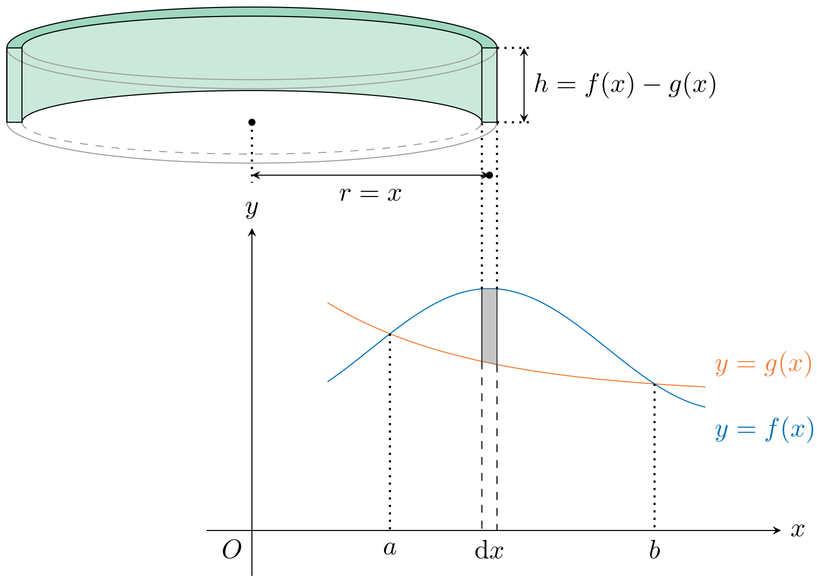
Therefore,

$$V = \int \mathrm{d}V = \int_a^b \pi \left([f(x)]^2 - [g(x)]^2\right) \mathrm{d}x. \tag{3.20}$$

When rotated about the y axis, we write:

$$V = \int \mathrm{d}V = \int_a^b \pi \left([f(y)]^2 - [g(y)]^2\right) \mathrm{d}y. \tag{3.21}$$

3.6.3 Shell



Consider again a small rectangle bounded between $y = f(x)$ and $y = g(x)$ with width $\mathrm{d}x$. However, this time, this rectangle is rotated about the y axis. What we get this time is a very thin cylindrical shell. Its height is $h = f(x) - g(x)$, radius is $r = x$, and the thickness is $\mathrm{d}x$. As this shell is very thin, we can approximate its volume by thinking it as a long cuboid with length $2\pi r$, height h, and width $\mathrm{d}x$.

$$\begin{aligned}\mathrm{d}V &= (\text{length of cuboid}) \cdot (\text{height of cuboid}) \cdot (\text{width of cuboid}) \\ &= 2\pi x[f(x) - g(x)]\mathrm{d}x.\end{aligned}$$

Therefore,

$$V = \int \mathrm{d}V = \int_a^b 2\pi x[f(x) - g(x)]\mathrm{d}x. \tag{3.22}$$

When rotated about the x axis, we write:

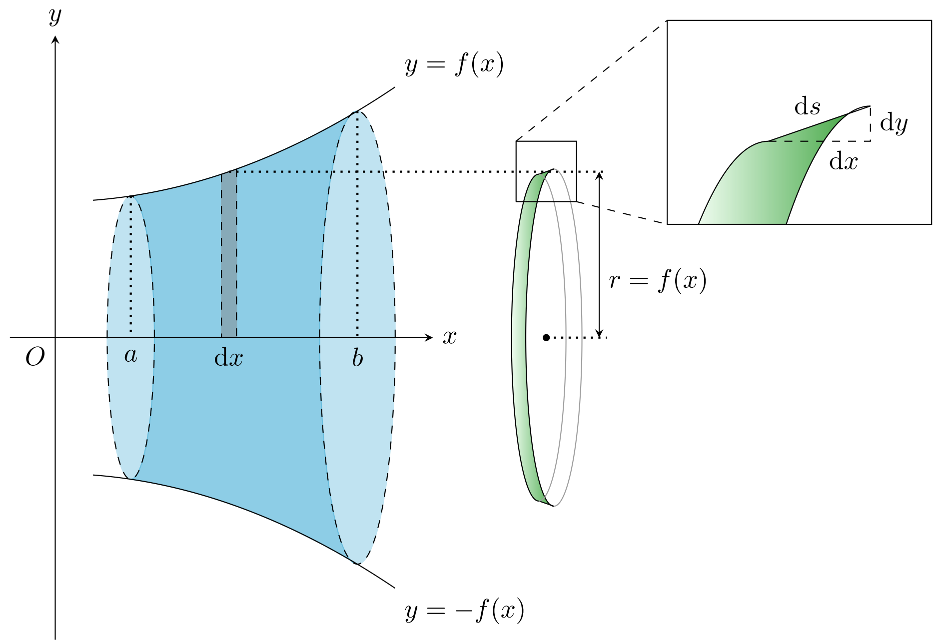
$$V = \int \mathrm{d}V = \int_a^b 2\pi y[f(y) - g(y)]\mathrm{d}y. \tag{3.23}$$

To sum up,

$$\begin{cases} \text{disk: } \mathrm{d}V = \pi r^2 \mathrm{d}x & V = \pi \displaystyle\int_a^b r^2 \mathrm{d}x \\ \text{washer: } \mathrm{d}V = \pi(R^2 - r^2)\mathrm{d}x & V = \pi \displaystyle\int_a^b (R^2 - r^2)\mathrm{d}x \\ \text{shell: } \mathrm{d}V = 2\pi r h \mathrm{d}x & V = 2\pi \displaystyle\int_a^b rh\mathrm{d}x. \end{cases} \tag{3.24}$$

One should pay attention to the difference between the shell method and the washer method. The washer has its height $\mathrm{d}x$, while the shell has its thickness $\mathrm{d}x$.

3.7 Surface Area of Revolution



Consider $\mathrm{d}x$. The curve length associated with this $\mathrm{d}x$ is $\mathrm{d}s$. As $\mathrm{d}x$ is so small that we can neglect the change in $f(x)$. Therefore, we can approximate the surface area of this surface of revolution by a rectangle of length $2\pi r$ and width $\mathrm{d}s$.

$$\begin{aligned}\mathrm{d}S &= (\text{length of rectangle}) \cdot (\text{width of rectangle}) \\ &= 2\pi r \mathrm{d}s \\ &= 2\pi y \sqrt{1 + \left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2} \mathrm{d}x.\end{aligned}$$

Therefore, the surface area of revolution is given by the formula:

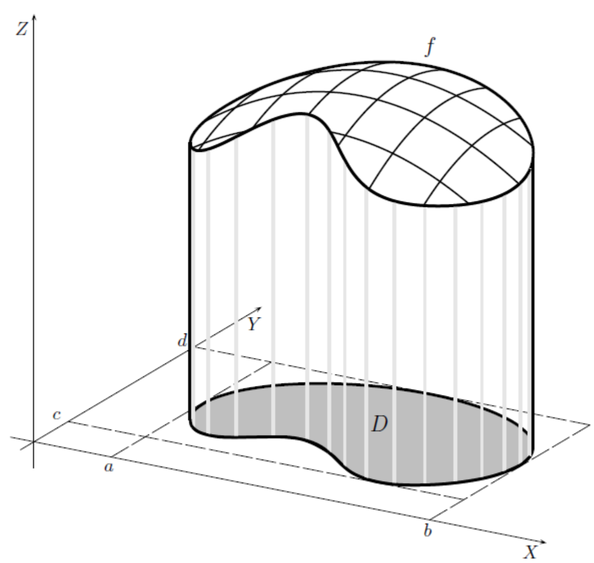
$$S = \int \mathrm{d}S = \int_a^b 2\pi y \sqrt{1 + \left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2} \mathrm{d}x. \tag{3.25}$$

Chapter 4

Partial Differentiation

4.1 Definition

Consider a function $u = u(x, y)$ of **2 independent variables** x and y. We can think of u as being the height of a surface above the (x, y) plane.



On the xOy plane we can find the domain of definition of this function.

To begin our discussion of partial derivatives, imagine a random point $P(x, y)$. Move a small distance δx in the x direction to $Q(x + \delta x, y)$ (i.e. keeping y fixed). Then, we define (if this limit exists) the partial derivative of u with respect to x

$$\frac{\partial u}{\partial x} = \lim_{\delta x \to 0} \left[\frac{u(x+\delta x, y) - u(x,y)}{\delta x} \right] \tag{4.1}$$

to be the rate of change of u with respect to x at P (keeping y fixed).

Similarly, we can define the partial derivative of u with respect to y:

$$\frac{\partial u}{\partial y} = \lim_{\delta y \to 0} \left[\frac{u(x, y+\delta y) - u(x,y)}{\delta y} \right]. \tag{4.2}$$

For example, consider $u = x^2 \sin y + y^3$:

$$\begin{aligned} \frac{\partial u}{\partial x} &= 2x \sin y \\ \frac{\partial u}{\partial y} &= x^2 \cos y + 3y^2. \end{aligned}$$

For higher order partial derivatives, we write:

$$\begin{aligned} \frac{\partial^2 u}{\partial x^2} &\equiv \frac{\partial}{\partial x}\left(\frac{\partial u}{\partial x}\right) \\ \frac{\partial^2 u}{\partial x \partial y} &\equiv \frac{\partial}{\partial x}\left(\frac{\partial u}{\partial y}\right), \text{ etc.} \end{aligned} \tag{4.3}$$

For the example above, we have:

$$\begin{aligned} \frac{\partial^2 u}{\partial x^2} &= 2 \sin y \\ \frac{\partial^2 u}{\partial y^2} &= -x^2 \sin y + 6y \\ \frac{\partial^2 u}{\partial y \partial x} &= 2x \cos y \\ \frac{\partial^2 u}{\partial x \partial y} &= 2x \cos y. \end{aligned}$$

We note that

$$\frac{\partial^2 u}{\partial y \partial x} = \frac{\partial^2 u}{\partial x \partial y}$$

in this case. This is a general result (requiring only the continuity of LHS and RHS). Therefore, we generally have

$$u_{xxyyx} = u_{yxyxx} = \dots.$$

Consider the function $u(x,t) = a\sin(x-ct)$.

$$\begin{aligned}
\frac{\partial u}{\partial x} &= a\cos(x-ct)\\
\frac{\partial u}{\partial t} &= -ac\cos(x-ct)\\
\frac{\partial^2 u}{\partial x^2} &= -a\sin(x-ct)\\
\frac{\partial^2 u}{\partial t^2} &= -ac^2\sin(x-ct)
\end{aligned}$$

From the above calculations, we can see that $u(x,t)$ satisfies **the wave equation**:

$$\frac{\partial^2 u}{\partial x^2} = \frac{1}{c^2}\frac{\partial^2 u}{\partial t^2}. \tag{4.4}$$

In fact, any reasonable function $f(x-ct)$ will satisfy this equation. It represents a wave form moving to the right ($c>0$).

Another equation about the partial derivatives is **the Laplace's Equation**. In the (x,y) plane, we write Laplace's Equation as:

$$\nabla^2 u \equiv \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0. \tag{4.5}$$

4.2 Gradient, Divergence, and Curl

4.2.1 Vector Differential Operator, $\boldsymbol{\nabla}$

The symbol ∇ is the vector differential operator. In the (x,y) plane,

$$\boldsymbol{\nabla} = \frac{\partial}{\partial x}\hat{x} + \frac{\partial}{\partial y}\hat{y}. \tag{4.6}$$

In a n-dimensional space,

$$\boldsymbol{\nabla} = \frac{\partial}{\partial x_1}\hat{x_1} + \frac{\partial}{\partial x_2}\hat{x_2} + \cdots + \frac{\partial}{\partial x_n}\hat{x_n} = \sum_{i=1}^{n}\frac{\partial}{\partial x_i}\hat{x_i}. \tag{4.7}$$

It is important to know the expression of the vector differential operator in two dimensional polar coordinates.

$$\boldsymbol{\nabla} = \frac{\partial}{\partial r}\hat{r} + \frac{1}{r}\frac{\partial}{\partial \theta}\hat{\theta}. \tag{4.8}$$

4.2.2 Gradient

The gradient of a scalar field f is denoted by $\boldsymbol{\nabla} f$.

In vector calculus, the gradient of a scalar-valued differentiable function f of several variables is the vector field $\boldsymbol{\nabla} f$ whose value at a point p is the "direction and rate of fastest increase". If the gradient of a function is non-zero at a point p, the direction of the gradient is the direction in which the function increases most quickly from p, and the magnitude of the gradient is the rate of increase in that direction, the greatest absolute directional derivative. Further, a point where the gradient is the zero vector is known as a stationary point. The gradient thus plays a fundamental role in optimization theory, where it is used to maximize a function by gradient ascent.

To sum up, the gradient acts on a scalar field, and the output is a vector field. The direction and magnitude of gradient represent "the direction and rate of fastest increase." The following expression means taking the gradient of f ($\boldsymbol{\nabla} f$):

$$\boldsymbol{\nabla} f = \frac{\partial f}{\partial x}\hat{x} + \frac{\partial f}{\partial y}\hat{y} + \frac{\partial f}{\partial z}\hat{z}. \tag{4.9}$$

4.2.3 Divergence

The divergence of a scalar field $\boldsymbol{F}$ is denoted by $\boldsymbol{\nabla} \cdot \boldsymbol{F}$.

In vector calculus, divergence is a vector operator that operates on a vector field, producing a scalar field giving the quantity of the vector field's source at each point. More technically, the divergence represents the volume density of the outward flux of a vector field from an infinitesimal volume around a given point.

As an example, consider air as it is heated or cooled. The velocity of the air at each point defines a vector field. While air is heated in a region, it expands in all directions, and thus the velocity field points outward from that region. The divergence of the velocity field in that region would thus have a positive value. While the air is cooled and thus contracting, the divergence of the velocity has a negative value.

To sum up, the divergence acts on a vector field, and the output is a scalar field. Physically, it represents the volume density of a flux. If divergence is not zero at a point, then this point acts as a source of the vector field. The following equation means taking the gradient of $\boldsymbol{F}$ ($\boldsymbol{\nabla} \cdot \boldsymbol{F}$):

$$\boldsymbol{\nabla} \cdot \boldsymbol{F} = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right) \cdot (F_x, F_y, F_z) = \frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} + \frac{\partial F_z}{\partial z}. \tag{4.10}$$

4.2.4 Curl

The curl of a vector field $\boldsymbol{F}$ is denoted by $\boldsymbol{\nabla} \times \boldsymbol{F}$.

In vector calculus, the curl is a vector operator that describes the infinitesimal circulation of a vector field in three-dimensional Euclidean space. The curl at a point in the field is represented by a vector whose length and direction denote the magnitude and axis of the maximum circulation. The curl of a field is formally defined as the circulation density at each point of the field.

A vector field whose curl is zero is called irrotational. The curl is a form of differentiation for vector fields. The corresponding form of the fundamental theorem of calculus is Stokes' theorem, which relates the surface integral of the curl of a vector field to the line integral of the vector field around the boundary curve.

To sum up, the curl acts on a vector field, and the output is a vector field. Physically, it represents the surface density of a circulation. If curl is not zero at a point, then this point acts as a source introducing swirls to the field. The following equation means taking the curl of $\boldsymbol{F}$ ($\boldsymbol{\nabla} \times \boldsymbol{F}$):

$$\begin{aligned} \boldsymbol{\nabla} \times \boldsymbol{F} &= \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right) \times (F_x, F_y, F_z) \\ &= \begin{vmatrix} \hat{x} & \hat{y} & \hat{z} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ F_x & F_y & F_z \end{vmatrix}. \end{aligned} \tag{4.11}$$

4.2.5 Summary

	Input	Output	Physical Meaning
Gradient ($\boldsymbol{\nabla}$)	Scalar Field	Vector Field	N.A.
Divergence ($\boldsymbol{\nabla}\cdot$)	Vector Field	Scalar Field	Volume Density of a Flux
Curl ($\boldsymbol{\nabla}\times$)	Vector Field	Vector Field	Surface Density of a Circulation

The applications of these operators involve the famous **Maxwell's Equations**:

$$\begin{aligned} \boldsymbol{\nabla} \cdot \boldsymbol{D} &= \rho \\ \boldsymbol{\nabla} \cdot \boldsymbol{B} &= 0 \\ \boldsymbol{\nabla} \times \boldsymbol{E} &= -\frac{\partial \boldsymbol{B}}{\partial t} \\ \boldsymbol{\nabla} \times \boldsymbol{H} &= \boldsymbol{J} + \frac{\partial \boldsymbol{D}}{\partial t}. \end{aligned} \tag{4.12}$$

The first equation says, the divergence of electric displacement is equal to the volume density of free charges (ρ). The second equation indicates that the magnetic field is a field without sources. The third equation says, the surface density of the circulation of electric field strength numerically equals to the rate of change of magnetic field strength. The last equation talks about the curl of the magnetic field. The curl of the magnetic field strength is equal to the total current density (conduction current and the displacement current).

It is also important to note these three relations:

$$\begin{aligned}\boldsymbol{D} &= \epsilon \boldsymbol{E} \\ \boldsymbol{B} &= \mu \boldsymbol{H} \\ \boldsymbol{J} &= \sigma \boldsymbol{E},\end{aligned} \tag{4.13}$$

where ϵ is the permittivity of the material, μ is the permeability of the material, σ is the electrical conductivity of the material.

As a reminder, here come the Maxwell's Equations in integration form:

$$\begin{aligned}&\oint_S \boldsymbol{D} \cdot \mathrm{d}\boldsymbol{S} = q_0 \\ &\oint_S \boldsymbol{B} \cdot \mathrm{d}\boldsymbol{S} = 0 \\ &\oint_L \boldsymbol{E} \cdot \mathrm{d}\boldsymbol{l} = -\iint_S \frac{\partial \boldsymbol{B}}{\partial t} \cdot \mathrm{d}\boldsymbol{S} \\ &\oint_L \boldsymbol{H} \cdot \mathrm{d}\boldsymbol{l} = I_0 + \iint_S \frac{\partial \boldsymbol{D}}{\partial t} \cdot \boldsymbol{S}.\end{aligned} \tag{4.14}$$

4.2.6 The Laplace Operator, ∇^2

In mathematics, the Laplace operator, or Laplacian, is a differential operator given by the divergence of the gradient of a scalar function on Euclidean space. It is usually denoted by $\boldsymbol{\nabla} \cdot \boldsymbol{\nabla}$, or ∇^2.

In two dimensional Cartesian Coordinates, we write:

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}. \tag{4.15}$$

It is also important to know the polar form of this operator:

$$\nabla^2 = \frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial^2}{\partial \theta^2}. \tag{4.16}$$

The Laplace Equation states that the sum of the second-order partial derivatives of R, the unknown function, equals zero:

$$\nabla^2 R = \frac{\partial^2 R}{\partial x^2} + \frac{\partial^2 R}{\partial y^2} = \frac{\partial^2 R}{\partial r^2} + \frac{1}{r}\frac{\partial R}{\partial r} + \frac{1}{r^2}\frac{\partial^2 R}{\partial \theta^2} = 0. \tag{4.17}$$

4.3 The Total Differential

When we have a function $f(x)$ of a single variable x and we make a small change $x \to x + \delta x$, so that $f \to f + \delta f$, then

$$\delta f \simeq \frac{\mathrm{d}f}{\mathrm{d}x}\delta x \text{ (increments).}$$

In the limit $\delta x \to 0$,

$$\mathrm{d}f = \frac{\mathrm{d}f}{\mathrm{d}x}\mathrm{d}x \text{ (differentials).}$$

Now, for a function of two variables $u(x, y)$, small changes $x \to x+\delta x$ and $y \to y+\delta y$ lead to $u \to u+\delta u$, with

$$\delta u \simeq \frac{\partial u}{\partial x}\delta x + \frac{\partial u}{\partial y}\delta y \text{ (increments).}$$

In the limit $\delta x \to 0$ and $\delta y \to 0$, we get:

$$\mathrm{d}u = \frac{\partial u}{\partial x}\mathrm{d}x + \frac{\partial u}{\partial y}\mathrm{d}y. \tag{4.18}$$

This is called the **total differential** of $u(x, y)$.

For example, see the function $u = x^2 \sin y + y^3$:

$$\mathrm{d}u = (2x \sin y)\mathrm{d}x + (x^2 \cos y + 3y^2)\mathrm{d}y.$$

4.4 The Chain Rule

Previously, when we have $u = f(x)$ and $x = g(t)$, we get:

$$\frac{\mathrm{d}u}{\mathrm{d}t} = \frac{\mathrm{d}u}{\mathrm{d}x} \cdot \frac{\mathrm{d}x}{\mathrm{d}t} = f'(x)g'(t) \equiv f'(g(t))g'(t).$$

Now, consider the function $u(x, y)$ with $x = x(t)$ and $y = y(t)$. By recalling the total differential

$$\mathrm{d}u = \frac{\partial u}{\partial x}\mathrm{d}x + \frac{\partial u}{\partial y}\mathrm{d}y$$

and divide the whole expression by $\mathrm{d}t$, we get:

$$\frac{\mathrm{d}u}{\mathrm{d}t} = \frac{\partial u}{\partial x}\frac{\mathrm{d}x}{\mathrm{d}t} + \frac{\partial u}{\partial y}\frac{\mathrm{d}y}{\mathrm{d}t}. \tag{4.19}$$

It is important to pay attention to the notations here. We use $\frac{\mathrm{d}x}{\mathrm{d}t}$ and $\frac{\mathrm{d}y}{\mathrm{d}t}$ instead of partials, because

x and y are functions of only t. We use $\dfrac{\partial u}{\partial x}$ and $\dfrac{\partial u}{\partial y}$ instead of d, because u is a function of both x and y. Also, in terms of t, u is again a function of **only** t, so we write $\dfrac{\mathrm{d}u}{\mathrm{d}t}$.

If we set $x = x(r, s)$ and $y = y(r, s)$, then

$$\begin{aligned}\frac{\partial u}{\partial r} &= \frac{\partial u}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial u}{\partial y}\frac{\partial y}{\partial r}\\ \frac{\partial u}{\partial s} &= \frac{\partial u}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial u}{\partial y}\frac{\partial y}{\partial s}.\end{aligned}$$

Now, consider the partial derivatives of a function $z = f(u, x, y)$ where $u = \psi(x, y)$.

There are several ways we could approach this problem.

Firstly, we can write down the total differential of z:

$$\mathrm{d}z = \frac{\partial f}{\partial u}\mathrm{d}u + \frac{\partial f}{\partial x}\mathrm{d}x + \frac{\partial f}{\partial y}\mathrm{d}y.$$

Then, as u is a function of x and y, $\mathrm{d}u$ could further be written as:

$$\mathrm{d}u = \frac{\partial u}{\partial x}\mathrm{d}x + \frac{\partial u}{\partial y}\mathrm{d}y.$$

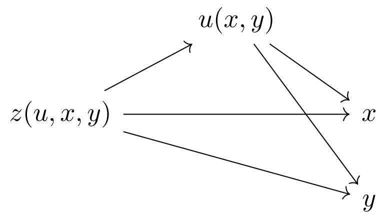
Plugging this expression back to the expression of $\mathrm{d}z$, and we get:

$$\begin{aligned}\mathrm{d}z &= \frac{\partial f}{\partial u}\mathrm{d}u + \frac{\partial f}{\partial x}\mathrm{d}x + \frac{\partial f}{\partial y}\mathrm{d}y\\ &= \frac{\partial f}{\partial u}\left(\frac{\partial u}{\partial x}\mathrm{d}x + \frac{\partial u}{\partial y}\mathrm{d}y\right) + \frac{\partial f}{\partial x}\mathrm{d}x + \frac{\partial f}{\partial y}\mathrm{d}y\\ &= \left(\frac{\partial f}{\partial u}\cdot\frac{\partial u}{\partial x} + \frac{\partial f}{\partial x}\right)\mathrm{d}x + \left(\frac{\partial f}{\partial u}\cdot\frac{\partial u}{\partial y} + \frac{\partial f}{\partial y}\right)\mathrm{d}y.\end{aligned}$$

We can recognise the above equation with the form $\mathrm{d}z = \dfrac{\partial z}{\partial x}\mathrm{d}x + \dfrac{\partial z}{\partial y}\mathrm{d}y$, and

$$\begin{aligned}\frac{\partial z}{\partial x} &= \frac{\partial f}{\partial u}\cdot\frac{\partial u}{\partial x} + \frac{\partial f}{\partial x}\\ \frac{\partial z}{\partial y} &= \frac{\partial f}{\partial u}\cdot\frac{\partial u}{\partial y} + \frac{\partial f}{\partial y}.\end{aligned} \tag{4.20}$$

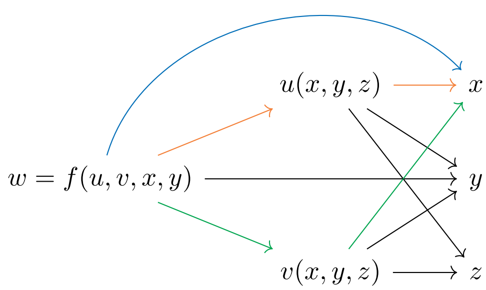
Pay attention here to the different meanings of $\dfrac{\partial z}{\partial x}$ and $\dfrac{\partial f}{\partial x}$. $\dfrac{\partial z}{\partial x}$ on the LHS sees u as the function of x and y, and it is the partial derivative of the composite function $z = f(u(x, y), x, y)$ with respect to x. $\dfrac{\partial f}{\partial x}$ on the RHS, on the other hand, treats z as a function of three variables u, x, and y. That is to say, $\dfrac{\partial f}{\partial x}$ means taking the partial derivative of x while keeping both u and y as constants. It is the derivative of the three-variable function $z = f(u, x, y)$.



There are too many types of composite functions to enumerate them one by one. If we draw the dependencies of these independent, intermediate, and dependent variables out by a graph like above, there can be a neat way to deal with the problem. For example, if we are trying to find $\dfrac{\partial z}{\partial x}$, just find all the arrows that connect z to x. For a single chain, take the (partial) derivative of the former variable with respect to the latter, and multiply them together. Then, add the results from all different chains together.

Consider the function $w = f(x + y + z, xyz, x, y)$ where f has continuous second partial derivatives. Calculate $\dfrac{\partial w}{\partial x}$ and $\dfrac{\partial^2 w}{\partial x \partial z}$.

Let $u = x + y + z$ and $v = xyz$ such that $w = f(u, v, x, y)$. We can draw out the diagram for this composite function. Then, we can mark all arrows that reach x, and trace them back.



Therefore,

$$
\begin{aligned}
\frac{\partial w}{\partial x} &= \frac{\partial f}{\partial x} + \frac{\partial f}{\partial u} \cdot \frac{\partial u}{\partial x} + \frac{\partial f}{\partial v} \cdot \frac{\partial v}{\partial x} \\
&= \frac{\partial f}{\partial x} + \frac{\partial f}{\partial u} + yz \frac{\partial f}{\partial v}.
\end{aligned}
$$

As the second order partial derivatives are continuous, we know that

$$
\frac{\partial^2 w}{\partial x \partial z} = \frac{\partial^2 w}{\partial z \partial x}.
$$

So, to get $\dfrac{\partial^2 w}{\partial x \partial z}$, just take the partial derivative of $\dfrac{\partial w}{\partial x}$ with respect to z:

$$\begin{aligned}
\frac{\partial^2 w}{\partial x \partial z} &= \frac{\partial}{\partial z}\left(\frac{\partial w}{\partial x}\right) \\
&= \frac{\partial}{\partial z}\left(\frac{\partial f}{\partial x} + \frac{\partial f}{\partial u}\cdot\frac{\partial u}{\partial x} + \frac{\partial f}{\partial v}\cdot\frac{\partial v}{\partial x}\right) \\
&= \frac{\partial}{\partial z}\left(\frac{\partial f}{\partial x} + \frac{\partial f}{\partial u} + yz\frac{\partial f}{\partial v}\right) \\
&= \frac{\partial^2 f}{\partial z \partial x} + \frac{\partial^2 f}{\partial z \partial u} + y\frac{\partial f}{\partial v} + yz\frac{\partial^2 f}{\partial z \partial v}.
\end{aligned}$$

Note that $\dfrac{\partial}{\partial z}\left(\dfrac{\partial f}{\partial x}\right)$ is not zero. While we take the derivative of f with respect to x by treating w as the function of four variables (x, y, u, and v), this partial derivative is still a function of x, y, u, and v, which indicates its dependence of z. Therefore, a slight change δz would still trigger a change in the partial derivative $\dfrac{\partial f}{\partial x}$, and $\dfrac{\partial}{\partial z}\left(\dfrac{\partial f}{\partial x}\right)$ is not zero.

4.5 Implicit Functions

Suppose we have a function defined implicitly:

$$F(x, y) = 0. \tag{4.21}$$

Then, F does not change as x and y change. However, we can still write the total differential:

$$\mathrm{d}F = \frac{\partial F}{\partial x}\mathrm{d}x + \frac{\partial F}{\partial y}\mathrm{d}y = 0.$$

Therefore, the derivative of y with respect to x is given by:

$$\frac{\mathrm{d}y}{\mathrm{d}x} = -\frac{F_x}{F_y}. \tag{4.22}$$

Now look at an implicit function that involves three variables:

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

While the function $F(x, y) = 0$ constrains the point (x, y) to be on a particular curve, the function

$F(x,y,z)=0$ constrains the point (x,y,z) to be on a particular surface. By writing the total differential,

$$\mathrm{d}F=\frac{\partial F}{\partial x}\mathrm{d}x+\frac{\partial F}{\partial y}\mathrm{d}y+\frac{\partial F}{\partial z}\mathrm{d}z=0,$$

we can see that:

$$\begin{cases}\text{At constant } y\text{:} & \dfrac{\partial z}{\partial x}=-\dfrac{F_x}{F_z}\\ \text{At constant } x\text{:} & \dfrac{\partial z}{\partial y}=-\dfrac{F_y}{F_z}\\ \text{At constant } z\text{:} & \dfrac{\partial y}{\partial x}=-\dfrac{F_x}{F_y}.\end{cases} \tag{4.23}$$

All of these equations could be easily derived from the total differential.

Also, note that e.g. $\left(\dfrac{\partial z}{\partial x}\right)_y=\dfrac{1}{\left(\frac{\partial x}{\partial z}\right)_y}$. This is because the variable y is being kept constant on both sides. We are just looking at the variations on a constant y slice of the $F=0$ surface.

4.6 Exact Differentials

We know that for a function of two variables $u(x,y)$, the total differential is:

$$\mathrm{d}u=\frac{\partial u}{\partial x}\mathrm{d}x+\frac{\partial u}{\partial y}\mathrm{d}y.$$

And of course, $\dfrac{\partial u}{\partial x}$ and $\dfrac{\partial u}{\partial y}$ will both, in general, be functions of x and y.

However, sometimes, we may meet the converse problem. Given $P(x,y)$ and $Q(x,y)$ in the expression called a differential form

$$P(x,y)\mathrm{d}x+Q(x,y)\mathrm{d}y, \tag{4.24}$$

when is it the case that this is the total differential of some function $u(x,y)$?

Clearly, if it is such, then $P(x,y)=\dfrac{\partial u}{\partial x}$ and $Q(x,y)=\dfrac{\partial u}{\partial y}$ for the function $u(x,y)$. This implies and is implied by the **Condition of Integrability**:

$$\frac{\partial P}{\partial y}\equiv\frac{\partial Q}{\partial x}\left(\equiv\frac{\partial^2 u}{\partial x\partial y}\right). \tag{4.25}$$

For example, look at the differential form $y^2\mathrm{d}x+\left(x^2+2y\right)\mathrm{d}y$. In this case, $P\equiv y^2$, and $Q\equiv x^2+2y$. So, we can see that

$$\frac{\partial P}{\partial y}=2y\neq\frac{\partial Q}{\partial x}=2x.$$

Therefore, this is not an exact differential.

Also, consider $(2xy + \cos x \cos y)\,\mathrm{d}x + \left(x^2 - \sin x \sin y\right)\mathrm{d}y$. In this case,

$$\frac{\partial P}{\partial y} = 2x - \cos x \sin y = \frac{\partial Q}{\partial x}.$$

And therefore, this is an exact differential.

Sometimes we would like to find out $u(x, y)$ from the differential form $P(x, y)\mathrm{d}x + Q(x, y)\mathrm{d}y$. If this is an exact differential, we have:

$$\begin{cases} \dfrac{\partial u}{\partial x} \equiv P(x, y) \\ \dfrac{\partial u}{\partial y} \equiv Q(x, y). \end{cases} \tag{4.26}$$

Therefore, we could first find out an expression for $u(x, y)$ by integrating $P(x, y)$ with respect to x:

$$u(x, y) = \int P(x, y)\mathrm{d}x. \tag{4.27}$$

However, pay attention to the "constant" in this case. As $u(x, y)$ is a function of two variables x and y, taking the partial derivative with respect to x would yield a zero for any function that only contains y. That is to say, the "constant" from this indefinite integral is a function of y, $f(y)$.

Look at the example above, where we have the differential form $(2xy + \cos x \cos y)\,\mathrm{d}x + \left(x^2 - \sin x \sin y\right)\mathrm{d}y$:

$$\begin{aligned} \frac{\partial u}{\partial x} &\equiv 2xy + \cos x \cos y \\ u(x, y) &\equiv \int (2xy + \cos x \cos y)\mathrm{d}x \\ &= x^2 y + \sin x \cos y + f(y). \end{aligned}$$

To find out $f(y)$, differentiate this function with respect to y:

$$\frac{\partial u}{\partial y} = x^2 - \sin x \sin y + f'(y).$$

Then, compare this expression with the function $Q(x, y) \equiv \dfrac{\partial u}{\partial y}$:

$$\frac{\partial u}{\partial y} = x^2 - \sin x \sin y + f'(y) \equiv Q(x, y) = x^2 - \sin x \sin y.$$

This means that $f(y) = C$, where C is a constant. Therefore, we can conclude that

$$u(x, y) = x^2 y + \sin x \cos y + C.$$

We can alternatively integrate $Q(x, y)$ with respect to y and get a function $g(x)$ to be determined. These two methods are exactly the same.

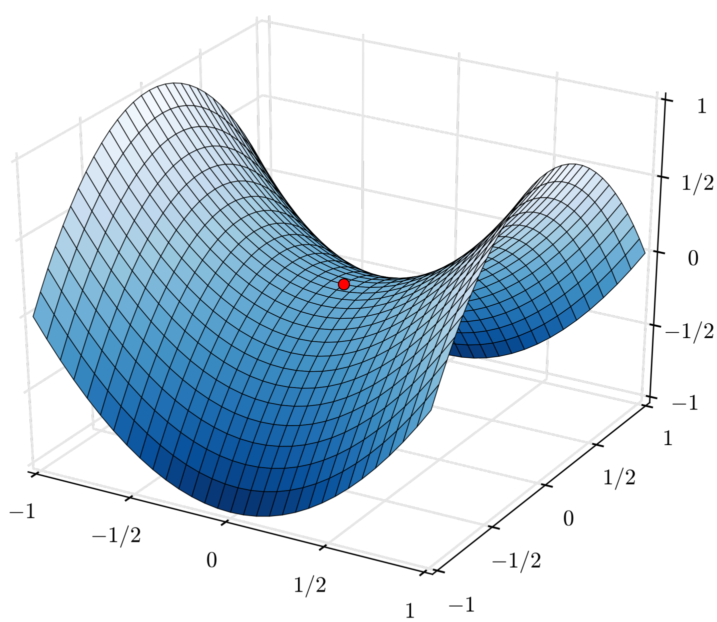
4.7 Stationary Points

We have already seen the stationary points for functions of one independent variable. For functions that involve two independent variables, there are three types of stationary points:

(1) Local Maximum ($u(x, y)$ decreasing away from the point);

(2) Local Minimum ($u(x, y)$ increasing away from the point);

(3) Saddle Point ($u(x, y)$ decreasing or increasing depending on direction).



For functions of two variables $u(x, y)$, stationary points are located at simultaneous solution of the two equations:

$$\mathrm{d}u = 0 \begin{cases} \dfrac{\partial u}{\partial x} = 0 \\ \dfrac{\partial u}{\partial y} = 0. \end{cases} \tag{4.28}$$

For a function of only one independent variable, a stationary point means a horizontal tangent line. Then, for a function of two independent variables, a stationary point means a horizontal tangent plane.

A stationary point (x_0, y_0) has character determined by the expression

$$E_0 \equiv \left(\frac{\partial^2 u}{\partial x \partial y}\right) - \left(\frac{\partial^2 u}{\partial x^2}\right)\left(\frac{\partial^2 u}{\partial y^2}\right)\bigg|_{(x_0,y_0)} \equiv B^2 - AC. \tag{4.29}$$

$$\begin{cases} E_0 > 0 & \text{Saddle Point} \\ E_0 < 0 \begin{cases} A < 0 & \text{Local Maximum} \\ A < 0 & \text{Local Minimum} \end{cases} \end{cases} \tag{4.30}$$

If $E_0 = 0$, higher order derivatives determine the issue.

When we are faced with a function of several variables and we need to find stationary points (and potential local maximum and minimum), we need to make sure that our independent variables are truly independent. For example, when maximising the volume V of a rectangular box given the surface area A fixed, the lengths of the three sides are not really independent:

$$\max V = xyz \text{ given that } A = 2xy + 2yz + 2xz \text{ is fixed.}$$

Here the three variables x, y, and z are constrained by the surface area A, making them not independent from each other. This is an example of conditional extremum.

4.7.1 Conditional Extremum

When we are trying to determine the maximum or minimum of a function $f(x_1, \ldots, x_n)$ under the condition that certain other functions assume given values like $\varphi(x_1, \ldots, x_n) = 0$, we are trying to get a conditional extremum.

For conditional extremum, we apply a method called **The Method of Lagrange Multiplier**. When finding the extreme value of a function $f(x_1, \ldots, x_n)$ under the constraint $\varphi(x_1, \ldots, x_n) = 0$, create a Lagrange function:

$$\mathcal{L} = f(x_1, \ldots, x_n) + \lambda\varphi(x_1, \ldots, x_n), \tag{4.31}$$

where λ is called the Lagrange multiplier. The set of stationary points of this Lagrange function would include the conditional extremum.

When finding the extreme value for a function $f(x_1, \ldots, x_n)$ under $m(m < n)$ conditions, we write:

$$\mathcal{L} = f(x_1, \ldots, x_n) + \sum_{i=1}^{m} \lambda_i \varphi_i(x_1, \ldots, x_n). \tag{4.32}$$

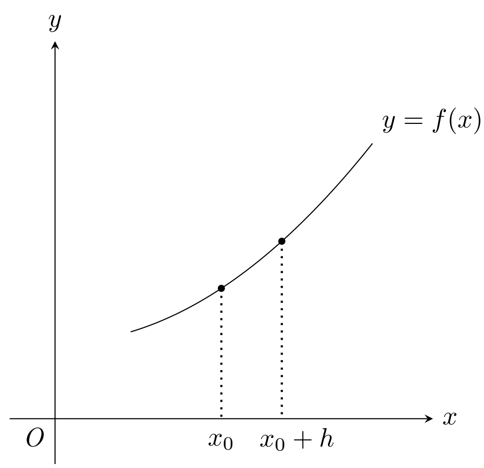
Chapter 5

Series

5.1 Taylor and Maclaurin Series

5.1.1 Taylor Series

Consider a function $f(x)$ of a single independent variable, where the value of all its derivatives are known at a point x_0. Try to evaluate the function at a nearby point $x_0 + h$, using a power series: (h is small)



$$f(x_0 + h) = \alpha_0 + \alpha_1 h + \alpha_2 h^2 + \cdots = \sum_{n=0}^{\infty} \alpha_n h^n,$$

with coefficients α_n to be determined.

Putting $h = 0$ into the above series, we get that $f(x_0) = \alpha_0$.

Then, differentiating with respect to h, we get:

$$f'(x_0 + h) = \alpha_1 + 2\alpha_2 h + 3\alpha_3 h^2 + \ldots$$

. Therefore, $h = 0 \Rightarrow f'(x_0) = \alpha_1$.

Continuing in this manner, we can easily prove that the general result is:

$$\alpha_n = \frac{f^{(n)}(x_0)}{n!}. \tag{5.1}$$

And, we obtain the **Taylor Series** of $f(x)$ about x_0 in the form

$$f(x_0 + h) = \sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!} h^n. \tag{5.2}$$

There is also an alternative expression:

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!} (x - x_0)^n. \tag{5.3}$$

The above is not a formal proof, since it does rely on various assumptions about e.g. differentiating an infinite series. For now, just consider the successive terms of the Taylor series as improvements to higher and higher order to a local polynomial approximation.

Normally, we may not achieve to Taylor expand a function and evaluate it to an infinite number of terms. What we usually do is to simply choose the first n terms. Therefore, we should consider the error (the remainder) by this approximation.

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2!}(x - x_0)^2 + \cdots + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n + R_n,$$

$$\text{with } R_n = \frac{f^{(n+1)}(\xi)}{n!}(x - x_0)^{n+1} \text{ with } x_0 < \xi < x. \tag{5.4}$$

The utility of this result depends on:

(1) $f(x)$ having the requisite number of derivatives;

(2) the behaviour of the remainder term as n increases.

5.1.2 Maclaurin Series

Simply put, the Maclaurin series of a function $f(x)$ is the Taylor series about the origin. Therefore, we can see that:

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n, \tag{5.5}$$

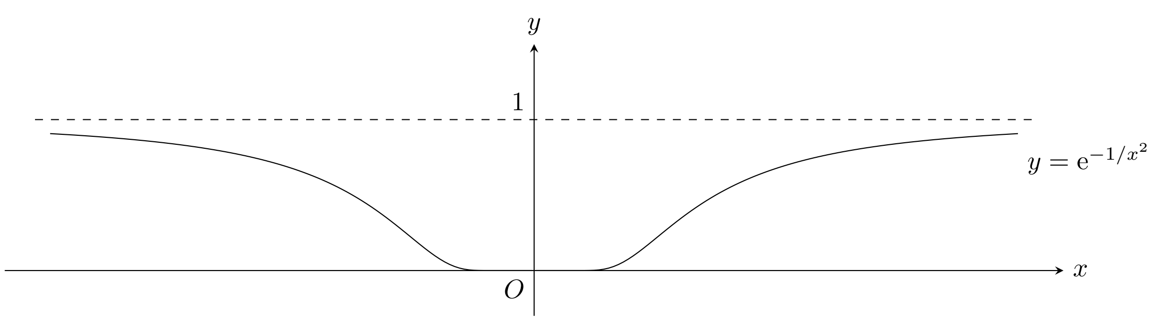
with the remarks about truncation, remainder terms for practical utility.

Here come some common examples of functions and their Maclaurin series.

$$
\begin{aligned}
e^x &= 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots \equiv \sum_{n=0}^{\infty} \frac{x^n}{n!} \\
\sin x &= x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots \equiv \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!} \\
\cos x &= 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots \equiv \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!} \\
\frac{1}{1-x} &= 1 + x + x^2 + \cdots \equiv \sum_{n=0}^{\infty} x^n \qquad \text{where } |x| < 1 \\
\ln(1+x) &= x - \frac{x^2}{2} + \frac{x^3}{3} - \cdots \equiv \sum_{n=0}^{\infty} (-1)^{n-1} \frac{x^n}{n} \qquad \text{where } -1 < x \leq 1 \\
\arctan x &= x - \frac{x^3}{3} + \frac{x^5}{5} - \cdots \equiv \sum_{n=0}^{\infty} (-1)^{n-1} \frac{x^{2n-1}}{2n-1} \qquad \text{where } |x| \leq 1.
\end{aligned}
\tag{5.6}
$$

It is important to note that some of the Maclaurin series are useful only in an interval of x. Also, we normally can rely on our expansion parameter being small enough for our expansion to be useful, but **the remainder can contain more information than we might anticipate**.

e.g. $f(x) = e^{-1/x^2}$



Here we can show that $f^{(n)}(0) = 0$, so

$$
f(x) = 0 + 0 \cdot x + 0 \cdot x^2 + \cdots + 0 \cdot x^n + R_n
$$

is the Maclaurin expansion. This function is very flat at $x = 0$ and is contained wholly in R_n.

5.2 L Hôpital's Rule (Revisited)

When we consider the limit

$$\lim_{x\to x_0}\frac{f(x)}{g(x)} \text{ with } \lim_{x\to x_0} f(x)=0 \text{ and } \lim_{x\to x_0} g(x)=0,$$

we used L Hôpital's Rule.

To find the justification for this rule, we can assume that $f(x)$ and $g(x)$ each have a Taylor series expansion in the neighborhood of x_0. Then, we can write the above limit as:

$$\begin{aligned}\lim_{x\to x_0}\frac{f(x)}{g(x)} &\equiv \lim_{h\to 0}\frac{f(x_0+h)}{g(x_0+h)}\\ &= \lim_{h\to 0}\frac{f(x_0)+hf'(x_0)+\frac{h^2}{2!}f''(x_0)+\dots}{g(x_0)+hg'(x_0)+\frac{h^2}{2!}g''(x_0)+\dots}.\end{aligned}$$

For $\lim_{x\to x_0} f(x)=f(x_0)=0$ and $\lim_{x\to x_0} g(x)=g(x_0)=0$,

$$\begin{aligned}\lim_{x\to x_0}\frac{f(x)}{g(x)} &\equiv \lim_{h\to 0}\frac{f'(x_0)+\frac{h}{2}f''(x_0)+\dots}{g'(x_0)+\frac{h}{2}g''(x_0)+\dots}\\ &= \frac{f'(x_0)}{g'(x_0)} \quad \text{if at least one of numerator and denominator is nonzero.}\end{aligned}$$

If $f'(x_0)=g'(x_0)=0$, then we go to the next terms and get $\dfrac{f''(x_0)}{g''(x_0)}$, and so on.

5.3 Double Taylor Series

We now consider a function $f(x,y)$ of two independent variables x and y in the neighborhood of (x_0,y_0). That is we seek an expansion in powers of $h=x-x_0$ and $k=y-y_0$.

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

where

$$\left(h\frac{\partial}{\partial x}+k\frac{\partial}{\partial y}\right)f(x_0,y_0) = h\left.\frac{\partial f}{\partial x}\right|_{(x_0,y_0)} + k\left.\frac{\partial f}{\partial y}\right|_{(x_0,y_0)},$$

and

$$\left(h\frac{\partial}{\partial x}+k\frac{\partial}{\partial y}\right)^m f(x_0,y_0) = \sum_{p=0}^{m}\binom{m}{p}h^p k^{m-p}\left.\frac{\partial^m f}{\partial x^p\partial y^{m-p}}\right|_{(x_0,y_0)}.$$

Note that we are assuming that $\dfrac{\partial^2 u}{\partial x \partial y} = \dfrac{\partial^2 u}{\partial y \partial x}$ etc.

We can write $D \equiv h\dfrac{\partial}{\partial x} + k\dfrac{\partial}{\partial y}$. Therefore, by writing $f(x_0, y_0) = f_0$,

$$f(x_0 + h, y_0 + k) = f_0 + Df_0 + \frac{D^2}{2!}f_0 + \cdots + \frac{D^n}{n!}f_0 + \ldots. \tag{5.7}$$

5.4 Stationary Points (Revisited)

We considered stationary points of a function $u(x, y)$ of two independent variables. We are now in a position to justify the criterion used there to determine their character.

Consider a stationary point (x_0, y_0) where we have $\left(\dfrac{\partial u}{\partial x}\right)_0 = \left(\dfrac{\partial u}{\partial y}\right)_0 = 0$.

Write our Taylor expansion for $u(x, y)$ about (x_0, y_0):

$$u(x_0 + h, y_0 + k) = u_0 + h\left(\frac{\partial u}{\partial x}\right)_0 + k\left(\frac{\partial u}{\partial y}\right)_0 + \frac{1}{2}\left(Ah^2 + 2Bhk + Ck^2\right) + \ldots,$$

where $A = \left(\dfrac{\partial^2 u}{\partial x^2}\right)_0$, $B = \left(\dfrac{\partial^2 u}{\partial x \partial y}\right)_0$, and $C = \left(\dfrac{\partial^2 u}{\partial y^2}\right)_0$. With $\left(\dfrac{\partial u}{\partial x}\right)_0 = \left(\dfrac{\partial u}{\partial y}\right)_0 = 0$,

$$u(x_0 + h, y_0 + k) = u(x_0, y_0) + \frac{1}{2}\left(Ah^2 + 2Bhk + Ck^2\right) + \ldots,$$

and

$$\begin{aligned}\delta u &= u(x_0 + h, y_0 + k) - u(x_0, y_0) \\ &= \frac{1}{2}\left(Ah^2 + 2Bhk + Ck^2\right) + \ldots.\end{aligned}$$

Evidently,

$$\begin{cases} \delta u > 0 \text{ for ANY small } h,\ k & u(x_0, y_0) \text{ is a local minimum;} \\ \delta u < 0 \text{ for ANY small } h,\ k & u(x_0, y_0) \text{ is a local maximum;} \\ \begin{cases} \delta u > 0 \text{ for some } h,\ k \\ \delta u < 0 \text{ for some } h,\ k \end{cases} & u(x_0, y_0) \text{ is a saddle point.} \end{cases} \tag{5.8}$$

We can arrange the expression for δu,

$$\delta u = \frac{1}{2}k^2\left[A\left(\frac{h}{k}\right)^2 + 2B\left(\frac{h}{k}\right) + C\right] + \ldots,$$

and consider $F(\lambda) \equiv A\lambda^2 + 2B\lambda + C$. We can see that

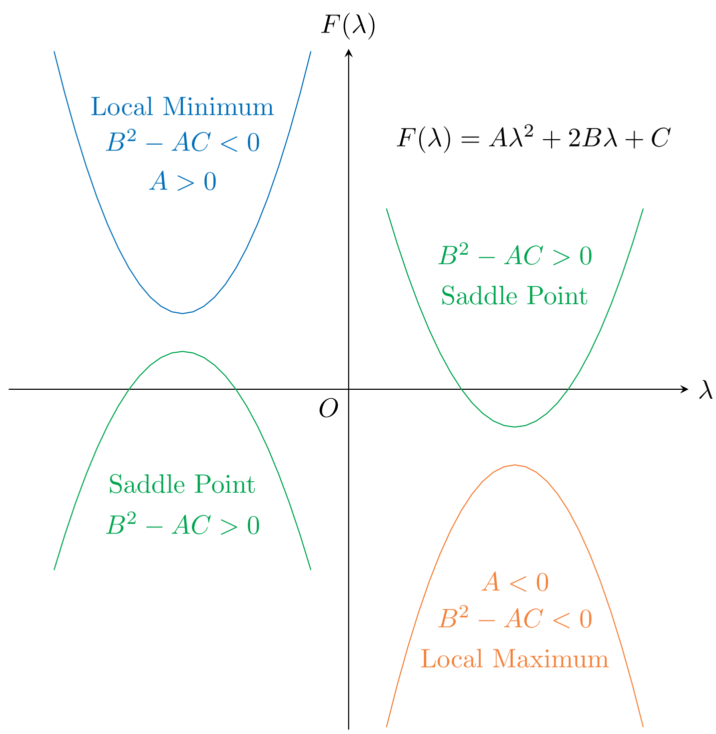
$$\delta u = \frac{1}{2}k^2 F\left(\frac{h}{k}\right),$$

and δu has the same sign as that of $F\left(\frac{h}{k}\right)$.

We can see that the character of the stationary point is largely determined by the values of A, B, and C. As $F(\lambda) \equiv A\lambda^2 + 2B\lambda + C$, the discriminant is:

$$\Delta = (2B)^2 - 4AC = 4(B^2 - AC).$$

Both local maximum and local minimum indicate no real roots, so we say $B^2 - AC < 0$. Also, as A is the coefficient of the dominant λ^2, the sign of A determines if the stationary point is a local maximum or a local minimum. That is to say, if $A > 0$, the stationary point is a local minimum; if $A < 0$, the point is a local maximum. When $B^2 - AC > 0$, there are roots on the λ axis, which manifests that the sign of $F(\lambda)$ depends on the value of λ. So, this is a saddle point.



5.5 Infinite Sums

Consider the infinite sum

$$S = \sum_{n=0}^{\infty} u_n \tag{5.9}$$

and its meaning. To understand it, we consider a truncation

$$S_N = \sum_{n=0}^{\infty} u_n$$

and examine whether $\lim_{N\to\infty} S_N$ exists; that is to say, we examine if this limit has a finite value. If so, we say the series S converges. If not, then the series diverges, and the infinite sum has no meaning.

It is important to note that divergence can involve $|S_N|$ increasing without bound as $N \to \infty$ or that $|S_N|$ just does not approach finite limit. For instance, it could just oscillate. Here come some examples of series.

(1) The series $\sum_{n=0}^{\infty}(-1)^n$ diverges, with $S_0 = 1$, $S_1 = 0$, $S_2 = 1$, etc.

(2) The **geometric series**:

Consider the finite sum of N terms:

$$S_N = 1 + x + x^2 + \cdots + x^N. \tag{5.10}$$

Evidently,

$$xS_N = x + x^2 + x^3 + \cdots + x^{N+1},$$

which gives

$$S_N(1 - x) = 1 - x^{N+1}.$$

Therefore,

$$S_N = \sum_{n=0}^{N} x^n = \frac{1 - x^{N+1}}{1 - x}, \tag{5.11}$$

and the limit $S = \lim_{N\to\infty} S_N$ exists only for $|x| < 1$. For $|x| \geq 1$ the series diverges. This is also the reason why the Maclaurin series expansion of $f(x) = \dfrac{1}{1-x}$ is only valid for $|x| < 1$.

Generally, we represent a geometric series by

$$S = \sum ar^{n-1}. \tag{5.12}$$

For a geometric series, when $|r| < 1$,

$$S_n = a \cdot \frac{1 - r^n}{1 - r},$$

and

$$\sum ar^{n-1} = \lim_{n\to\infty} S_n = a \cdot \frac{1}{1 - r}. \tag{5.13}$$

(3) The **p-series** and **harmonic series**:

We define $\sum \frac{1}{n^p}$ to be the p series. When $p = 1$, this yields the harmonic series:

$$S = \sum_{n=1}^{\infty} \frac{1}{n}. \tag{5.14}$$

This series seems to converge to a finite value, but actually it diverges.

Proof. We can combine the terms of the harmonic series as follows:

$$S = 1 + \frac{1}{2} + \left(\frac{1}{3} + \frac{1}{4}\right) + \left(\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8}\right) + \dots.$$

Every pair of brackets contains a value that is clearly larger than $\frac{1}{2}$, and:

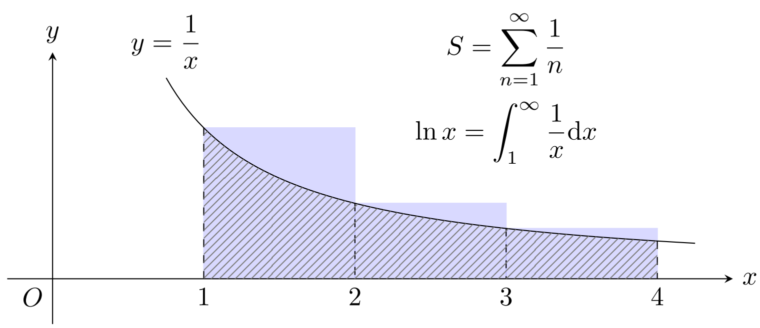
$$\begin{aligned} S &= 1 + \frac{1}{2} + \left(\frac{1}{3} + \frac{1}{4}\right) + \left(\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8}\right) + \dots \\ &\geq 1 + \frac{1}{2} + \left(\frac{1}{4} + \frac{1}{4}\right) + \left(\frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8}\right) + \dots \\ &= 1 + \frac{1}{2} + \frac{1}{2} + \frac{1}{2} + \dots. \end{aligned}$$

Therefore, as the number of terms increases to infinity, this series diverges. □

Recall our definition of the natural logarithm:

$$\ln x = \int_1^{\infty} \frac{1}{x} \mathrm{d}x.$$

This comes in a similar manner, but in a slightly different way. If we are to represent the harmonic sum, we should draw:



Clearly, The harmonic series is a left Riemann sum of the integral $\int_1^{\infty} \frac{1}{x} \mathrm{d}x$, with step sizes of 1. And, in fact, the difference between the harmonic series and $\ln N$, as N approaches infinity, is the **Euler-Mascheroni**

Constant:

$$\gamma = \lim_{N\to\infty}\left(\sum_{n=1}^{N}\frac{1}{n} - \ln N\right) = 0.577215\ldots. \tag{5.15}$$

This constant is not even known to be irrational. If it is a rational number, then the denominator is over 10^{244663}.

5.6 More Definitions and Theorems

An infinite series $S = \sum_{n=0}^{\infty} u_n$ is said to be **absolutely convergent** if

$$\sum_{n=0}^{\infty} |u_n| \tag{5.16}$$

is convergent. This is useful because absolute convergence indicates convergence. We can examine the absolute convergence when some terms are negative or complex.

For example, consider the series

$$S_1 = \frac{1}{2} - \frac{1}{4} + \frac{1}{8} - \frac{1}{16} + \ldots.$$

This series is absolutely convergent, because

$$\begin{aligned} S_2 &= \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \ldots \\ &= \frac{1/2}{1-1/2} \\ &= 1. \end{aligned}$$

Therefore, convergence of S_2 indicates the convergence of S_1. In fact,

$$\begin{aligned} S_1 &= \frac{1}{2} - \frac{1}{4} + \frac{1}{8} - \frac{1}{16} + \ldots \\ &= \frac{1/2}{1-(-1/2)} \\ &= \frac{1}{3}. \end{aligned}$$

If $\sum_{n=0}^{\infty} u_n$ converges and $\sum_{n=0}^{\infty} |u_n|$ does not, then we say $\sum_{n=0}^{\infty} u_n$ is **conditionally convergent**.

For example, $S = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \ldots$ is conditionally convergent, since it converges (to $\ln 2$), but the series $S = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \ldots$ diverges.

5.7 Convergence Tests

5.7.1 Convergence Theorems of Infinite Series

(1) If $\sum a_n$ converges, then $\lim_{n\to\infty} a_n = 0$.

It is **necessary but not sufficient** for convergence that

$$\lim_{n\to\infty} a_n = 0.$$

If this limit is not zero, then this series diverges; if it is zero, then the series may still diverge (the harmonic series).

$$\begin{cases} \text{If } \sum a_n \text{ converges, then } \lim_{n\to\infty} a_n = 0. & \text{Correct} \\ \text{If } \lim_{n\to\infty} a_n = 0, \text{ then } \sum a_n \text{ converges.} & \text{Wrong} \\ \text{If } \sum a_n \text{ diverges, then } \lim_{n\to\infty} a_n \neq 0. & \text{Wrong} \\ \text{If } \lim_{n\to\infty} a_n \neq 0, \text{ then } \sum a_n \text{ diverges.} & \text{Correct} \end{cases} \tag{5.17}$$

That is to say, if a series converges, then the limit of the n-th term of this series must be zero; nevertheless, $\lim_{n\to\infty} a_n = 0$ does not necessarily indicate that the series converges.

If we are to determine if a series converges, we can examine first if $\lim_{n\to\infty} a_n = 0$. If it does not, then the series must diverge; if it is, then we continue the discussion.

(2) $\sum_{n=1}^{\infty} a_n$ and $\sum_{n=m}^{\infty} a_n$ both converge or diverge.

(3) $\sum a_n$ and $\sum (c \cdot a_n)$ $(c \neq 0)$ both converge or diverge.

5.7.2 Tests for Infinite Series

The n-th Term Test

If $\lim_{n\to\infty} a_n \neq 0$, then $\sum a_n$ diverges.

The Geometric Series Test

A geometric series $\sum ar^{n-1}$ converges if and only if $|r| < 1$.

The Comparison Test

If a_n is given and we can find a converging series $\sum_{n=0}^{\infty} b_n$ with non-negative b_n such that $|a_n| \leq b_n$ for all n, then $\sum a_n$ is absolutely convergent.

Similarly, if we can find a diverging series $\sum_{n=0}^{\infty} b_n$ with non-negative b_n such that $a_n \geq b_n$ for all n, then $\sum a_n$ is divergent.

For example, $\sum_{n=1}^{\infty} \frac{1}{\sqrt{n}}$ is divergent because $\frac{1}{\sqrt{n}} \geq \frac{1}{n}$ for all n and we know $\sum_{n=1}^{\infty} \frac{1}{n}$ is divergent.

5.7.3 Tests for Positive Infinite Series

The Integral Test

$\sum a_n$ and $\int_1^{\infty} f(x)\mathrm{d}x$ both converge or diverge, where $a_n = f(n)$.

The P-series Test

A p-series $\sum \frac{1}{n^p}$ converges when $p > 1$, but diverges when $p \leq 1$.

The Limit Comparison Test

If $\lim_{n\to\infty} \frac{a_n}{b_n} = L$, where $0 < L < \infty$, then $\sum a_n$ and $\sum b_n$ both converge or diverge.
If $\lim_{n\to\infty} \frac{a_n}{b_n} = 0$, and $\sum b_n$ converges, then $\sum a_n$ converges.
If $\lim_{n\to\infty} \frac{a_n}{b_n} = \infty$, and $\sum a_n$ diverges, then $\sum b_n$ diverges.

The Ratio Test

For $\sum_{n=0}^{\infty} u_n$ we suppose $u_n \neq 0$ for all n. Define

$$L = \lim_{n\to\infty} \left| \frac{u_{n+1}}{u_n} \right|, \tag{5.18}$$

then:

$$\begin{cases} L < 1 & \text{absolute convergence} \Rightarrow \text{convergence;} \\ L > 1 & \text{divergence;} \\ L = 1 & \text{inconclusive.} \end{cases}$$

The Root Test

$$\lim_{n\to\infty} \sqrt[n]{a_n} = L \begin{cases} L < 1 & \text{converge} \\ L = 1 & \text{inconclusive} \\ L > 1 & \text{diverge.} \end{cases} \tag{5.19}$$

5.7.4 Alternating Series - Leibniz Test

If we have $\sum_{n=0}^{\infty}(-1)^n a_n$ with

(1) positive a_n,

(2) a_n decreasing, and

(3) $\lim_{n\to\infty} a_n = 0$,

then our series converges.

For example, the series $1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \dots$ is convergent. (Its sum is actually $\arctan(1) = \frac{1}{4}$.)

A natural question to ask is whether condition (2) us actually needed when condition (3) is satisfied. To prove that condition (2) is necessary, consider the alternating series

$$S = \frac{1}{\sqrt{2}-1} - \frac{1}{\sqrt{2}+1} + \frac{1}{\sqrt{3}-1} - \frac{1}{\sqrt{3}+1} + \frac{1}{\sqrt{4}-1} - \frac{1}{\sqrt{4}+1} + \dots.$$

It is easy to see that conditions (1) and (3) are satisfied, but the condition (2) is not. Therefore, this series fails for the alternating series test. By combining the series, we can see that

$$\begin{aligned}
S &= \left(\frac{1}{\sqrt{2}-1} - \frac{1}{\sqrt{2}+1}\right) + \left(\frac{1}{\sqrt{3}-1} - \frac{1}{\sqrt{3}+1}\right) + \left(\frac{1}{\sqrt{4}-1} - \frac{1}{\sqrt{4}+1}\right) + \dots \\
&= \frac{2}{2-1} + \frac{2}{3-1} + \frac{2}{4-1} + \dots \\
&= 2\left(1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \dots\right),
\end{aligned}$$

and that S diverges. Therefore, the condition (2) is also needed in the alternating series test.

5.7.5 Radius of Convergence of Taylor / Maclaurin Series

The ratio test allows us to determine the range of convergence of Taylor / Maclaurin series.

Consider $f(x) = \mathrm{e}^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$.

We can see that

$$L = \lim_{n\to\infty}\left|\frac{u_{n+1}}{u_n}\right| = \lim_{n\to\infty}\left|\frac{x^{n+1}/(n+1)!}{x^n/n!}\right| = \lim_{n\to\infty}\frac{|x|}{n+1} = 0 < 1$$

for all fixed x. Therefore, the Maclaurin series for e^x converges for all x (both real and complex).

Within the 'radius of convergence' we have convergence. However, at that value we do not know if the series converges. For instance, consider the series

$$\sum_{n=1}^{\infty}\frac{x^n}{n} = x + \frac{x^2}{2} + \frac{x^3}{3} + \dots.$$

Using the ratio test,

$$L = \lim_{n\to\infty} \left| \frac{x^{n+1}}{n+1} \cdot \frac{n}{x^n} \right| = |x|.$$

Therefore, we know that the series

$$\begin{cases} \text{converges} & \text{if } |x| < 1 \\ \text{diverges} & \text{if } |x| > 1. \end{cases}$$

At $x = 1$ we have $1 + \frac{1}{2} + \frac{1}{3} + \ldots$. This is the harmonic series, and we know that it diverges. At $x = -1$ we have $-1 + \frac{1}{2} - \frac{1}{3} + \cdots \equiv -\ln 2$. This series converges.

Furthermore, we use the word "radius" here to describe a circle of convergence in the complex plane.

It is important to note that a function and its derivative have the same radius of convergence, but their interval of convergence can be different. That is to say, they may not both converge or diverge on the endpoints.

5.7.6 The Mysterious Zeta Function

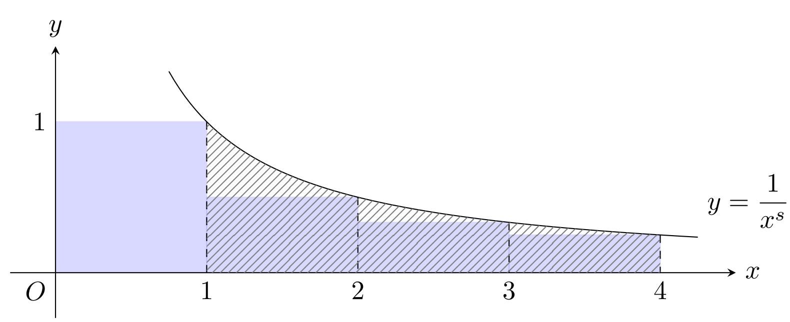
The Zeta function of Riemann is the function

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}, \tag{5.20}$$

which has attained a mysterious, indeed almost mythical, status.

$\zeta(1)$ is of course the harmonic series, which diverges.

$\zeta(s)$ can be shown to converge if $\mathrm{Re}(s) > 1$. This function is bounded by an integral. We can plot this infinite sum as follows:



The 'tail' (terms excluding the first one) is bounded above by an integral that is convergent when $\mathrm{Re}(s) > 1$. That is to say,

$$\frac{1}{2^s} + \frac{1}{3^s} + \cdots < \int_1^{\infty} \frac{1}{x^s} \mathrm{d}x = \frac{1}{s-1},$$

and

$$\zeta(s) < u_1 + \frac{1}{s-1} = 1 + \frac{1}{s-1}.$$

This argument works for complex s too.

We can find that

$$\zeta(2) = \frac{\pi^2}{6}$$
$$\zeta(3) = 1.2020569\ldots \text{ (irrational)}$$
$$\zeta(4) = \frac{\pi^4}{90}.$$

The Riemann hypothesis is that all non-trivial zeros of the ζ function have real part $\frac{1}{2}$. That is to say, $s_0 = \frac{1}{2} + it$. This is known as the critical line in the complex s plane. A very large number have been computed, but so far there is no proof for this.